Algebra Primer

① Varieties

- Let $\mathbb{K}$ be a field, typically $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{Q}\}$
  statistics, algebraic geometry, computations
- $\mathbb{K}[\underline{x}] := \mathbb{K}[x_1, \dots, x_m]$ is the ring of polynomial functions in the indeterminates $x_1, \dots, x_m$

Def: Let $S \subseteq \mathbb{K}[\underline{x}]$. The variety defined by S is
$$V(S) := \{a \in \mathbb{K}^m \mid \forall f \in S : f(a) = 0\}.$$
$V(S)$ is also called the vanishing locus / zero locus of S.

Ex: $\mathbb{K} = \mathbb{R}$

$V(x_2 - x_1^2)$

$V(x_1^2 + x_2^2 - 1)$

$\left.\begin{matrix} V(x_2 - x_1^2, \\ x_1^2 + x_2^2 - 1) \end{matrix}\right\} =: Z$

$\mathbb{K}$	$\mathbb{Q}$	$\mathbb{R}$	$\mathbb{C}$
$\lvert Z \rvert$	0	2	4

② Ideals

Def: Let $Z \subseteq \mathbb{K}^m$. The vanishing ideal / defining ideal of Z is
$$I(Z) := \{f \in \mathbb{K}[\underline{x}] \mid \forall a \in Z : f(a) = 0\}.$$

- $I(Z)$ is an ideal $\left[\begin{matrix} \text{i.e.,} & f, g \in I(Z) & \Rightarrow f + g \in I(Z) \\ & f \in I(Z), h \in \mathbb{K}[\underline{x}] & \Rightarrow hf \in I(Z) \end{matrix}\right]$
- For $S \subseteq \mathbb{K}[\underline{x}]$, we write $\langle S \rangle := \{\sum_{i=1}^{k} h_i f_i \mid k \in \mathbb{Z}_{>0}, f_i \in S, h_i \in \mathbb{K}[\underline{x}]\}$ for the ideal generated by S.

Hilbert Basis Theorem
For every I in $\mathbb{K}[\underline{x}]$ there is a finite subset $S \subseteq I$ such that $I = \langle S \rangle$.

Lemma: Let $I \subseteq \mathbb{K}[\underline{x}]$ be an ideal. $\Rightarrow I \subseteq I(V(I))$

Ex: $I = \langle x_1^2 \rangle \subseteq \mathbb{R}[x_1, x_2]$. $\Rightarrow V(I) = \{(a_1, a_2) \in \mathbb{R}^2 \mid a_1 = 0\}$

$\Rightarrow I(V(I)) = \langle x_1 \rangle \supsetneq \langle x_1^2 \rangle$

Def: Let $I \subseteq \mathbb{K}[\underline{x}]$ be an ideal. The radical of I is
$$\sqrt{I} := \{f \in \mathbb{K}[\underline{x}] \mid \exists k \in \mathbb{Z}_{>0} : f^k \in I\}.$$
I is called radical if $\sqrt{I} = I$.

note: $\sqrt{I}$ is an ideal

Prop: Let $Z \subseteq \mathbb{K}^m$. $\Rightarrow I(Z)$ is a radical ideal.

Nullstellensatz: Let $\mathbb{K}$ be algebraically closed and let $I \subseteq \mathbb{K}[\underline{x}]$ be an ideal.
$$\Rightarrow I(V(I)) = \sqrt{I}.$$

$\forall f \in \mathbb{K}[x_1]$ $\exists a \in \mathbb{K} : f(a) = 0$

e.g. $\mathbb{K} = \mathbb{C}$

③ Ideal-Variety-Correspondence

Let $\mathbb{K}$ be algebraically closed.

$$\left\{\begin{matrix} \text{radical ideals} \\ \text{in } \mathbb{K}[\underline{x}] \end{matrix}\right\} \underset{I(\cdot)}{\overset{V(\cdot)}{\rightleftarrows}} \left\{\begin{matrix} \text{varieties} \\ \text{in } \mathbb{K}^m \end{matrix}\right\}$$

inclusion-reversing bijections

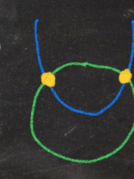
$I_1 \subseteq I_2 \Rightarrow V(I_2) \subseteq V(I_1)$
$Z_1 \subseteq Z_2 \Rightarrow I(Z_2) \subseteq I(Z_1)$

