Lecture 7: Approximate Border Bases

— The Hybrid Lecture —

Martin Kreuzer

Fakultät für Informatik und Mathematik

Universität Passau

martin.kreuzer@uni-passau.de

Sophus Lie Center

Nordfjordeid

June 19, 2009

Contents

Contents

Contents

Contents

Contents

Contents

Contents

1 – Approximate Data and Polynomials

Ralph W. Emerson:

I hate quotations.

1 – Approximate Data and Polynomials

Ralph W. Emerson:

I hate quotations.

Tell me what you know.

1 – Approximate Data and Polynomials

Ralph W. Emerson:

I hate quotations.

Tell me what you know.

$P = \mathbb{R}[x_1, \ldots, x_n]$ polynomial ring over the real number field

$\mathbb{X} = \{p_1, \ldots, p_s\}$ finite set of points in $\mathbb{R}^n$

1 – Approximate Data and Polynomials

Ralph W. Emerson:
I hate quotations.
Tell me what you know.

$P = \mathbb{R}[x_1, \ldots, x_n]$ polynomial ring over the real number field

$\mathbb{X} = \{p_1, \ldots, p_s\}$ finite set of points in $\mathbb{R}^n$

Recall that the map $\mathrm{eval} : P \longrightarrow \mathbb{R}^s$ given by $f \mapsto (f(p_1), \ldots, f(p_s))$ is called the **evaluation map** associated to $\mathbb{X}$.

The ideal $I_{\mathbb{X}} = \ker(\mathrm{eval})$ is called the **vanishing ideal** of $\mathbb{X}$.

1 – Approximate Data and Polynomials

Ralph W. Emerson:

I hate quotations.

Tell me what you know.

$P = \mathbb{R}[x_1, \ldots, x_n]$ polynomial ring over the real number field

$\mathbb{X} = \{p_1, \ldots, p_s\}$ finite set of points in $\mathbb{R}^n$

Recall that the map $\mathrm{eval} : P \longrightarrow \mathbb{R}^s$ given by $f \mapsto (f(p_1), \ldots, f(p_s))$ is called the **evaluation map** associated to $\mathbb{X}$.

The ideal $I_{\mathbb{X}} = \ker(\mathrm{eval})$ is called the **vanishing ideal** of $\mathbb{X}$.

The Gretchen Question:

1 – Approximate Data and Polynomials

Ralph W. Emerson:
I hate quotations.
Tell me what you know.

$P = \mathbb{R}[x_1, \ldots, x_n]$ polynomial ring over the real number field

$\mathbb{X} = \{p_1, \ldots, p_s\}$ finite set of points in $\mathbb{R}^n$

Recall that the map $\mathrm{eval} : P \longrightarrow \mathbb{R}^s$ given by $f \mapsto (f(p_1), \ldots, f(p_s))$ is called the **evaluation map** associated to $\mathbb{X}$.

The ideal $I_{\mathbb{X}} = \ker(\mathrm{eval})$ is called the **vanishing ideal** of $\mathbb{X}$.

The Gretchen Question: What happens if the points of $\mathbb{X}$ are only **empirical points**, e.g. points whose coordinates are derived from measured data?

In the following we let $\varepsilon > 0$ be a given **threshold number**.

In the following we let $\varepsilon > 0$ be a given **threshold number**.

Definition 1.1 A polynomial $f \in P$ is said to **vanish ε-approximately** at a point $p \in \mathbb{R}^n$ if $|f(p)| < \varepsilon$.

In the following we let $\varepsilon > 0$ be a given **threshold number**.

Definition 1.1 A polynomial $f \in P$ is said to **vanish ε-approximately** at a point $p \in \mathbb{R}^n$ if $|f(p)| < \varepsilon$.

Problem: The polynomials which vanish ε-approximately at $\mathbb{X}$ **do not form an ideal**.

For instance, if $|f(p)| = 0.001 < \varepsilon = 0.1$ then $|(1000\, f)(p)| = 1 > \varepsilon$.

Hence the question whether f vanishes at p or not depends on the **size of f**, i.e. we need a **metric** on P.

Definition 1.2 Let $f = a_1 t_1 + \cdots + a_s t_s \in P$, where $a_1, \ldots, a_s \in \mathbb{R} \setminus \{0\}$ and $t_1, \ldots, t_s \in \mathbb{T}^n$.

Then the number $\|f\| = \|(a_1, \ldots, a_s)\|$ is called the (Euclidean) **norm** of f.

Clearly, this definition turns P into a normed vector space.

Definition 1.2 Let $f = a_1 t_1 + \cdots + a_s t_s \in P$, where $a_1, \ldots, a_s \in \mathbb{R} \setminus \{0\}$ and $t_1, \ldots, t_s \in \mathbb{T}^n$.

Then the number $\|f\| = \|(a_1, \ldots, a_s)\|$ is called the (Euclidean) **norm** of f.

Clearly, this definition turns P into a normed vector space.

Reformulated Problem: A **small** polynomial **always** vanishes ε-approximately at $\mathbb{X}$!

Definition 1.2 Let $f = a_1 t_1 + \cdots + a_s t_s \in P$, where $a_1, \ldots, a_s \in \mathbb{R} \setminus \{0\}$ and $t_1, \ldots, t_s \in \mathbb{T}^n$.

Then the number $\|f\| = \|(a_1, \ldots, a_s)\|$ is called the (Euclidean) **norm** of f.

Clearly, this definition turns P into a normed vector space.

Reformulated Problem: A **small** polynomial **always** vanishes ε-approximately at $\mathbb{X}$!

Henceforth we use the condition that polynomials $f \in P$ with $\|f\| = 1$ vanish ε-approximately at p.

2 – Approximate Border Bases

It's kind of fun to do the impossible.
(Walt Disney)

Definition 2.1 An ideal $I \subseteq P$ is called an **ε-approximate vanishing ideal** of $\mathbb{X}$ if there exists a system of generators $\{f_1, \ldots, f_r\}$ of I such that $\|f_i\| = 1$ and f_i vanishes ε-approximately at $\mathbb{X}$ for $i = 1, \ldots, r$.

