

Moment varieties of local mixtures and Terracini loci

Douglas A. Torrance

October 11, 2026

Georgia Tech

AMS Fall Southeastern Sectional Meeting, Kennesaw State University



Pearson's crabs



In 1894, Pearson studied measurements from 1000 crabs. They appeared to come from two populations.

Image credits: Pearson (public domain), crab (Luis Miguel Bugallo Sánchez, CC BY-SA 3.0)

Method of moments

Idea

Suppose the distribution is a mixture of two Gaussians:

$$x \mapsto \lambda \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} + (1-\lambda) \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}$$

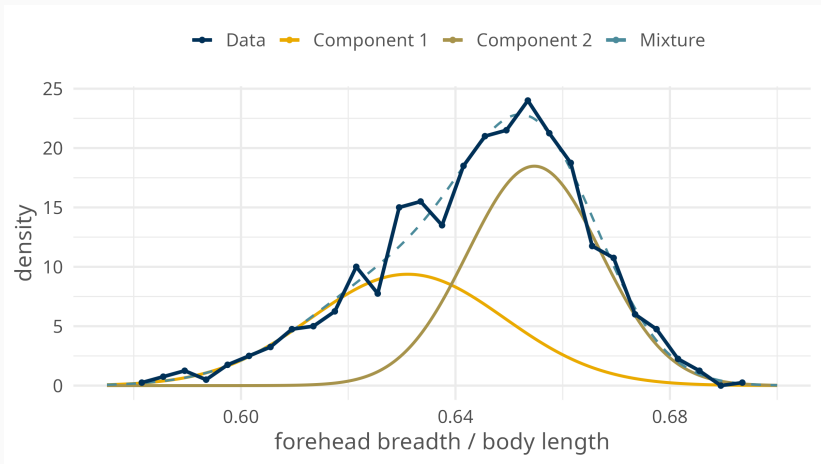
Then estimate $\lambda, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ by solving the polynomial system

$$E_{(\lambda, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2)}(X) = \frac{1}{1000} \sum_{i=1}^{1000} x_i$$

$$\vdots$$

$$E_{(\lambda, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2)}(X^5) = \frac{1}{1000} \sum_{i=1}^{1000} x_i^5.$$

Pearson's crabs - result



Moment varieties

Suppose $\mathcal{P} = \{P_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\}$ is a family of n -variate probability distributions (satisfying various conditions).

Definition

The d th *moment variety* of $\mathcal{P}$ is

$$\mathcal{M}_{\mathcal{P},d} = \overline{\{[(E_{\boldsymbol{\theta}}(\mathbf{X}^{\boldsymbol{\alpha}}))_{|\boldsymbol{\alpha}| \leq d}] : \boldsymbol{\theta} \in \Theta\}} \subset \mathbb{P}^{\binom{n+d}{d}-1}.$$

Mixtures

Definition

An r -mixture of $\mathcal{P}$ is a probability distribution of the form

$$\lambda_1 P_{\theta_1} + \cdots + \lambda_r P_{\theta_r}$$

for $\lambda \in \Delta_{r-1}$, $P_{\theta_1}, \dots, P_{\theta_r} \in \mathcal{P}$. Let $\mathcal{P}_r$ be the family of all r -mixtures of $\mathcal{P}$.

Fact

$$\mathcal{M}_{\mathcal{P}_r, d} = \sigma_r(\mathcal{M}_{\mathcal{P}, d})$$

Secant varieties

Definition

Given a projective variety $X \subset \mathbb{P}^N$, its r th *secant variety* is

$$\sigma_r(X) = \overline{\bigcup \{\langle p_1, \dots, p_r \rangle : p_i \in X\}}.$$

If $\dim \sigma_r(X) = \min\{r(\dim X + 1) - 1, N\}$, then $\sigma_r(X)$ is *nondefective*.

Algebraic identifiability

Definition

A family $\mathcal{P}$ of probability distributions is *algebraically identifiable* (from its moments of degree at most d) if the map $\overline{\Theta} \rightarrow \mathcal{M}_{\mathcal{P},d}$ is generically finite-to-one.

Fact

If $\mathcal{P}$ is algebraically identifiable, $\sigma_r(\mathcal{M}_{\mathcal{P},d})$ is nondefective, and $r(\dim X + 1) \leq \binom{n+d}{d} + 1$ (*subabundant*), then $\mathcal{P}_r$ is algebraically identifiable.

Pearson's crabs and algebraic identifiability

Theorem (Améndola, Ranestad, Sturmfels (2018))

If $\mathcal{P}$ is the family of univariate Gaussian distributions, then $\sigma_r(\mathcal{M}_{\mathcal{P},d})$ is always nondefective.