Ex: $I_1 = \langle x_2 \rangle \subseteq \mathbb{K}[x_1, x_2]$
$I_2 = \langle x_1, x_2 \rangle$
$\Rightarrow I_1 \subseteq I_2$

$V(I_1)$

$V(I_2)$

Prop: Let $I_1, I_2 \subseteq \mathbb{K}[\underline{x}]$ ideals. $\Rightarrow$ a) $V(I_1 + I_2) = V(I_1) \cap V(I_2)$
b) $V(I_1 \cap I_2) = V(I_1) \cup V(I_2)$



Let $Z_1, Z_2 \subseteq \mathbb{K}^m$. $\Rightarrow$ b) $I(Z_1 \cup Z_2) = I(Z_1) \cap I(Z_2)$

a) If $\mathbb{K}$ algebraically closed, then $I(Z_1 \cap Z_2) = \sqrt{I(Z_1) + I(Z_2)}$.

④ Zariski Topology

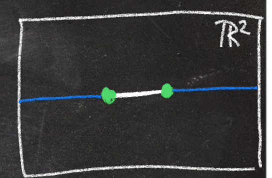
Def: The closed sets of the Zariski topology on $\mathbb{K}^m$ are the varieties in $\mathbb{K}^m$.

Prop: Let $Z \subseteq \mathbb{K}^m$. $\Rightarrow$ $V(I(Z))$ is the Zariski closure of Z,
[i.e. the smallest Zariski closed set (=variety) containing Z]

Ex: $Z = \{(a,0)\in\mathbb{R}^2 \mid 0<a<1\}$

$\Rightarrow$ Euclidean closure of Z is $\{(a,0)\in\mathbb{R}^2 \mid 0\le a\le 1\}$

Zariski closure of Z is $\{(a,0)\in\mathbb{R}^2\}$ $[I(Z)=\langle x_2\rangle]$



- Zariski closed sets are Euclidean closed, but generally not vice versa!
- The complement of a Zariski closed set is called a Zariski open set.

⑤ Example: Binomial Random Variables

- Consider the polynomial map $\phi: \mathbb{C} \longrightarrow \mathbb{C}^{m+1}$ where

$$\phi_i(t) = \binom{m}{i} t^i (1-t)^{m-i} \quad \text{for } i=0,1,\dots,m.$$

- If $\theta\in[0,1]\subseteq\mathbb{R}$ is the probability of getting head in 1 flip of a biased coin, $\phi_i(\theta)$ is the probability of getting i heads in m independent flips of the coin.

$\Rightarrow$ vector $\phi(\theta)$ is probability distribution of a binomial random variable

- $\phi([0,1])$ is a curve in the probality simplex

$$\Delta_m := \{p\in\mathbb{R}^{m+1}_{\ge 0} \mid \sum_{i=0}^{m} p_i = 1\}$$

- $\phi(\mathbb{C})$ is a curve in $\mathbb{C}^{m+1}$

$m=2$

⑤ Example: Binomial Random Variables

- Consider the polynomial map $\phi: \mathbb{C} \longrightarrow \mathbb{C}^{m+1}$ where

$$\phi_i(t) = \binom{m}{i} t^i (1-t)^{m-i} \quad \text{for } i=0,1,\dots,m.$$

- $\phi([0,1])$ is a curve in the probality simplex

$$\Delta_m := \{p\in\mathbb{R}^{m+1}_{\ge 0} \mid \sum_{i=0}^{m} p_i = 1\}$$

- $\phi(\mathbb{C})$ is a curve in $\mathbb{C}^{m+1}$:
it is the variety defined by $\sum_{i=0}^{m} p_i - 1 = 0$ and
the vanishing of all 2×2 minors of $\begin{bmatrix} p_0 & p_1/m & p_2/\binom{m}{2} & \cdots & p_{m-1}/m \\ p_1/m & p_2/\binom{m}{2} & p_3/\binom{m}{3} & \cdots & p_m \end{bmatrix}$