2 – Approximate Border Bases

It's kind of fun to do the impossible.
(Walt Disney)

Definition 2.1 An ideal $I \subseteq P$ is called an **ε-approximate vanishing ideal** of $\mathbb{X}$ if there exists a system of generators $\{f_1, \ldots, f_r\}$ of I such that $\|f_i\| = 1$ and f_i vanishes ε-approximately at $\mathbb{X}$ for $i = 1, \ldots, r$.

Problems: (1) Approximate vanishing ideals are **not at all unique**. They are not necessarily zero-dimensional either!

2 – Approximate Border Bases

It's kind of fun to do the impossible.
(Walt Disney)

Definition 2.1 An ideal $I \subseteq P$ is called an **ε-approximate vanishing ideal** of $\mathbb{X}$ if there exists a system of generators $\{f_1, \ldots, f_r\}$ of I such that $\|f_i\| = 1$ and f_i vanishes ε-approximately at $\mathbb{X}$ for $i = 1, \ldots, r$.

Problems: (1) Approximate vanishing ideals are **not at all unique**. They are not necessarily zero-dimensional either!

(2) If the coordinates of the points are very small, **every** polynomial of norm 1 in $\langle x_1, \ldots, x_n \rangle$ vanishes at $\mathbb{X}$. If they are very large, **no** polynomial of norm 1 vanishes at $\mathbb{X}$.

Definition 2.2 Let $\mathcal{O} = \{t_1, \ldots, t_\mu\}$ be an order ideal, let $\partial\mathcal{O} = \{b_1, \ldots, b_\nu\}$, and let $\varepsilon > 0$.

Definition 2.2 Let $\mathcal{O} = \{t_1, \ldots, t_\mu\}$ be an order ideal, let $\partial\mathcal{O} = \{b_1, \ldots, b_\nu\}$, and let $\varepsilon > 0$.

A set of polynomials $G = \{g_1, \ldots, g_\nu\}$ is called an **ε-approximate $\mathcal{O}$-border basis** if the following conditions are satisfied:

Definition 2.2 Let $\mathcal{O} = \{t_1, \ldots, t_\mu\}$ be an order ideal, let $\partial\mathcal{O} = \{b_1, \ldots, b_\nu\}$, and let $\varepsilon > 0$.

A set of polynomials $G = \{g_1, \ldots, g_\nu\}$ is called an **ε-approximate $\mathcal{O}$-border basis** if the following conditions are satisfied:

(a) The set G is an $\mathcal{O}$-border prebasis, i.e. g_j is of the form $g_j = b_j - \sum_{i=1}^{\mu} c_{ij} t_i$.

Definition 2.2 Let $\mathcal{O} = \{t_1, \ldots, t_\mu\}$ be an order ideal, let $\partial\mathcal{O} = \{b_1, \ldots, b_\nu\}$, and let $\varepsilon > 0$.

A set of polynomials $G = \{g_1, \ldots, g_\nu\}$ is called an **ε-approximate $\mathcal{O}$-border basis** if the following conditions are satisfied:

(a) The set G is an $\mathcal{O}$-border prebasis, i.e. g_j is of the form $g_j = b_j - \sum_{i=1}^{\mu} c_{ij} t_i$.

(b) For all pairs (i, j) such that (b_i, b_j) are neighbors, we have $\| \mathrm{NR}_{\mathcal{O},G}(S_{ij}) \| < \varepsilon$.

Definition 2.2 Let $\mathcal{O} = \{t_1, \ldots, t_\mu\}$ be an order ideal, let $\partial\mathcal{O} = \{b_1, \ldots, b_\nu\}$, and let $\varepsilon > 0$.

A set of polynomials $G = \{g_1, \ldots, g_\nu\}$ is called an **ε-approximate $\mathcal{O}$-border basis** if the following conditions are satisfied:

(a) The set G is an $\mathcal{O}$-border prebasis, i.e. g_j is of the form $g_j = b_j - \sum_{i=1}^{\mu} c_{ij} t_i$.

(b) For all pairs (i, j) such that (b_i, b_j) are neighbors, we have $\|\,\mathrm{NR}_{\mathcal{O},G}(S_{ij})\| < \varepsilon$.

Remark 2.3 If $G = \{g_1, \ldots, g_\nu\}$ is an ε-approximate border basis then the point $(c_{11}, \ldots, c_{\mu\nu})$ in $\mathbb{R}^{\mu\nu}$ given by its coefficients is **close to** the **border basis scheme**.

Example 2.4 Let $\mathcal{O} = \{1,\, x,\, y,\, xy\}$. Then the set

$$g_1 = x^2 + 0.02xy - 0.01y - 1.01 \qquad g_2 = x^2y + 0.03x - 0.98\,y$$
$$g_3 = xy^2 - 1.02\,x \qquad g_4 = y^2 - 0.99$$

is an approximate $\mathcal{O}$-border basis. The ideal $I = \langle g_1, g_2, g_3, g_4 \rangle$ is the unit ideal, since $g_3 - x\,g_4 = 0.03\,x$ shows $-g_1 \equiv 0.01\,y + 1.01$ and $g_4 \equiv 1.01^2 - 0.99 \pmod{I}$.

Example 2.4 Let $\mathcal{O} = \{1,\, x,\, y,\, xy\}$. Then the set

$$g_1 = x^2 + 0.02xy - 0.01y - 1.01 \qquad g_2 = x^2y + 0.03x - 0.98\,y$$
$$g_3 = xy^2 - 1.02\,x \qquad g_4 = y^2 - 0.99$$

is an approximate $\mathcal{O}$-border basis. The ideal $I = \langle g_1, g_2, g_3, g_4 \rangle$ is the unit ideal, since $g_3 - x\,g_4 = 0.03\,x$ shows $-g_1 \equiv 0.01\,y + 1.01$ and $g_4 \equiv 1.01^2 - 0.99 \pmod{I}$.

Goal of the AVI Algorithm: Given a set of (empirical) points $\mathbb{X} = \{p_1, \ldots, p_s\}$ in $\mathbb{R}^n$ and $\varepsilon > 0$, find an order ideal $\mathcal{O}$ and an approximate $\mathcal{O}$-border basis G such that the polynomials in G vanish ε-approximately at the points of $\mathbb{X}$.

