Indam

Numerical Solution of Stochastic Partial Differential Equations

Politecnico di Torino, May 10-13, 2010

Computer Commutative Algebra and Polynomials in Normal Variables

Giovanni Pistone `giovanni.pistone@gmail.com`

POLITECNICO DI TORINO Collegio Carlo Alberto

Eva Riccomagno `riccomag@dima.unige.it`

UNIVERSITÀ DEGLI STUDI DI GENOVA

May 11, 2010

Abstract and a reference

Computational methods based on polynomial algebra software such as

- CoCoATeam, CoCoA*: a system for doing Computations in Commutative Algebra*, Available at cocoa.dima.unige.it,

have been used in Statistics for Design of Experiments DoE and for various problems in statistical modeling. A recent overview of this new field, termed Algebraic Statistics, is in

- P. Gibilisco, E. Riccomagno, M. Rogantin, H.P. Wynn, eds., *Algebraic and Geometric Methods in Statistics* (Cambridge University Press, 2009).

In this approach to DoE, the set of design points is described as the solution of a system of polynomial equations and the identification of various classes of models is computed by the use of special bases of the ideal generated. Here we present the first results of a research in progress in which we explore the applicability of these ideas when the defining equations are derived from Hermite polynomials, e.g. the system is

$$\boxed{x^3 - 3x = 0, y^3 - 3y = 0, x^2 - 1 = y^2 - 1}$$

Hermite polynomials

Definition

1. Define $\delta f(x) = xf(x) - f'(x) = -e^{x^2/2}\frac{d}{dx}\left(f(x)e^{-x^2/2}\right)$. If $Z \sim \mathcal{N}(0,1)$,
$$\mathrm{E}\left(g(Z)\delta f(Z)\right) = \mathrm{E}\left(dg(Z)f(Z)\right),$$
i.e. δ is the transpose of the derivative w.r.t. the standard Gaussian measure.
2. Define $H_0 = 1$, $H_n(x) = \delta^n 1$, $n > 0$, e.g.
$$H_1(x) = x, H_2(x) = x^2 - 1, H_3(x) = x^3 - 3x, H_4(x) = x^4 - 6x^2 + 3, \ldots$$

Properties

1. The transposition formula shows that the H_n's are orthogonal.
2. $d\delta - \delta d = \mathrm{id}$, $dH_n = nH_{n-1}$, $H_{n+1} = xH_n - nH_{n-1}$.

- P. Malliavin, *Integration and probability*, Vol. 157 of *Graduate Texts in Mathematics* (Springer-Verlag, New York, 1995), ISBN 0-387-94409-5, with the collaboration of Hélène Airault, Leslie Kay and Gérard Letac, Edited and translated from the French by Kay, With a foreword by Mark Pinsky

Zeros of H_n.

Theorem

1. *Each Hermite polynomial H_n, $n \geq 1$, has n distinct real roots.*
2. *The roots of H_{n+1} are separated by the roots of H_n, $n \geq 1$.*

Theorem

1. $H_k H_n = H_{n+k} + \sum_{i=1}^{n\wedge k} \binom{n}{i}\binom{k}{i} i! H_{n+k-2i}$, $n, k \geq 1$.
2. $\mathrm{E}(H_n^2(Z)) = n!$, $n \geq 0$.
3. *If* $H_n(x) = 0$, *then* $H_{n+k}(x) + \sum_{i=1}^{n\wedge k} \binom{n}{i}\binom{k}{i} i! H_{n+k-2i}(x) = 0$, $n \geq 1$.

In statistical language, item 3 shows an aliasing relation on the design $\mathcal{D} = \{x \colon H_n(x) = 0\}$.

- W. Gautschi, *Orthogonal polynomials: computation and approximation*, Numerical Mathematics and Scientific Computation (Oxford University Press, New York, 2004), ISBN 0-19-850672-4, oxford Science Publications

Algebraic DoE: basics

- Given univariate polynomials $f_1(x_1), \dots, f_m(x_m) \in \mathbb{Q}[x_1, \dots, x_m]$, we consider the design ideal
$$\mathrm{Ideal}\,(f_1(x_1), \dots, f_m(x_m)) = \left\{ \sum_{i=1}^{m} a_i f_i \colon a_i \in \mathbb{Q}[x_1, \dots, x_m] \right\}.$$
- We assume all zeros to be real and simple; they form the full design $\mathcal{D}$.
- Two polynomials h, k, are aliased if $h - k$ is zero on $\mathcal{D}$, i.e. if $h - k$ belong to the design ideal.
- A fraction is a subset $\mathcal{F}$ of $\mathcal{D}$. It is obtained by adding new equations $g_1, \dots, g_l$, called defining equations, to the design ideal.
- The indicator polynomial F of the fraction $\mathcal{F}$ is a polynomial whose restriction to $\mathcal{D}$ is the indicator function of the fraction.
- The main interest of this setting is the availability of symbolic software for the computation of ideals in the ring $\mathbb{Q}[x_1, \dots, x_m]$, e.g. CoCoA, `4ti2`, Singular, Maple, Matematica, Maxima, Macaulay2, . . . (subjective order).