$m=2$

$I(\phi(\mathbb{C})) = \langle p_0+p_1+p_2-1,\ 4p_0p_2-p_1^2\rangle$

- $\phi([0,1]) = \phi(\mathbb{C}) \cap \Delta_m$
- $\phi(\mathbb{C})$ is the Zariski closure of the model $\phi([0,1])$ over $\mathbb{K}=\mathbb{C}$

⑥ Mixture Models

- Let $P\subseteq\Delta_m$ be a statistical model, i.e. a family of probability distributions
- For $s\in\mathbb{Z}_{>0}$, the s-th mixture model is

$$\mathrm{Mixt}^s(P) = \Big\{\sum_{j=1}^{s} \pi_j p^{(j)} \;\Big|\; \pi\in\Delta_{s-1} \ \&\ \forall j: p^{(j)}\in P\Big\}$$

Ex: $m=2$, $P=\phi([0,1])$, $s=2$

$\Rightarrow$ $\mathrm{Mixt}^2(P)$

$*$ The Zariski closure of $\mathrm{Mixt}^2(P)$ is the whole plane $V(p_0+p_1+p_2-1)$.

$*$ $\mathrm{Mixt}^2(P) = \Delta_2 \cap \underbrace{\{p\in\mathbb{R}^3 \mid 4p_0p_2-p_1^2 \ge 0\}}_{\text{Semialgebraic set}}$

⑦ Implicitization

Problem: What is the image of a given polynomial map $\phi: \mathbb{K}^d \to \mathbb{K}^m$?

Ex: model of binomial random variables & its mixture models

More general problem: What is the image of a given rational map $\phi: \mathbb{K}^d \dashrightarrow \mathbb{K}^m$?

$\forall i = 1, \dots, m: \ \phi_i = \frac{f_i}{g_i}$ where $f_i, g_i \in \mathbb{K}[t_1, \dots, t_d]$

"$\dashrightarrow$" : ϕ not defined on all of $\mathbb{K}^d$

ϕ well-defined on the Zariski open set $\mathbb{K}^d \setminus (V(g_1) \cup \dots \cup V(g_m))$

The image of ϕ is not a variety in general!

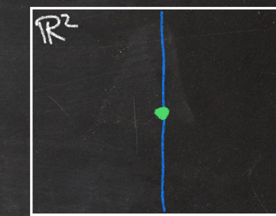
Ex: $\phi: \mathbb{C}^2 \to \mathbb{C}^2$

$(t_1, t_2) \mapsto (t_1, t_1 t_2)$

$\Rightarrow \phi(\mathbb{C}^2) = \{(a_1, a_2) \in \mathbb{C}^2 \mid a_1 = 0 \Rightarrow a_2 = 0\}$

$= (\mathbb{C}^2 \setminus V(a_1)) \cup V(a_1, a_2)$

$= \mathbb{C}^2 \setminus (V(a_1) \setminus V(a_1, a_2))$



Thm: Let $\mathbb{K}$ be algebraically closed, $V \subseteq \mathbb{K}^d$ a variety and $\phi: V \dashrightarrow \mathbb{K}^m$ a rational map.

$\Rightarrow \phi(V)$ is a constructible set, i.e.

there are finitely many varieties $Z_1 \supseteq Z_2 \supseteq \dots \supseteq Z_k$ in $\mathbb{K}^m$ such that $\phi(V) = Z_1 \setminus (Z_2 \setminus (\cdots \setminus (Z_{k-1} \setminus Z_k) \cdots))$.

Ex: $\phi: \mathbb{R} \to \mathbb{R}$

$t \mapsto t^2$

$\Rightarrow \phi(\mathbb{R}) = \mathbb{R}_{\geq 0}$ is not constructible

Tarski-Seidenberg Theorem

Let $V \subseteq \mathbb{R}^d$ be a semialgebraic set and $\phi: V \dashrightarrow \mathbb{R}^m$ a rational map.

$\Rightarrow \phi(V)$ is a semialgebraic set, i.e.

a finite union of sets defined by a finite number of polynomial equations and inequalities.

Constructible subsets of $\mathbb{R}^n$ are semialgebraic, but generally not vice versa!

