Replicated measurements, ideal and derivatives

Roberto Notari

roberto.notari @ polimi.it

Dipartimento di Matematica, Politecnico di Milano

Introduction

The first part of this talk reports on the paper

–, E. Riccomagno, *Replicated measurements and algebraic statistics*, in Algebraic and Geometric Methods in Statistics, P. Gibilisco, E. Riccomagno, M.P. Rogantin, H. Wynn Eds., 2010, Cambridge University Press.

The second part is based on various papers by J. Elias, A. Iarrobino, V. Kanev, and M.E. Rossi.

Introduction

A basic application of algebraic statistics to design and analysis of experiments considers a design $\mathcal{D}$ as a finite set of distinct points in $\mathbb{R}^n$. This set can be equivalently described as the zero set of a system of polynomial equations, that is to say, of an ideal $I(\mathcal{D})$ in a polynomial ring. Then a subset of a basis of the quotient ring $R/I(\mathcal{D})$ is used as support for an identifiable regression model.

We consider this identifiability problem in the case where more than one measurement is taken at a design point. As we are after saturated regression models, this is essentially an interpolation problem.

Introduction

We focus on the case where a set of sample points $\omega_i \in \Omega$ are such that the corresponding design points $d(\omega_i)$ are unknown and identified with a single point d (error-in-variables models and random effect models).

Namely, consider clouds of points with unknown coordinates. Each cloud is close to a point d whose coordinates are known. The measured responses for each point in a cloud $y_i = y(d(\omega_i))$ are known.

We might include non replicated points as well.

Introduction

The problems we want to tackle are

I. determine a reasonable algebraic notion of a cloud of points close to a point $\leftrightarrow$ an analogue of $I(\mathcal{D})$;

II. determine conditions that ensure the good behavior of the interpolating polynomial $\leftrightarrow$ the analogue of $R/I(\mathcal{D})$.

In the second part of the talk, I will discuss the apolar correspondence and some of its applications.

Points with multiplicities

Let $\mathbb{K}$ be a field, and let $\mathbb{A}^n$ be the affine space over $\mathbb{K}$ of dimension n. If we fix a coordinate system, we can identify $\mathbb{A}^n$ with $\mathbb{K}^n$, and in particular, every point $P \in \mathbb{A}^n$ is represented by its coordinates $(a_1, \ldots, a_n)$ or equivalently, by its defining ideal $I(P) = \langle x_1 - a_1, \ldots, x_n - a_n \rangle$.

If we want to consider sets consisting of finitely many distinct points, e.g. $X = \{P_1, \ldots, P_r\}$, the defining ideals can be computed as

$$I(X) = I(P_1) \cap \cdots \cap I(P_r)$$

and consists exactly of the polynomials $f(x_1, \ldots, x_n)$ vanishing at all the points in X, i.e. $f(P_i) = 0$ for each $i = 1, \ldots, r$.

Points with multiplicities

In many situations, the computation of the coordinates of points is enough for solving specific problems, in some others, it is tto a poor information.

Example 1 *Consider the two intersection problems*

$$\begin{cases} y - x^2 = 0 \\ y = 1 \end{cases} \qquad \begin{cases} y - x^2 = 0 \\ y = 0 \end{cases}$$

The first system has $A(1,1)$ and $B(-1,1)$ as only solutions, and so we say that the conic and the line meet at the points A and B, and there is nothing more to say.
In the second case, we find that $O(0,0)$ is the only intersection point, but the line is tangent to the conic at the origin. So, the coordinates are not enough.

Points with multiplicities

Example 2 *Determine the polynomials vanishing at the origin, or vanishing at the origin with all their first derivatives.*

In the first case, as said before, we get exactly the polynomials in $I(O) = \langle x_1, \dots, x_n \rangle$.
In the second case, we get the subset of $I(O)$ consisting of all the polynomials having no linear part, i.e., if we write a polynomial as sum of homogeneous forms,

$$f = f_s + f_{s-1} + \cdots + f_1 + f_0, \quad \text{with } \deg(f_j) = j,$$

then f solves the second problem if and only if $f_1 = f_0 = 0$.
The solutions form the ideal I generated by all the degree 2 monomials.
We set $f_s = LF(f)$.