Example 2.4 Let $\mathcal{O} = \{1,\, x,\, y,\, xy\}$. Then the set

$$g_1 = x^2 + 0.02xy - 0.01y - 1.01 \qquad g_2 = x^2y + 0.03x - 0.98\,y$$

$$g_3 = xy^2 - 1.02\,x \qquad g_4 = y^2 - 0.99$$

is an approximate $\mathcal{O}$-border basis. The ideal $I = \langle g_1, g_2, g_3, g_4 \rangle$ is the unit ideal, since $g_3 - x\,g_4 = 0.03\,x$ shows $-g_1 \equiv 0.01\,y + 1.01$ and $g_4 \equiv 1.01^2 - 0.99 \pmod{I}$.

Goal of the AVI Algorithm: Given a set of (empirical) points $\mathbb{X} = \{p_1, \ldots, p_s\}$ in $\mathbb{R}^n$ and $\varepsilon > 0$, find an order ideal $\mathcal{O}$ and an approximate $\mathcal{O}$-border basis G such that the polynomials in G vanish ε-approximately at the points of $\mathbb{X}$.

Notice that, in general, we have $\#\mathcal{O} << \#\mathbb{X}$, and the ideal $\langle G \rangle$ is the unit ideal.

3 – The Singular Value Decomposition

Under capitalism, man exploits man.

3 – The Singular Value Decomposition

Under capitalism, man exploits man.
Under communism, it's just the opposite.
(John Kenneth Galbraith)

3 – The Singular Value Decomposition

Under capitalism, man exploits man.
Under communism, it's just the opposite.
(John Kenneth Galbraith)

Theorem 3.1 *Let* $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{R})$*.*
There are orthogonal matrices $\mathcal{U} \in \mathrm{Mat}_{m,m}(\mathbb{R})$ *and* $\mathcal{V} \in \mathrm{Mat}_{n,n}(\mathbb{R})$ *and a matrix* $\mathcal{S} \in \mathrm{Mat}_{m,n}(\mathbb{R})$ *of the form* $\mathcal{S} = \begin{pmatrix} \mathcal{D} & 0 \\ 0 & 0 \end{pmatrix}$ *such that*

$$\mathcal{A} = \mathcal{U} \cdot \mathcal{S} \cdot \mathcal{V}^{\mathrm{tr}} = \mathcal{U} \cdot \begin{pmatrix} \mathcal{D} & 0 \\ 0 & 0 \end{pmatrix} \cdot \mathcal{V}^{\mathrm{tr}}$$

where $\mathcal{D} = \mathrm{diag}(s_1, \ldots, s_r)$ *is a diagonal matrix.*

In this decomposition, it is possible to achieve:

In this decomposition, it is possible to achieve:

(1) $s_1 \geq s_2 \geq \cdots \geq s_r > 0$. The numbers $s_1, \ldots, s_r$ depend only on $\mathcal{A}$ and are called the **singular values** of $\mathcal{A}$.

In this decomposition, it is possible to achieve:

(1) $s_1 \geq s_2 \geq \cdots \geq s_r > 0$. The numbers $s_1, \ldots, s_r$ depend only on $\mathcal{A}$ and are called the **singular values** of $\mathcal{A}$.

(2) The number r is the rank of $\mathcal{A}$.

In this decomposition, it is possible to achieve:

(1) $s_1 \geq s_2 \geq \cdots \geq s_r > 0$. The numbers $s_1, \ldots, s_r$ depend only on $\mathcal{A}$ and are called the **singular values** of $\mathcal{A}$.

(2) The number r is the rank of $\mathcal{A}$.

(3) The matrices $\mathcal{U}$ and $\mathcal{V}$ have the following interpretation:

$$
\begin{array}{ccc}
\text{first } r \text{ columns of } \mathcal{U} & \equiv & \text{ONB of the column space of } \mathcal{A} \\
\text{last } m-r \text{ columns of } \mathcal{U} & \equiv & \text{ONB of the kernel of } \mathcal{A}^{\mathrm{tr}} \\
\text{first } r \text{ columns of } \mathcal{V} & \equiv & \text{ONB of the row space of } \mathcal{A} \\
 & \equiv & \text{ONB of the column space of } \mathcal{A}^{\mathrm{tr}} \\
\text{last } n-r \text{ columns of } \mathcal{V} & \equiv & \text{ONB of the kernel of } \mathcal{A}
\end{array}
$$

Definition 3.2 Let $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{R})$, and let $\varepsilon > 0$ be given. Let $k \in \{1, \ldots, r\}$ be chosen such that $s_k > \varepsilon \geq s_{k+1}$. Form the matrix $\widetilde{\mathcal{A}} = \mathcal{U}\,\widetilde{\mathcal{S}}\,\mathcal{V}^{\mathrm{tr}}$ by setting $s_{k+1} = \cdots = s_r = 0$ in $\mathcal{S}$. Then $\widetilde{\mathcal{A}}$ is called the **singular value truncation** of $\mathcal{A}$ at ε.

Definition 3.2 Let $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{R})$, and let $\varepsilon > 0$ be given. Let $k \in \{1, \ldots, r\}$ be chosen such that $s_k > \varepsilon \geq s_{k+1}$. Form the matrix $\widetilde{\mathcal{A}} = \mathcal{U}\,\widetilde{\mathcal{S}}\,\mathcal{V}^{\mathrm{tr}}$ by setting $s_{k+1} = \cdots = s_r = 0$ in $\mathcal{S}$. Then $\widetilde{\mathcal{A}}$ is called the **singular value truncation** of $\mathcal{A}$ at ε.

Corollary 3.3 *Let $\widetilde{\mathcal{A}}$ be the singular value truncation of $\mathcal{A}$ at ε.*

Definition 3.2 Let $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{R})$, and let $\varepsilon > 0$ be given. Let $k \in \{1, \ldots, r\}$ be chosen such that $s_k > \varepsilon \geq s_{k+1}$. Form the matrix $\widetilde{\mathcal{A}} = \mathcal{U}\,\widetilde{\mathcal{S}}\,\mathcal{V}^{\mathrm{tr}}$ by setting $s_{k+1} = \cdots = s_r = 0$ in $\mathcal{S}$. Then $\widetilde{\mathcal{A}}$ is called the **singular value truncation** of $\mathcal{A}$ at ε.