- G. Pistone, E. Riccomagno, H.P. Wynn, *Algebraic statistics. Computational commutative algebra in statistics*, Vol. 89 of *Monographs on Statistics and Applied Probability* (Chapman & Hall/CRC, Boca Raton, FL, 2001), ISBN 1-58488-204-2.

Gröbner basis, normal form

Definition

1. A term-order is a total order on terms x^α compatible with the product. Given a term-order, the leading term $\mathrm{LT}(f)$ of each polynomial f is defined and each polynomial is an ordered list of coefficients.
2. A finite subset $\{g_1, \dots, g_r\}$ of an ideal I is a Gröbner basis if, and only if, the leading terms $\mathrm{LT}(g_i)$, $i = 1, \dots, r$, generate the leading terms of I.

Theorem

1. *Given a term ordering and an ideal I, a Gröbner basis $g_1, \dots, g_r$ can be computed by a finite (and highly complex) algorithm.*
2. *For each polynomial f there exist a unique polynomial r such that $f - r \in I$ and none of its terms is divided by any $\mathrm{LT}(g_i)$'s.*
3. *Such a remainder r is called normal form of f, $\mathrm{NF}(f) = r$.*

- D. Cox, J. Little, D. O'Shea, *Ideals, varieties, and algorithms: An introduction to computational algebraic geometry and commutative algebra*, Undergraduate Texts in Mathematics, 2nd edn. (Springer-Verlag, New York, 1997), ISBN 0-387-94680-2

CoCoA: Monomial basis `1.coc`

Given the input:

```
Use R::= Q[x,y];  --- the ring, standar monomial order x>y
L:=[x^3-3x,y^3-3y,x^2-y^2]; --- the list of polynomials
I:=Ideal(L); --- the ideal generated by the list
GBasis(I); --- computation of the Groebner basis
QuotientBasis(I); --- vector basis of Q[x,y]/I
```

the Gröbner basis of the ideal is

$$x^2 - y^2, y^3 - 3y, -xy^2 + 3x,$$

and the monomial vector basis is

$$\frac{\mathbb{Q}[x,y]}{\text{Ideal}\left(x^3-3x, y^3-3y, x^2-y^2\right)} = \text{Span}\left(1, y, y^2, x, xy\right)$$

CoCoA: The monomial order 2.coc

The monomial basis depends on the monomial order.

```
Use R::= Q[y,x], Lex;  --- the ring, standard monomial order y>x
L:=[x^3-3x,y^3-3y,x^2-y^2]; --- the list of polynomials
I:=Ideal(L); --- the ideal generated by the list
GBasis(I); --- computation of the Groebner basis
QuotientBasis(I); --- vector basis of Q[y,x]/I
```

gives the Gröbner basis

$$-y^2 + x^2, x^3 - 3x, -yx^2 + 3y$$

and the monomial basis

$$1, x, x^2, y, yx$$

Not all monomial bases are obtained this way

- G. Pistone, E. Riccomagno, M. Rogantin, in *Search for Optimality in Design and Statistics: Algebraic and Dynamical System Methods*, edited by L. Pronzato, A. Zhigljavsky (Springer-Verlag, 2009), Number 28 in Optimizations and its applications, pp. 97–132

CoCoA: Indicator polynomial `3.coc`

- $x^3 - 3x = 0, y^3 - 3y = 0$ is the full design; $x^2 - y^2 = 0$ is the generating equation.
- $1 - f = h(x^2 - y^2)$ means $1 = f$ if the generating equation holds.
- $f(x^2 - y^2) = 0$ means $f = 0$ if the generating equation is violated.
- The h-elimination ideal of $I = \text{Ideal}\left(x^3 - 3x, y^3 - 3y, 1 - f - h(x^2 - y^2), f(x^2 - y^2)\right)$ is $I \cap \mathbb{Q}[f, y, x]$

```
Use R::= Q[h,f,y,x], Lex;  Lex monomial order
L:=[x^3-3x,y^3-3y,1-f-h(x^2-y^2),f(x^2-y^2)];
I:=Ideal(L); --- the ideal generated by the list
J:=Elim(h,I); --- elimination ideal
ReducedGBasis(J); --- elimination ideal in Lex order
```

produces

```
x^3 - 3x,
y^3 - 3y,
f - 2/9y^2x^2 + 1/3y^2 + 1/3x^2 - 1
```

where the last equation is the indicator polynomial.