Points with multiplicities

Now, we give the definition of a point with multiplicity.

Definition 3 *An ideal I defines the point P with multiplicity m if there exists $k \in \mathbb{N}$ such that*

$$I(P)^k \subseteq I \subseteq I(P) \quad \textit{and} \quad \dim_{\mathbb{K}} R/I = m$$

where R/I is the quotient ring and $\dim_{\mathbb{K}}$ is the dimension as $\mathbb{K}$–vector space.

Equivalently, we say that I is an $I(P)$–primary ideal of degree m. The point P is called the support of I.

If we want to consider more that one point as support, it is enough to consider the intersection of the corresponding primary ideals.

Points with multiplicities

In the first example, the ideal associated to the intersection of $y = 0$ and $y - x^2 = 0$ is $I = \langle y, x^2 \rangle$ that defines the origin O with multiplicity 2, while in the second example, $I = I(O)^2$ and its multiplicity is $n + 1$, where n is the number of variables.

Also if we consider only one point as support, the degree m does not allow us to uniquely find the ideal unless $m = 1$. We give some examples in $\mathbb{K}[x, y]$ of multiple point supported at the origin.
$m = 2$ (double point): $I = \langle y, x^2 \rangle$ and $J = \langle y - x, x^2 \rangle$;
$m = 3$ (triple point): $I = \langle y, x^3 \rangle$ and $J = \langle x^2, xy, y^2 \rangle$.
$I(P)$ instead is the only ideal that defines P with $m = 1$.

Some algebraic tools

A tangent line to a curve can be seen as the limit position of a moving secant line. Equivalently, a double point can be seen as the limit position of two points that collapse to the same support. How to handle points that move?

Requirements:

• one more variable t to describe the movement;

• an ideal $J \subseteq \mathbb{K}[x_1, \ldots, x_n, t]$ that defines points whose coordinates depend on t;

• no point can appear or disappear during the movement.

Some algebraic tools

The requirements motivate the following definition.

Definition 4 *An ideal* $J \subseteq S = \mathbb{K}[x_1, \ldots, x_n, t]$ *defines a flat family of points if there exists* m *such that, for every* $t_0 \in \mathbb{K}$,

$$\dim_{\mathbb{K}} S/\langle J, t - t_0 \rangle = m.$$

For example, let $\mathbb{K} = \mathbb{R}$. Then, $J = \langle x^2 + y^2 - t^2, xy \rangle$ is a flat family.

For $t_0 \neq 0$, $\langle J, t - t_0 \rangle$ defines the 4 points $(\pm t_0, 0), (0, \pm t_0)$.

For $t_0 = 0$, $\langle J, t \rangle = \langle x^2 + y^2, xy, t \rangle$ defines the origin with multiplicity 4.

Some algebraic tools

In a flat family, for almost every $t_0 \in \mathbb{K}$, the corresponding set of points has the same geometric properties.
They are called the *general element* of the family.
The remaining $t_0 \in \mathbb{K}$ give rise to the *special elements* of the family.
In the previous example, the origin with multiplicity 4 was the special element of the family.

Some algebraic tools

In particular, it is possible to explicitly compute the special element of a family of points collapsing to one point along rays, given the ideal of the starting points.

Theorem 5 *Consider* $\mathcal{D} = \{P_1, \ldots, P_r\} \subseteq \mathbb{A}^n$ *a set of* r *distinct points, and let* $I(\mathcal{D})$ *be its defining ideal. Consider the flat family of points obtained by moving every* P_i *to the origin* O *along a straight line. Then, the special element is the origin with multiplicity* r *and it is defined by the* $I(O)$*–primary ideal*

$$I_0 = \{F \text{ homogeneous } | F = LF(f) \text{ for some } f \in I(\mathcal{D})\}.$$

I_0 is homogeneous.

Some algebraic tools

It is possible to generalize the previous theorem to a more general situation.