Corollary 3.3 *Let $\widetilde{\mathcal{A}}$ be the singular value truncation of $\mathcal{A}$ at ε.*

(a) $\|\mathcal{A} - \widetilde{\mathcal{A}}\| = s_{k+1} = \min\{\|\mathcal{A} - \mathcal{B}\| : \mathrm{rank}(\mathcal{B}) \leq k\}$

Definition 3.2 Let $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{R})$, and let $\varepsilon > 0$ be given. Let $k \in \{1, \ldots, r\}$ be chosen such that $s_k > \varepsilon \geq s_{k+1}$. Form the matrix $\widetilde{\mathcal{A}} = \mathcal{U}\,\widetilde{\mathcal{S}}\,\mathcal{V}^{\mathrm{tr}}$ by setting $s_{k+1} = \cdots = s_r = 0$ in $\mathcal{S}$. Then $\widetilde{\mathcal{A}}$ is called the **singular value truncation** of $\mathcal{A}$ at ε.

Corollary 3.3 *Let $\widetilde{\mathcal{A}}$ be the singular value truncation of $\mathcal{A}$ at ε.*

(a) $\|\mathcal{A} - \widetilde{\mathcal{A}}\| = s_{k+1} = \min\{\|\mathcal{A} - \mathcal{B}\| : \mathrm{rank}(\mathcal{B}) \leq k\}$

(b) *The vector subspace* $\mathrm{apker}(\mathcal{A}, \varepsilon) = \ker(\widetilde{\mathcal{A}})$ *is the largest dimensional kernel of a matrix whose Euclidean distance from $\mathcal{A}$ is at most ε. It is called the* **ε-approximate kernel** *of $\mathcal{A}$.*

Definition 3.2 Let $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{R})$, and let $\varepsilon > 0$ be given. Let $k \in \{1, \ldots, r\}$ be chosen such that $s_k > \varepsilon \geq s_{k+1}$. Form the matrix $\widetilde{\mathcal{A}} = \mathcal{U}\,\widetilde{\mathcal{S}}\,\mathcal{V}^{\mathrm{tr}}$ by setting $s_{k+1} = \cdots = s_r = 0$ in $\mathcal{S}$. Then $\widetilde{\mathcal{A}}$ is called the **singular value truncation** of $\mathcal{A}$ at ε.

Corollary 3.3 *Let $\widetilde{\mathcal{A}}$ be the singular value truncation of $\mathcal{A}$ at ε.*

(a) $\|\mathcal{A} - \widetilde{\mathcal{A}}\| = s_{k+1} = \min\{\|\mathcal{A} - \mathcal{B}\| : \mathrm{rank}(\mathcal{B}) \leq k\}$

(b) *The vector subspace* $\mathrm{apker}(\mathcal{A}, \varepsilon) = \ker(\widetilde{\mathcal{A}})$ *is the largest dimensional kernel of a matrix whose Euclidean distance from $\mathcal{A}$ is at most ε. It is called the* **ε-approximate kernel** *of $\mathcal{A}$.*

(c) *The last $n-k$ columns $v_{k+1}, \ldots, v_n$ of $\mathcal{V}$ are an ONB of* $\mathrm{apker}(\mathcal{A}, \varepsilon)$. *They satisfy $\|\mathcal{A}v_i\| < \varepsilon$.*

4 – Stabilizing Gaußian Reduction

Question to Radio Eriwan:

Is it true that there are polynomials in the approximate vanishing ideal which do not vanish approximately?

4 – Stabilizing Gaußian Reduction

Question to Radio Eriwan:

Is it true that there are polynomials in the approximate vanishing ideal which do not vanish approximately?

Radio Eriwan answers:

In principle yes. Approximately.

4 – Stabilizing Gaußian Reduction

Question to Radio Eriwan:

Is it true that there are polynomials in the approximate vanishing ideal which do not vanish approximately?

Radio Eriwan answers:

In principle yes. Approximately.

$P = \mathbb{R}[x_1, \ldots, x_n]$ polynomial ring over $\mathbb{R}$

$V \subset P$ K-vector subspace with basis $B = \{f_1, \ldots, f_r\}$

σ term ordering on $\mathbb{T}^n$

4 – Stabilizing Gaußian Reduction

Question to Radio Eriwan:

Is it true that there are polynomials in the approximate vanishing ideal which do not vanish approximately?

Radio Eriwan answers:

In principle yes. Approximately.

$P = \mathbb{R}[x_1, \ldots, x_n]$ polynomial ring over $\mathbb{R}$

$V \subset P$ K-vector subspace with basis $B = \{f_1, \ldots, f_r\}$

σ term ordering on $\mathbb{T}^n$

Problem: Find all **approximate leading terms** in V, i.e. all all leading terms of unitary polynomials in V whose leading coefficient is larger than a given thereshold number $\tau > 0$.

The first step is to transform this to a matrix problem:

The first step is to transform this to a matrix problem:

Definition 4.1 Let $S = \mathrm{Supp}(f_1) \cup \cdots \cup \mathrm{Supp}(f_r)$ and write $S = \{t_1, \ldots, t_s\}$ where the terms $t_i \in \mathbb{T}^n$ are ordered such that $t_1 \geq_\sigma t_2 \geq_\sigma \cdots \geq_\sigma t_s$.

The first step is to transform this to a matrix problem:

Definition 4.1 Let $S = \mathrm{Supp}(f_1) \cup \cdots \cup \mathrm{Supp}(f_r)$ and write $S = \{t_1, \ldots, t_s\}$ where the terms $t_i \in \mathbb{T}^n$ are ordered such that $t_1 \geq_\sigma t_2 \geq_\sigma \cdots \geq_\sigma t_s$. We have $\mathrm{Supp}(V) \subseteq S$. We write

$$f_i = c_{i1} t_1 + \cdots + c_{is} t_s \qquad \text{with} \quad c_{ij} \in \mathbb{R}$$

for $i = 1, \ldots, r$. Then the matrix $M_{\sigma,B} = (c_{ij}) \in \mathrm{Mat}_{r,s}(\mathbb{R})$ is called the **coefficient matrix** of V with respect to σ and B.