For example, $\sigma_2(\mathcal{M}_{\mathcal{P},5}) \subset \mathbb{P}^5$ is nondefective, and since $2(2+1) \leq 5+1$, subabundant. So 2-mixtures of univariate Gaussians are algebraically identifiable from their moments of degree at most 5.

This is why Pearson was able to solve the crab problem!

Location families

Definition

Suppose $\mathcal{P} = \{P_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \mathbb{R}^n\}$ is a family of probability distributions such that P_0 has density f . If the density of $P_{\boldsymbol{\theta}}$ is $\mathbf{x} \mapsto f(\mathbf{x} - \boldsymbol{\theta})$ for all $\boldsymbol{\theta}$, then $\mathcal{P}$ is a *location family*.

Theorem (T.)

If $\mathcal{P}$ is a location family, then $\mathcal{M}_{\mathcal{P},d}$ is projectively equivalent to the Veronese variety $\nu_d(\mathbb{P}^n)$.

Algebraic identifiability of location families

Theorem (Alexander, Hirschowitz (1995))

Given (d, n, r) , r -mixtures of n -variate location families are algebraically identifiable unless

- (a) $d = 2$, $n \geq 2$, and $r \in \left\{ 2, \dots, \left\lfloor \frac{1}{n+1} \binom{n+2}{2} \right\rfloor \right\}$,
- (b) $(d, n, r) \in \{(3, 4, 7), (4, 2, 5), (4, 4, 14)\}$, or
- (c) $r > \frac{1}{n+1} \binom{n+d}{d}$.

Definition

Let s be a nonnegative integer and $\mathcal{P} = \{P_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\}$ a family of probability distributions where $P_{\boldsymbol{\theta}}$ has density function $f(\cdot; \boldsymbol{\theta})$ (and satisfying various other conditions). If for a given $\boldsymbol{\theta}_0 \in \Theta$ and local mixing parameters $\boldsymbol{\gamma} = (\gamma_{\boldsymbol{\beta}})_{|\boldsymbol{\beta}| \leq s}$,

$$\mathbf{x} \mapsto \sum_{|\boldsymbol{\beta}| \leq s} \gamma_{\boldsymbol{\beta}} \partial_{\boldsymbol{\theta}}^{\boldsymbol{\beta}} f(\mathbf{x}; \boldsymbol{\theta}) \Big|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0}$$

is a probability density, then the corresponding distribution is an *s-local mixture* of $\mathcal{P}$ at $\boldsymbol{\theta}_0$. The family of all *s-local mixtures* of $\mathcal{P}$ is denoted $\mathcal{P}^{(s)}$.

Local mixtures of location families

Theorem (T.)

If $\mathcal{P}$ is a location family, then $\mathcal{M}_{\mathcal{P}^{(s)},d}$ is projectively equivalent to the s th osculating variety of the Veronese variety, $\tau_s(\nu_d(\mathbb{P}^n))$.

Theorem (Abo, Vannieuwenhoven (2018))

Given (d, n, r) , r -mixtures of 1-local mixtures of n -variate location families are algebraically identifiable unless

- (a) $d = 2$, $n \geq 6$, and $r \in \left\{ 2, \dots, \left\lfloor \frac{1}{2n+1} \binom{n+2}{2} \right\rfloor \right\}$,
- (b) $(d, n, r) = (3, 2, 2)$, or
- (c) $r > \frac{1}{2n+1} \binom{n+d}{d}$.

Terracini loci

Lemma (Terracini (1911))

If $p_1, \dots, p_r \in X$ are general, then

$$\dim \sigma_r(X) = \dim \langle T_{p_1} X, \dots, T_{p_r} X \rangle .$$

Definition

The r th *Terracini locus* of X , $\text{Ter}_r(X)$, is (the Zariski closure of) all the sets $\{p_1, \dots, p_r\} \subset X$ of smooth points for which

$$\dim \langle T_{p_1} X, \dots, T_{p_r} X \rangle < r(\dim X + 1) - 1.$$

Terracini loci of Veronese varieties

Theorem (Laface, Massarenti (2024))

If $r \leq \lfloor \frac{d+1}{2} \rfloor$, then $\text{Ter}_r(\nu_d(\mathbb{P}^n)) = \emptyset$.

Theorem (Galuppi, Santarsiero, T., Turatti (2025))

Suppose $n \geq 2$ and $d \geq 3$.