Theorem 6 *Assume* $X_i = \{P_{i1}, \ldots, P_{i,r_i}\}$ *is a set of distinct points that collapse to* A_i *for* $i = 1, \ldots, s$. *Assume further that* $X_1 \cup \cdots \cup X_s$ *is a set of* $r_1 + \cdots + r_s = r$ *distinct points. If* I_j *is the* $I(A_j)$*–primary ideal defined in the previous Theorem, then*

$$I_1 \cap \cdots \cap I_s$$

is the special element of the flat family of points obtained by moving all the r *points at the same time.*

Interpolation

As before, we first approach the problem in a special case, then we generalize it.
Let $\mathcal{D} = \{P_1, \ldots, P_r\} \subseteq \mathbb{A}^n$ be a set of distinct points, and let $y_1, \ldots, y_r \in \mathbb{K}$.

The multivariate interpolation problem can be stated as: find a polynomial $f \in R = \mathbb{K}[x_1, \ldots, x_n]$ such that $f(P_i) = y_i$ for each $i = 1, \ldots, r$.

The problem has a unique solution if we look for $f \in R/I(\mathcal{D})$ instead of $f \in R$. In fact, if f is a solution of the interpolation problem, then $f + g$ interpolates the same values for every $g \in I(\mathcal{D})$.

Interpolation

First step: collapse the points to the origin along rays, and study the coefficients of the interpolating polynomial on the distinct points. Of course, we choose to vary the computed value y_i at P_i by means of a polynomial $y_i(t)$.

Theorem 7 *Let* $X = \{P_1, \ldots, P_r\} \subseteq \mathbb{A}^n$ *be a set of distinct points, and let* $y_1, \ldots, y_r \in \mathbb{K}$. *Let* $M_1 = 1, \ldots, M_r$ *be a monomial base of* $R/I(X)$ *and assume that* $\deg(M_i) = d_i$ *with* $0 = d_1 \leq \cdots \leq d_r$. *Let* $y_i(t) \in \mathbb{K}[t]$ *verify* $y_i(1) = y_i$ *for each* $i = 1, \ldots, r$. *Then, there exists a unique interpolating polynomial* $F = c_1 M_1 + \cdots + c_r M_r$ *with* $c_i(t) \in \mathbb{K}[t]_t$ *such that* $F(t_0 P_i) = y_i(t_0)$ *for every* i *and* $t_0 \neq 0$.

$c_i \in \mathbb{K}[t]_t$ means that c_i is a ratio whose denominator is t^d.

Interpolation

Second step: look for sufficient conditions on the $y_i(t)$ to assure that $c_i(t)$ can be evaluated at $t = 0$. The evaluation is possible if $c_i \in \mathbb{K}[t]$, or, if $\mathbb{K} = \mathbb{R}$, if $\lim_{t\to 0} c_i(t)$ exists and it is finite.

Theorem 8 *Let A be the $r \times r$ matrix such that $a_{ij} = M_j(P_i)$ evaluation of the monomial M_j at the point P_i. Moreover, let*

$$\overline{y} = (y_1(t), \dots, y_r(t))^T = y^0 + ty^1 + \dots + t^b y^b$$

with $y^i \in \mathbb{K}^r$. Then, $c_i(t) \in \mathbb{K}[t]$ if and only if $y^j \in$ Span$\langle A_h | d_h \leq j \rangle$ where A_h is the h–th column of A.

The advantage of $y_i(t)$ to be a polynomial is that the condition is not only sufficient, but it is also necessary.

Interpolation

Third step: as in the computation of the primary ideal when collapsing a cloud of points at the origin, now we want to explicitly compute the limit of the interpolating polynomial.

Theorem 9 *In the hypotheses and notation of previous Theorems,*

$$c_i(0) = \frac{\det(A_{i,d_i})}{\det(A)}$$

where A_{i,d_i} is the matrix obtained from A by substituting its i–th column with y^{d_i}.

Hence, also in this case, the new variable t plays no role. As last remark, $M_1 = 1$, and so $y^0 = y_0 M_1$, i.e. $c_1(0) = y_0$ for some $y_0 \in \mathbb{K}$ arbitrarily chosen.

Interpolation

Last step: we want to collapse the starting points to more centers, and we want to get one interpolating polynomial fitting both the initial data, and the movement of the points.