The first step is to transform this to a matrix problem:

Definition 4.1 Let $S = \mathrm{Supp}(f_1) \cup \cdots \cup \mathrm{Supp}(f_r)$ and write $S = \{t_1, \ldots, t_s\}$ where the terms $t_i \in \mathbb{T}^n$ are ordered such that $t_1 \geq_\sigma t_2 \geq_\sigma \cdots \geq_\sigma t_s$. We have $\mathrm{Supp}(V) \subseteq S$. We write

$$f_i = c_{i1}t_1 + \cdots + c_{is}t_s \qquad \text{with} \quad c_{ij} \in \mathbb{R}$$

for $i = 1, \ldots, r$. Then the matrix $M_{\sigma,B} = (c_{ij}) \in \mathrm{Mat}_{r,s}(\mathbb{R})$ is called the **coefficient matrix** of V with respect to σ and B.

Remark 4.2 To find a **RREF** **(Reduced Row Echelon Form)** of $M_{\sigma,B}$, we could use Gaußian reduction. However, it is known that this algorithm is **numerically unstable**.

The first step is to transform this to a matrix problem:

Definition 4.1 Let $S = \mathrm{Supp}(f_1) \cup \cdots \cup \mathrm{Supp}(f_r)$ and write $S = \{t_1, \ldots, t_s\}$ where the terms $t_i \in \mathbb{T}^n$ are ordered such that $t_1 \geq_\sigma t_2 \geq_\sigma \cdots \geq_\sigma t_s$. We have $\mathrm{Supp}(V) \subseteq S$. We write

$$f_i = c_{i1} t_1 + \cdots + c_{is} t_s \qquad \text{with} \quad c_{ij} \in \mathbb{R}$$

for $i = 1, \ldots, r$. Then the matrix $M_{\sigma,B} = (c_{ij}) \in \mathrm{Mat}_{r,s}(\mathbb{R})$ is called the **coefficient matrix** of V with respect to σ and B.

Remark 4.2 To find a **RREF (Reduced Row Echelon Form)** of $M_{\sigma,B}$, we could use Gaußian reduction. However, it is known that this algorithm is **numerically unstable**.

A more stable method would be to use **complete pivoting**. But since we have to respect the ordering of the columns, we can do only **partial pivoting**.

Proposition 4.3 (Stabilized Reduced Row Echelon Form)
Let $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ *and* $\tau > 0$ *be given. Let* $a_1, \ldots, a_n$ *be the columns of* A*. Consider the following instructions.*

Proposition 4.3 (Stabilized Reduced Row Echelon Form) *Let $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ and $\tau > 0$ be given. Let $a_1, \ldots, a_n$ be the columns of A. Consider the following instructions.*

(1) Let $\lambda_1 = \|a_1\|$. If $\lambda_1 < \tau$, let $R = (0, \ldots, 0) \in \mathrm{Mat}_{m,1}(\mathbb{R})$. Otherwise, let $Q = ((1/\lambda_1)\, a_1) \in \mathrm{Mat}_{m,1}(\mathbb{R})$ and $R = (\lambda_1, 0, \ldots, 0) \in \mathrm{Mat}_{m,1}(\mathbb{R})$.

Proposition 4.3 (Stabilized Reduced Row Echelon Form)
Let $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ and $\tau > 0$ be given. Let $a_1, \ldots, a_n$ be the columns of A. Consider the following instructions.

(1) Let $\lambda_1 = \|a_1\|$. If $\lambda_1 < \tau$, let $R = (0, \ldots, 0) \in \mathrm{Mat}_{m,1}(\mathbb{R})$. Otherwise, let $Q = ((1/\lambda_1)\, a_1) \in \mathrm{Mat}_{m,1}(\mathbb{R})$ and $R = (\lambda_1, 0, \ldots, 0) \in \mathrm{Mat}_{m,1}(\mathbb{R})$.

(2) For $i = 2, \ldots, n$, compute $q_i = a_i - \sum_{j=1}^{i-1} \langle a_i, q_j \rangle\, q_j$ and $\lambda_i = \|q_i\|$. If $\lambda_i < \tau$, append a zero column to R. Otherwise, append the column $(1/\lambda_i)\, q_i$ to Q and the column $(\lambda_i \langle a_1, q_1 \rangle, \ldots, \lambda_i \langle a_{i-1}, q_{i-1} \rangle, \lambda_i, 0, \ldots, 0)$ to R.

Proposition 4.3 (Stabilized Reduced Row Echelon Form)
Let $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ and $\tau > 0$ be given. Let $a_1, \ldots, a_n$ be the columns of A. Consider the following instructions.

(1) Let $\lambda_1 = \|a_1\|$. If $\lambda_1 < \tau$, let $R = (0, \ldots, 0) \in \mathrm{Mat}_{m,1}(\mathbb{R})$. Otherwise, let $Q = ((1/\lambda_1)\, a_1) \in \mathrm{Mat}_{m,1}(\mathbb{R})$ and $R = (\lambda_1, 0, \ldots, 0) \in \mathrm{Mat}_{m,1}(\mathbb{R})$.

(2) For $i = 2, \ldots, n$, compute $q_i = a_i - \sum_{j=1}^{i-1} \langle a_i, q_j \rangle\, q_j$ and $\lambda_i = \|q_i\|$. If $\lambda_i < \tau$, append a zero column to R. Otherwise, append the column $(1/\lambda_i)\, q_i$ to Q and the column $(\lambda_i \langle a_1, q_1 \rangle, \ldots, \lambda_i \langle a_{i-1}, q_{i-1} \rangle, \lambda_i, 0, \ldots, 0)$ to R.

(3) Starting with the last row and working upwards, use the first non-zero entry of each row of R to clean out the non-zero entries above it.

(4) For $i = 1, \ldots, m$, compute the norm ϱ_i of the i-th row of R. If $\varrho_i < \tau$, set this row to zero. Otherwise, divide this row by ϱ_i. Then return the matrix R.

(4) For $i = 1, \ldots, m$, compute the norm ϱ_i of the i-th row of R. If $\varrho_i < \tau$, set this row to zero. Otherwise, divide this row by ϱ_i. Then return the matrix R.