(a) If $r = \lfloor \frac{d+1}{2} \rfloor + 1$, then

$$\text{Ter}_r(\nu_d(\mathbb{P}^n)) = \overline{\{\text{sets of } r \text{ collinear points}\}}$$

(b) $\text{Ter}_4(\nu_3(\mathbb{P}^n)) = \overline{\{\text{sets of 4 coplanar points}\}}$

(c) If $d \geq 4$ and $r = \lfloor \frac{d+1}{2} \rfloor + 2$, then

$$\text{Ter}_r(\nu_d(\mathbb{P}^n)) = \overline{\{\text{sets of } r \text{ points with } r - 1 \text{ collinear}\}}$$

Terracini loci of tangential varieties of Veroneseans

An *intratangent pair* is a pair of points in $\tau(\nu_d(\mathbb{P}^n))$ lying on the same tangent space to $\nu_d(\mathbb{P}^n)$.

Theorem (T.)

Suppose $\Gamma \subset \tau(\nu_d(\mathbb{P}^n))$ is a set of r smooth points. If Γ contains an intratangent pair, then $\Gamma \in \text{Ter}_r(\tau(\nu_d(\mathbb{P}^n)))$.

Furthermore,

(a) If $d \geq 5$, then $\text{Ter}_2(\tau(\nu_d(\mathbb{P}^n))) = \overline{\{\text{intratangent pairs}\}}$.

(b) If $r \leq \frac{d+1}{3}$, then

$\text{Ter}_r(\tau(\nu_d(\mathbb{P}^1))) = \overline{\{r \text{ points w/ an intratangent pair}\}}$.

Terracini loci of tangential varieties of Veroneseans (cont.)

Theorem (T.)

Suppose $\Gamma = \{[\ell_1^{d-1}m_1], \dots, [\ell_r^{d-1}m_r]\} \subset \tau(\nu_d(\mathbb{P}^n))$ is a set of r smooth points.

- (a) *Suppose $d \geq 3$, $n \geq 2$, and $r > \frac{1}{2n} \left(\binom{n+d-2}{d-1} + \binom{n+d-1}{d} \right)$. If $[\ell_1], \dots, [\ell_r]$ lie on a hyperplane of $\mathbb{P}^n$, then $\Gamma \in \text{Ter}_r(\tau(\nu_d(\mathbb{P}^n)))$.*
- (b) *Suppose $d \geq 2$. If $\{[\ell_1^d], \dots, [\ell_r^d]\} \in \text{Ter}_r(\nu_d(\mathbb{P}^n))$, then $\Gamma \in \text{Ter}_r(\tau(\nu_d(\mathbb{P}^n)))$.*

Algebraic identifiability at a moment vector of a mixture

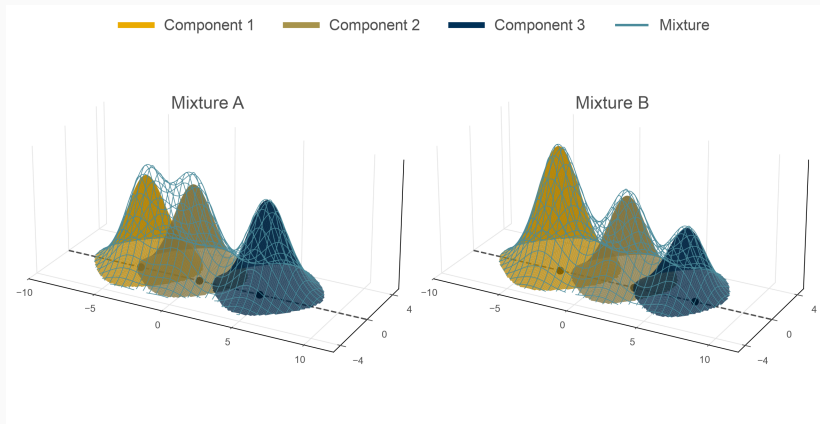
Definition

Suppose $q \in \mathcal{M}_{\mathcal{P}_r, d}$. If there are only finitely many linearly independent $\Gamma \subset \mathcal{M}_{\mathcal{P}, d}$ with $|\Gamma| = r$ and $q \in \langle \Gamma \rangle$, then q is *algebraically r -identifiable*.

Theorem (Ballico, Chiantini (2021))

If $q \in \mathcal{M}_{\mathcal{P}_r, d}$ is not algebraically r -identifiable, then there exist infinitely many $\Gamma \subset \mathcal{M}_{\mathcal{P}, d}$ with $q \in \langle \Gamma \rangle$ and $\Gamma \in \text{Ter}_r(\mathcal{M}_{\mathcal{P}, d})$.

A Terracini locus example



3-mixtures of 1-local mixtures of bivariate logistic distributions (w/ fixed scale parameter) $\Rightarrow$ moment vectors from collinear location parameters are in $\text{Ter}_3(\mathcal{M}_{\mathcal{P}(1),5})$.

 Thank you! 