Theorem 10 *In the set-up of Theorem 6, let* I_j *the* $I(A_j)$*–primary ideal obtained by collapsing* X_j *to* A_j *and let* $I = I_1 \cap \cdots \cap I_s$*. Let* $F_j \in R/I_j$ *be the limit interpolating polynomial computed in Theorem 9, for* $j = 1, \ldots, s$*. Then, there exists a unique polynomial* $F \in R/I$ *such that* $F \mod I_j = F_j$ *for every* j*.*

This result is a sort of gluing of partial interpolators. Once again, all the computations can be performed without using the extra variable t.

Projection to the support

In most practical cases, after various observations are taken over each point of the design, one is interested in determining a saturated, linear, regression model identifiable by the design. Hence, we start comparing the rings R/I and $R/I(Y)$ where $Y = \{A_1, \dots, A_s\}$. Of course, Y is the design, while I describes the intersection of the $I(A_j)$–primary ideals obtained by collapsing various points at each A_j.

At first, we show that the comparison is possible.

Theorem 11 *The inclusion $I \subseteq I(Y)$ induces a surjective map*

$$\psi : \frac{R}{I} \to \frac{R}{I(Y)}$$

defined as $\psi(F) = F \mod I(Y)$.

Projection to the support

Now, we adapt the previous Theorem to our situation.

Theorem 12 *Let* $F_j \in R/I_j$ *be the limit interpolating polynomial for* $j = 1, \ldots, s$, *and let* $F \in R/I$ *be the gluing of* $F_1, \ldots, F_s$. *Let* $G \in R/I(Y)$ *be the only solution of the interpolating problem* $G(A_j) = F_j(A_j)$ *for* $j = 1, \ldots, s$. *Then,* $\psi(F) = G$.

Often, $F_j(A_j)$ is the mean value of the observations over each design point, but the results hold true also for different choices of $F_j(A_j)$.

Primary ideals and derivatives

Now, we want to describe the apolar correspondence, that allows to describe primary ideals by means of derivatives. Let $\mathbb{K}$ be a field of characteristic 0, and let $R = \mathbb{K}[x_1, \ldots, x_n], S = \mathbb{K}[y_1, \ldots, y_n]$ be polynomial rings in the same number of indeterminates. We define a multiplication between a monomial $y_1^{a_1} \ldots y_n^{a_n} \in S$ and a monomial $x_1^{b_1} \ldots x_n^{b_n} \in R$ as

$$y_1^{a_1} \ldots y_n^{a_n} \circ x_1^{b_1} \ldots x_n^{b_n} = \frac{b_1!}{(b_1 - a_1)!} \cdots \frac{b_n!}{(b_n - a_n)!} x_1^{b_1 - a_1} \ldots x_n^{b_n - a_n}$$

if $b_i \geq a_i$ for each $i = 1, \ldots, n$, and $y_1^{a_1} \ldots y_n^{a_n} \circ x_1^{b_1} \ldots x_n^{b_n} = 0$ otherwise. Essentially, we think of y_i as $\partial/\partial x_i$.

Primary ideals and derivatives

This new product can be extended from monomials to polynomials by using the distributive law.
In algebraic terms, the operation $\circ$ makes R an S–module.
This means that in R the product of polynomials cannot be considered.

Definition 13 *Let $g \in S$ and $f \in R$ be polynomials. We say that g and f are apolar each other if $g \circ f = 0$.*

Easy generalization

1. $(I)^{\perp} = \{f \in R \mid g \circ f = 0$ for every $g \in I\}$ given the ideal $I \subseteq S$; $(I)^{\perp}$ is a submodule of R.
2. $(M)^{\perp} = \{g \in S \mid g \circ f = 0$ for every $f \in M\}$ given the submodule $M \subseteq R$; $(M)^{\perp}$ is an ideal of S.

Primary ideals and derivatives

The S–module R is not Noetherian, i.e. there are submodules of R that are not finitely generated. For example, consider $M = \langle x_1^n \mid n \in \mathbb{N} \rangle$.

Under the standard grading, $S_d \times R_d \to \mathbb{K}$ identifies S_d with the dual of R_d, because it is a non–degenerate bilinear map.

First basic result result on the apolar correspondence.

Theorem 14 *There is a 1-to-1 correspondence between Artinian homogeneous ideals in S and finitely generated, graded, S–submodules of R.*

Artinian homogeneous ideals in S represent $I(O)$–primary ideals in S, i.e. the origin with some multiplicity.