This is an algorithm which computes a matrix R in reduced row echelon form. The row space of R is contained in the row space of the matrix $\overline{A}$ which is obtained from A by setting columns whose norm is less than τ to zero. Here the pivot elements of R are not 1, but its rows are unitary vectors.

(4) For $i = 1, \ldots, m$, compute the norm ϱ_i of the i-th row of R. If $\varrho_i < \tau$, set this row to zero. Otherwise, divide this row by ϱ_i. Then return the matrix R.

This is an algorithm which computes a matrix R in reduced row echelon form. The row space of R is contained in the row space of the matrix $\overline{A}$ which is obtained from A by setting columns whose norm is less than τ to zero. Here the pivot elements of R are not 1, but its rows are unitary vectors.

Furthermore, if the rows of A are unitary and mutually orthogonal, the row vectors of R differ by less than $\tau\, m \sqrt{n}$ from unitary vectors in the row space of A.

5 – The AVI Algorithm

A good algorithm is
10% inspiration, 15% perspiration,

5 – The AVI Algorithm

A good algorithm is
10% inspiration, 15% perspiration,
and 75% desperation.

5 – The AVI Algorithm

A good algorithm is
10% inspiration, 15% perspiration,
and 75% desperation.

Theorem 5.1 **(The AVI Algorithm)**
The following algorithm computes an approximate border basis of an approximate vanishing ideal of a finite set of points $\mathbb{X} \subseteq [-1,1]^n$.

5 – The AVI Algorithm

A good algorithm is
10% inspiration, 15% perspiration,
and 75% desperation.

Theorem 5.1 **(The AVI Algorithm)**
The following algorithm computes an approximate border basis of an approximate vanishing ideal of a finite set of points $\mathbb{X} \subseteq [\mathbf{-1, 1}]^{\mathbf{n}}$.

A1. Start with lists $G = \emptyset$, $\mathcal{O} = [1]$, a matrix $\mathcal{M} = (1, \ldots, 1)^{\mathrm{tr}} \in \mathrm{Mat}_{s,1}(\mathbb{R})$, and $d = 0$.

5 – The AVI Algorithm

A good algorithm is
10% inspiration, 15% perspiration,
and 75% desperation.

Theorem 5.1 (The AVI Algorithm)
The following algorithm computes an approximate border basis of an approximate vanishing ideal of a finite set of points $\mathbb{X} \subseteq [\mathbf{-1}, \mathbf{1}]^{\mathbf{n}}$.

A1. Start with lists $G = \emptyset$, $\mathcal{O} = [1]$, a matrix $\mathcal{M} = (1, \ldots, 1)^{\mathrm{tr}} \in \mathrm{Mat}_{s,1}(\mathbb{R})$, and $d = 0$.

A2. Increase d by one and let L be the list of all terms of degree d in $\partial\mathcal{O}$, ordered decreasingly w.r.t. σ. If $L = \emptyset$, return the pair $(G, \mathcal{O})$ and stop. Otherwise, let $L = (t_1, \ldots, t_\ell)$.

A3. Let m be the number of columns of $\mathcal{M}$. Form the matrix

$$\mathcal{A} = (\mathrm{eval}(t_1), \ldots, \mathrm{eval}(t_\ell), \mathcal{M}) \in \mathrm{Mat}_{s,\ell+m}(\mathbb{R}).$$

Using its **SVD**, calculate a matrix $\mathcal{B}$ whose column vectors are an ONB of the **approximate kernel** $\mathrm{apker}(\mathcal{A}, \varepsilon)$.

A3. Let m be the number of columns of $\mathcal{M}$. Form the matrix

$$\mathcal{A} = (\mathrm{eval}(t_1), \ldots, \mathrm{eval}(t_\ell), \mathcal{M}) \in \mathrm{Mat}_{s,\ell+m}(\mathbb{R}).$$

Using its **SVD**, calculate a matrix $\mathcal{B}$ whose column vectors are an ONB of the **approximate kernel** $\mathrm{apker}(\mathcal{A}, \varepsilon)$.

A4. Compute the **stabilized reduced row echelon form** of $\mathcal{B}^{\mathrm{tr}}$ with respect to the given τ. The result is a matrix $\mathcal{C} = (c_{ij}) \in \mathrm{Mat}_{k,\ell+m}(\mathbb{R})$ such that $c_{ij} = 0$ for $j < \nu(i)$. Here $\nu(i)$ is the column index of the pivot element in the i^{th} row of $\mathcal{C}$.

A3. Let m be the number of columns of $\mathcal{M}$. Form the matrix

$$\mathcal{A} = (\mathrm{eval}(t_1), \ldots, \mathrm{eval}(t_\ell), \mathcal{M}) \in \mathrm{Mat}_{s,\ell+m}(\mathbb{R}).$$

Using its **SVD**, calculate a matrix $\mathcal{B}$ whose column vectors are an ONB of the **approximate kernel** $\mathrm{apker}(\mathcal{A}, \varepsilon)$.

A4. Compute the **stabilized reduced row echelon form** of $\mathcal{B}^{\mathrm{tr}}$ with respect to the given τ. The result is a matrix $\mathcal{C} = (c_{ij}) \in \mathrm{Mat}_{k,\ell+m}(\mathbb{R})$ such that $c_{ij} = 0$ for $j < \nu(i)$. Here $\nu(i)$ is the column index of the pivot element in the i^{th} row of $\mathcal{C}$.

A5. For all $j \in \{1, \ldots, \ell\}$ which are column indices of pivot elements, append the polynomial

$$c_{ij} t_j + \textstyle\sum_{j'=j+1}^{\ell} c_{ij'} t_{j'} + \sum_{j'=\ell+1}^{\ell+m} c_{ij'} u_{j'}$$

to the list G, where $u_{j'}$ is the $(j' - \ell)^{\mathrm{th}}$ element of $\mathcal{O}$.

A6. For all $j = \ell, \ell - 1, \ldots, 1$ such that the j^{th} column of $\mathcal{C}$ contains no pivot element, append the term t_j as a new first element to $\mathcal{O}$ and append the column $\text{eval}(t_j)$ as a new first column to $\mathcal{M}$.