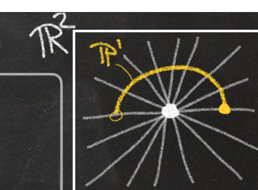
---

⑧ Projective Varieties

Def: The m-dimensional projective space is

$\mathbb{P}^m := \{\text{lines through origin in } \mathbb{K}^{m+1}\}$

$= (\mathbb{K}^{m+1} \setminus \{0\}) /_{\sim}$ where $a \sim b :\Leftrightarrow \exists \lambda \in \mathbb{K} \setminus \{0\}: a = \lambda b$



$\mathbb{P}^1 = \mathbb{K}^1 \cup \{\infty\}$

Notation: $(a_0 : a_1 : \dots : a_m)$ = equivalence class of $(a_0, \dots, a_m) \in \mathbb{K}^{m+1} \setminus \{0\}$

$\Rightarrow \mathbb{P}^m = \mathbb{K}^m \,\dot\cup\, \mathbb{P}^{m-1}$

$a_0 \neq 0$: $(a_0 : \dots : a_m) = (1 : \frac{a_1}{a_0} : \dots : \frac{a_m}{a_0})$

$a_0 = 0$: "hyperplane at ∞"

Projective space $\mathbb{P}^m$ is compact (in Euclidean topology), unlike affine space $\mathbb{K}^m$.

Ex: $(-1:-1) = (1:1) \in \mathbb{P}^1$ — not homogeneous!

$f(x_0, x_1) := x_1^2 - x_0 \Rightarrow f(1,1) = 0 \neq f(-1,-1)$

Def: A polynomial is homogeneous if all of its terms have the same degree.
An ideal is homogeneous if it is generated by a set of homogeneous polynomials.

f homogeneous $\Rightarrow f(\lambda a) = \lambda^{\deg(f)} \cdot f(a)$
$\Rightarrow V(f)$ well-defined in projective space

Def: The projective variety defined by a homogeneous ideal $I \subseteq \mathbb{K}[x_0, \ldots, x_m]$ is $V(I) := \{a \in \mathbb{P}^m \mid \forall \text{ homogeneous } f \in I : f(a) = 0\}$.

- Any 2 lines in $\mathbb{P}^2$ intersect

  line = zero locus of a homogeneous linear polynomial, e.g. $V(x_0 + 2x_1 - x_2)$ … ∞

- Any 2 conics in $\mathbb{P}^2_{\mathbb{C}}$ intersect in 4 points (counted with multiplicity)

Projective space $\mathbb{P}^m$ is compact (in Euclidean topology), unlike affine space $\mathbb{K}^m$.

$\mathbb{P}^m = \mathbb{K}^m \,\dot\cup\, \mathbb{P}^{m-1}$

$\{a \in \mathbb{P}^m \mid l(a) \neq 0\} \cong \{a \in \mathbb{K}^{m+1} \mid l(a) = 1\} \rightarrow l(a) \neq 0$ $\qquad l(a) = 0$

l = homog. linear pol.

statistical model, complicated problem

$\Delta_m \subseteq \{a \in \mathbb{R}^{m+1} \mid \sum_{i=0}^m a_i = 1\} \subseteq \{a \in \mathbb{C}^{m+1} \mid \sum_{i=0}^m a_i = 1\} \subseteq \mathbb{P}^m_{\mathbb{C}}$ easy algebra

⑨ Gröbner Bases

Linear algebra

All undergraduate students learn about Gaussian elimination, a general method for solving linear systems of algebraic equations:

Input:

$$x + 2y + 3z = 5$$
$$7x + 11y + 13z = 17$$
$$19x + 23y + 29z = 31$$

Output:

$$x = -35/18$$
$$y = 2/9$$
$$z = 13/6$$

Solving very large linear systems is central to applied mathematics.

Nonlinear algebra

Lucky students also learn about Gröbner bases, a general method for non-linear systems of algebraic equations:

Input:

$$x^2 + y^2 + z^2 = 2$$
$$x^3 + y^3 + z^3 = 3$$
$$x^4 + y^4 + z^4 = 4$$

Output:

$$3z^{12} - 12z^{10} - 12z^9 + 12z^8 + 72z^7 - 66z^6 - 12z^4 + 12z^3 - 1 = 0$$

$$4y^2 + (36z^{11} + 54z^{10} - 69z^9 - 252z^8 - 216z^7 + 573z^6 + 72z^5 - 12z^4 - 99z^3 + 10z + 3)\, y + 36z^{11} + 48z^{10} - 72z^9 - 234z^8 - 192z^7 + 564z^6 - 48z^5 + 96z^4 - 96z^3 + 10z^2 + 8 = 0$$

$$4x + 4y + 36z^{11} + 54z^{10} - 69z^9 - 252z^8 - 216z^7 + 573z^6 + 72z^5 - 12z^4 - 99z^3 + 10z + 3 = 0$$

This is very hard for large systems, but . . .