Primary ideals and derivatives

We want to translate operations on ideals in S to operations on submodules in R.

1. $I \subseteq J$ if, and only if, $I^\perp \supseteq J^\perp$.
2. $\dim_{\mathbb{K}} \left(\frac{J}{I}\right)_t = \dim_{\mathbb{K}} \left(\frac{I^\perp}{J^\perp}\right)_t$ for each $t \in \mathbb{Z}$.
3. $(I \cap J)^\perp = I^\perp + J^\perp$;
4. $(I + J)^\perp = I^\perp \cap J^\perp$;
5. $(I : J)^\perp = J \circ I^\perp$, for whatever homogeneous ideal J;
6. I is monomial if, and only if, $I^\perp$ is monomial.

Moreover, $I^\perp$ is isomorphic to the canonical module of S/I.

Gorenstein ideals

A particular class of multiple points are the Gorenstein ones, that we want to define and characterize in terms of apolarity.

Let $I \subseteq S$ a homogeneous Artinian ideal. We define the *socle* of S/I as the ideal $Soc(S/I) = 0 :_{S/I} I(O)$ of S/I.

Definition 15 *We say that either I or S/I is Gorenstein if* $\dim_{\mathbb{K}} Soc(S/I) = 1.$

The apolar submodule of a Gorenstein ideal is quite simple. In fact,

Theorem 16 *I is Gorenstein if, and only if, $I^{\perp}$ is cyclic.*

This result is known as Macaulay' s correspondence. Hence, to construct Gorenstein multiple points it is enough to choose a homogeneous polynomial in R and compute its apolar ideal.

Gorenstein ideals

Gorenstein Artinian homogeneous ideals are then the basic bricks to construct every other Artinian homogeneous ideal. In fact, it holds

Theorem 17 *Every Artinian ideal I is the intersection of finitely many Artinian Gorenstein ideals.*

Let $g_1, \dots, g_s$ be the largest degree generators of $I^\perp$. Then, the ideal $J = \cap_{i=1}^{s} (g_i)^\perp$ is uniquely determined by I. It would be interesting to deeply investigate the relation between the ideals I and J.

What about the monomial Gorenstein Artinian ideals in S?

Theorem 18 *Let $I \subseteq S$ be an Artinian monomial ideal. Then, I is Gorenstein if, and only if, $(I)^\perp = \langle x_1^{b_1} \dots x_n^{b_n} \rangle$, i.e. I is the complete intersection ideal generated by $y_i^{b_i+1}$.*

Gorenstein ideals

Assume $I = (f)^{\perp}$ for some $f \in R_d$. Then,

$$\dim_{\mathbb{K}} \left(\frac{S}{I}\right)_t = \dim_{\mathbb{K}} \left(\frac{S}{I}\right)_{d-t}$$

for every $t \in \mathbb{Z}$, i.e., there is a symmetry around $d/2$. In particular, $\dim_{\mathbb{K}}(S/I)_d = 1$ and $\dim_{\mathbb{K}}(S/I)_t = 0$ for $t > d$.

Theorem 19 *Let $f \in R_d$ be a non–zero form of degree $d \geq 1$. Then, $f^{\perp}$ has minimal generators in degree $> d$ if, and only if, $f = \ell^d$ for a suitable linear form $\ell \in R$, or equivalently, $\dim_{\mathbb{K}}(S/I)_t = 1$ for every $0 \leq t \leq d$.*

Gorenstein ideals

Assume $d = 2h + 1$, and let $f \in R_d$ be a polynomial. The ideal $(f)^\perp$ defines the origin with multiplicity $m = \sum_{t=0}^{d} \dim_{\mathbb{K}} (S/I)_t$. For a general choice of f, we get $m = 2\binom{n+h}{n}$.

The parameter space for Gorenstein ideals has dimension $\dim_{\mathbb{K}} R_d - 1 = \binom{2h+n}{n} - 1$.

The parameters to determine m distinct points in $\mathbb{A}^n$ are $mn = 2n\binom{n+h}{n}$.

Theorem 20 *Not every primary ideal is limit of distinct points.*

Proof. Asymptotically, $\binom{2h+n}{n} - 1 > 2n\binom{n+h}{n}$, and so there are too many Gorenstein ideals with respect to the set of m distinct points. □