A6. For all $j = \ell, \ell - 1, \ldots, 1$ such that the j^{th} column of $\mathcal{C}$ contains no pivot element, append the term t_j as a new first element to $\mathcal{O}$ and append the column $\mathrm{eval}(t_j)$ as a new first column to $\mathcal{M}$.

A7. Using the SVD of $\mathcal{M}$, calculate a matrix $\mathcal{B}$ whose column vectors are an ONB of $\mathrm{apker}(\mathcal{M}, \varepsilon)$.

A6. For all $j = \ell, \ell - 1, \ldots, 1$ such that the j^{th} column of $\mathcal{C}$ contains no pivot element, append the term t_j as a new first element to $\mathcal{O}$ and append the column $\operatorname{eval}(t_j)$ as a new first column to $\mathcal{M}$.

A7. Using the SVD of $\mathcal{M}$, calculate a matrix $\mathcal{B}$ whose column vectors are an ONB of $\operatorname{apker}(\mathcal{M}, \varepsilon)$.

A8. Repeat steps **A4** – **A7** until $\mathcal{B}$ is empty. Then continue with step **A2**.

A6. For all $j = \ell, \ell - 1, \ldots, 1$ such that the j^{th} column of $\mathcal{C}$ contains no pivot element, append the term t_j as a new first element to $\mathcal{O}$ and append the column $\operatorname{eval}(t_j)$ as a new first column to $\mathcal{M}$.

A7. Using the SVD of $\mathcal{M}$, calculate a matrix $\mathcal{B}$ whose column vectors are an ONB of $\operatorname{apker}(\mathcal{M}, \varepsilon)$.

A8. Repeat steps **A4** – **A7** until $\mathcal{B}$ is empty. Then continue with step **A2**.

The resulting set $\mathcal{O} = \{t_1, \ldots, t_\mu\}$ is an order ideal such that no unitary polynomial in $\langle \mathcal{O} \rangle_K$ vanishes ε-approximately on $\mathbb{X}$.

The resulting set G is a δ-approximate $\mathcal{O}$-border basis. (An explicit bound for δ can be given.)

A6. For all $j = \ell, \ell-1, \ldots, 1$ such that the j^{th} column of $\mathcal{C}$ contains no pivot element, append the term t_j as a new first element to $\mathcal{O}$ and append the column $\operatorname{eval}(t_j)$ as a new first column to $\mathcal{M}$.

A7. Using the SVD of $\mathcal{M}$, calculate a matrix $\mathcal{B}$ whose column vectors are an ONB of $\operatorname{apker}(\mathcal{M}, \varepsilon)$.

A8. Repeat steps **A4** – **A7** until $\mathcal{B}$ is empty. Then continue with step **A2**.

The resulting set $\mathcal{O} = \{t_1, \ldots, t_\mu\}$ is an order ideal such that no unitary polynomial in $\langle \mathcal{O} \rangle_K$ vanishes ε-approximately on $\mathbb{X}$.

The resulting set G is a δ-approximate $\mathcal{O}$-border basis. (An explicit bound for δ can be given.)

The AVI algorithm has been implemented in the **ApCoCoA** library (see `http://www.apcocoa.org`)

6 – Approximate Examples

Cheer up! The worst is yet to come!

(Mark Twain)

6 – Approximate Examples

Cheer up! The worst is yet to come!
(Mark Twain)

Example 6.1 **(Four Almost Aligned Points)**
Let $\mathbb{X} = \{(0, 0.01),\ (0.34, 0.32),\ (0.65, 0.68),\ (0.99, 1)\}$, and let $\varepsilon = 0.05$. When we apply the AVI algorithm to this case, we get $\mathcal{O} = \{1,\ y,\ y^2\}$ and the approximate border basis $G = \{x - 0.99y,\ xy - 1.01y^2 + 0.02y,\ xy^2 - 1.59y^2 + 0.62y - 0.02,\ y^3 - 1.59y^2 + 0.60y - 0.02\}$.

6 – Approximate Examples

Cheer up! The worst is yet to come!
(Mark Twain)

Example 6.1 **(Four Almost Aligned Points)**
Let $\mathbb{X} = \{(0, 0.01), (0.34, 0.32), (0.65, 0.68), (0.99, 1)\}$, and let $\varepsilon = 0.05$. When we apply the AVI algorithm to this case, we get $\mathcal{O} = \{1,\, y,\, y^2\}$ and the approximate border basis $G = \{x - 0.99y,\, xy - 1.01y^2 + 0.02y,\, xy^2 - 1.59y^2 + 0.62y - 0.02,\, y^3 - 1.59y^2 + 0.60y - 0.02\}$.

Surprisingly, the set $\mathcal{O}$ corresponds to only **three** points. They are $\mathbb{X}' = \{(0.03, 0.04), (0.52, 0.52), (0.97, 0.98)\}$.

6 – Approximate Examples

Cheer up! The worst is yet to come!
(Mark Twain)

Example 6.1 **(Four Almost Aligned Points)**
Let $\mathbb{X} = \{(0, 0.01), (0.34, 0.32), (0.65, 0.68), (0.99, 1)\}$, and let $\varepsilon = 0.05$. When we apply the AVI algorithm to this case, we get $\mathcal{O} = \{1, y, y^2\}$ and the approximate border basis $G = \{x - 0.99y,\ xy - 1.01y^2 + 0.02y,\ xy^2 - 1.59y^2 + 0.62y - 0.02,\ y^3 - 1.59y^2 + 0.60y - 0.02\}$.

Surprisingly, the set $\mathcal{O}$ corresponds to only **three** points. They are $\mathbb{X}' = \{(0.03, 0.04), (0.52, 0.52), (0.97, 0.98)\}$.
What has happened is that AVI found a curve of degree 3 passing close to $\mathbb{X}$, namely $g_4 = 0.51y^3 - 0.80y^2 + 0.30y - 0.01$.

Example 6.2 **(Five Points on a Conic)**

Let $\mathbb{X} = \{(0,1),\ (0.2, 0.4),\ (0.28, 0.28),\ (0.4, 0.2),\ (1, 0)\}$.

Example 6.2 **(Five Points on a Conic)**

Let $\mathbb{X} = \{(0,1),\ (0.2, 0.4),\ (0.28, 0.28),\ (0.4, 0.2),\ (1,0)\}$.