The world is non-linear!

Many models in the sciences and engineering are characterized by polynomial equations. Such a set is an algebraic variety.

- Algebraic statistics
- Machine learning
- Optimization
- Computer vision
- Robotics
- Complexity theory
- Cryptography
- Biology
- Economics
- ...

Def: A term order $<$ on $\mathbb{K}[\underline{x}] = \mathbb{K}[x_1, \dots, x_m]$ is a total order on the set of monomials in $\mathbb{K}[\underline{x}]$ such that

a) $\forall u \in \mathbb{Z}^m_{\geq 0}: 1 = x^0 \leq x^u$ and

b) $\forall u, v \in \mathbb{Z}^m_{\geq 0}: [x^u < x^v \Rightarrow \forall w \in \mathbb{Z}^m_{\geq 0}: x^w \cdot x^u < x^w \cdot x^v]$

Ex: The lexicographic term order $<_{lex}$ is defined by

$x^u <_{lex} x^v :\Leftrightarrow$ the leftmost nonzero entry in $v-u$ is positive.

e.g. $x_3^3 <_{lex} x_2 <_{lex} x_1^2 x_3^3 <_{lex} x_1^2 x_2 <_{lex} x_1^3$

Here we assumed: $x_m <_{lex} x_{m-1} <_{lex} \dots <_{lex} x_1$.

Any permutation of the indeterminates yields a **different** lexicographic term order!

Def: The initial monomial / initial term / leading term $\mathrm{in}_<(f)$ of $f \in \mathbb{K}[\underline{x}]$ with respect to a term order $<$ is the largest monomial with nonzero coefficient in f.

Ex: $\mathrm{in}_{<_{lex}}(x_1^2 - 3x_1^2 x_2 + \pi x_2^4) = x_1^2 x_2$

Def: The initial ideal of an ideal $I \subseteq \mathbb{K}[\underline{x}]$ with respect to a term order $<$ is $\mathrm{in}_<(I) := \langle \mathrm{in}_<(f) \mid f \in I \rangle$.

Ex: $I = \langle x_1^2, \; x_1x_2 + x_2^2 \rangle \Rightarrow \mathrm{in}_{<_{lex}}(I) = \langle x_1^2, x_1x_2, x_2^3 \rangle$

$[x_2^3 = x_2 \cdot x_1^2 - (x_1 - x_2) \cdot (x_1x_2 + x_2^2) \in I]$ $\neq \langle x_1^2, x_1x_2 \rangle$

$I = \langle S \rangle$ does in general **not** imply that $\mathrm{in}_<(I) = \langle \mathrm{in}_<(f) \mid f \in S \rangle$!

Def: A Gröbner basis of an ideal $I \subseteq \mathbb{K}[\underline{x}]$ with respect to a term order $<$ is a finite subset $\mathcal{G} \subseteq I$ such that $\mathrm{in}_<(I) = \langle \mathrm{in}_<(g) \mid g \in \mathcal{G} \rangle$.

Ex: $I = \langle x_1^2, x_1x_2 + x_2^2 \rangle$ has Gröbner basis $x_1^2, \; x_1x_2 + x_2^2, \; x_2^3$

- equivalently, a finite subset $\mathcal{G} \subseteq I$ is a Gröbner basis iff $\forall f \in I \setminus \{0\} \; \exists g \in \mathcal{G}: \mathrm{in}_<(g) \mid \mathrm{in}_<(f)$
- Gröbner bases always exist (by Hilbert basis theorem)
- If $\mathcal{G}$ is a Gröbner basis of I, then $I = \langle \mathcal{G} \rangle$.
- heart of computational algebra software:

Let Z be affine variety. From a Gröbner basis of $I(Z)$, can easily compute
- * dimension of Z
- * $I(\Phi(Z))$ for a rational map Φ (implicitization problem)
- * much more!