(a) When we choose $\varepsilon = 0.1$, we get $\mathcal{O} = \{1, x, y\}$ and the (unitary) approximate border basis $G = 0.52x^2 - 0.77x - 0.25y + 0.25,\ 0.94xy + 0.18x + 0.18y - 0.18,\ 0.51y^2 - 0.26x - 0.77y + 0.26\}$. The set G approximately defines $\mathbb{X}' = \{(0, 0.98),\ (0.28, 0.29),\ (0.98, 0)\}$.

Example 6.2 **(Five Points on a Conic)**

Let $\mathbb{X} = \{(0,1),\ (0.2,0.4),\ (0.28,0.28),\ (0.4,0.2),\ (1,0)\}$.

(a) When we choose $\varepsilon = 0.1$, we get $\mathcal{O} = \{1, x, y\}$ and the (unitary) approximate border basis $G = 0.52x^2 - 0.77x - 0.25y + 0.25,\ 0.94xy + 0.18x + 0.18y - 0.18,\ 0.51y^2 - 0.26x - 0.77y + 0.26\}$. The set G approximately defines $\mathbb{X}' = \{(0,0.98),\ (0.28,0.29),\ (0.98,0)\}$.

(b) The choice $\varepsilon = 0.01$ leads to $\mathcal{O}' = \{1, x, y, y^2\}$ and $G' = \{0.3x^2+0.3y^2-0.6x-0.6y+0.3,\ 0.94xy+0.18x+0.18y-0.18,\ 0.95xy^2+0.19y^2-0.03x-0.22y+0.03,\ 0.42y^3-0.77y^2+0.10x+0.44y-0.10\}$.

Example 6.2 **(Five Points on a Conic)**

Let $\mathbb{X} = \{(0,1),\ (0.2,0.4),\ (0.28,0.28),\ (0.4,0.2),\ (1,0)\}$.

(a) When we choose $\varepsilon = 0.1$, we get $\mathcal{O} = \{1, x, y\}$ and the (unitary) approximate border basis $G = 0.52x^2 - 0.77x - 0.25y + 0.25,\ 0.94xy + 0.18x + 0.18y - 0.18,\ 0.51y^2 - 0.26x - 0.77y + 0.26\}$. The set G approximately defines $\mathbb{X}' = \{(0,0.98),\ (0.28,0.29),\ (0.98,0)\}$.

(b) The choice $\varepsilon = 0.01$ leads to $\mathcal{O}' = \{1, x, y, y^2\}$ and $G' = \{0.3x^2+0.3y^2-0.6x-0.6y+0.3,\ 0.94xy+0.18x+0.18y-0.18,\ 0.95xy^2 + 0.19y^2 - 0.03x - 0.22y + 0.03,\ 0.42y^3 - 0.77y^2 + 0.10x + 0.44y - 0.10\}$.

The set G' approximately defines the four points $\mathbb{X}'' = \{(0,0.99),\ (0.21,0.37),\ (0.37,0.21),\ (0.99,0)\}$. Thus even this small choice of ε leads to a decrease in the codimension of $I_{\mathbb{X}}$.

Finally we consider an industrial application of the AVI algorithm.

Finally we consider an industrial application of the AVI algorithm.

Example 6.3 Suppose that $\mathbb{X}$ consists of 6000 points in $\mathbb{R}^5$. The relative errors in the coordinates of the points of $\mathbb{X}$ are about 10%. It is reasonable to use the AVI algorithm with $\varepsilon = 0.1$.

Finally we consider an industrial application of the AVI algorithm.

Example 6.3 Suppose that $\mathbb{X}$ consists of 6000 points in $\mathbb{R}^5$. The relative errors in the coordinates of the points of $\mathbb{X}$ are about 10%. It is reasonable to use the AVI algorithm with $\varepsilon = 0.1$.

The result (after ca. 0.5s CPU time on my laptop) is the order ideal $\mathcal{O} = \{1, x_5, x_4, x_3, x_2, x_1, x_5^2, x_4x_5, x_3x_5, x_2x_5,$ $x_1x_5, x_4^2, x_3x_4, x_2x_4, x_1x_4, x_3^2, x_2x_3, x_5^3, x_4x_5^2, x_3x_5^2\}$ consisting of only **20** terms. The approximate border basis G has 43 elements.

Finally we consider an industrial application of the AVI algorithm.

Example 6.3 Suppose that $\mathbb{X}$ consists of 6000 points in $\mathbb{R}^5$. The relative errors in the coordinates of the points of $\mathbb{X}$ are about 10%. It is reasonable to use the AVI algorithm with $\varepsilon = 0.1$.

The result (after ca. 0.5s CPU time on my laptop) is the order ideal $\mathcal{O} = \{1, x_5, x_4, x_3, x_2, x_1, x_5^2, x_4x_5, x_3x_5, x_2x_5, x_1x_5, x_4^2, x_3x_4, x_2x_4, x_1x_4, x_3^2, x_2x_3, x_5^3, x_4x_5^2, x_3x_5^2\}$ consisting of only **20** terms. The approximate border basis G has 43 elements.

It is **not true** that there are 20 real points such that G is an approximate border basis of their vanishing ideal.

Finally we consider an industrial application of the AVI algorithm.

Example 6.3 Suppose that $\mathbb{X}$ consists of 6000 points in $\mathbb{R}^5$. The relative errors in the coordinates of the points of $\mathbb{X}$ are about 10%. It is reasonable to use the AVI algorithm with $\varepsilon = 0.1$.

The result (after ca. 0.5s CPU time on my laptop) is the order ideal $\mathcal{O} = \{1, x_5, x_4, x_3, x_2, x_1, x_5^2, x_4x_5, x_3x_5, x_2x_5,$ $x_1x_5, x_4^2, x_3x_4, x_2x_4, x_1x_4, x_3^2, x_2x_3, x_5^3, x_4x_5^2, x_3x_5^2\}$ consisting of only **20** terms. The approximate border basis G has 43 elements.

It is **not true** that there are 20 real points such that G is an approximate border basis of their vanishing ideal.

One can only say that the 20-dimensional space $\langle \mathcal{O} \rangle_{\mathbb{R}}$ suffices to interpolate approximately at the given 6000 points.