Aliasing computation

- The computation of the normal form introduces a notion of confounding. For example from $H_{n+1}(x) = xH_n(x) - nH_{n-1}(x)$ and for $\equiv$ meaning equality holds over $\mathcal{D}_n = \{x : H_n(x) = 0\}$, we obtain $H_{n+1}(x) \equiv -nH_{n-1}(x)$.
- In general let $H_{n+k} \equiv \sum_{j=0}^{n-1} h_j^{n+k} H_j$ be the representation of H_{n+k} at $\mathcal{D}_n$. Substituting in the product formula gives

$$\begin{aligned}\mathrm{NF}(H_{n+k}) &\equiv -\sum_{i=1}^{n\wedge k} \binom{n}{i}\binom{k}{i} i!\,\mathrm{NF}(H_{n+k-2i}) \\ &= -\sum_{i=1}^{n\wedge k} \binom{n}{i}\binom{k}{i} i! \sum_{j=0}^{n-1} h_j^{n+k-2i} H_j\end{aligned}$$

Equating coefficients gives a general recursive formula

$$h_j^{n+k} = -\sum_{i=1}^{n\wedge k} \binom{n}{i}\binom{k}{i} i!\, h_j^{n+k-2i}$$

The first confounding relationships are

k	expansion
1	$-nH_{n-1}$
2	$-n(n-1)H_{n-2}$
3	$-n(n-1)(n-2)H_{n-3} + 3nH_{n-1}$
4	$-n(n-1)(n-2)(n-3)H_{n-4} + 8n(n-1)H_{n-2}$
5	$-\frac{n!}{(n-5)!}H_{n-5} + 5nH_{n-1} + 15n(n-1)(n-2)H_{n-3}$
6	$-\frac{n!}{(n-6)!}H_{n-6} + 24n(n-1)(n-2)(n-3)H_{n-4} + 10n(n-1)(2n-5)H_{n-2}$

- For $f = \sum_{i=0}^{n+1} c_i(f)H_i$, we have $k = 1$ and

$$\begin{aligned} \mathrm{NF}(f) &= \sum_{i=0}^{n-1} c_i(f)H_i + \underline{c_n(f)H_n} + c_{n+1}(f)\,\mathrm{NF}(H_{n+1}) \\ &\equiv \sum_{i=0}^{n-2} c_i(f)H_i + \left(c_{n-1}(f) - nc_{n+1}(f)\right)H_{n-1} \end{aligned}$$

Expectation and NF

- Let f be a polynomial in one variable with real coefficients and by polynomial division $f(x) = q(x)H_n(x) + r(x)$ where r has degree smaller than H_n and $r(x) = f(x)$ on $H_n(x) = 0$. The $n-1$ degree polynomial r is the remainder or normal form $\mathrm{NF}(f) = r$.
- Then

$$\begin{aligned} \mathrm{E}\left(f(Z)\right) &= \mathrm{E}\left(q(Z)H_n(Z)\right) + \mathrm{E}\left(r(Z)\right) \\ &= \mathrm{E}\left(q(Z)\,\delta 1^n\right) + \mathrm{E}\left(r(Z)\right) \\ &= \mathrm{E}\left(d^n q(Z)\right) + \mathrm{E}\left(r(Z)\right) = \mathrm{E}\left(r(Z)\right) \qquad \text{iff } \mathrm{E}\left(d^n q(Z)\right) = 0. \end{aligned}$$

- Note that $d^n q(Z) = 0$ if and only if q has degree smaller than n and this is only if f has degree smaller or equal to $2n-1$. But also

$$\mathrm{E}\left(d^n q(Z)\right) = \mathrm{E}\left(d^n \sum_{i=0}^{\infty} c_i(q) H_i\right)$$

$$= \langle H_n, \sum_{i=0}^{\infty} c_i(q) H_i \rangle = n! c_n(q) = 0$$

iff $c_n(q) = 0$.

A bi-dimensional example

The previous argument generalizes to fractions in the multivariate case.

- Let $\mathcal{F}$ be the zero set of the Gröbner basis

$$\begin{cases} g_1 = x^2 - y^2 & = H_2(x) - H_2(y) = 0 \\ g_2 = y^3 - 3y & = H_3(y) = 0 \\ g_3 = xy^2 - 3x & = H_1(x)\left(H_2(y) - 2H_0\right) = 0 \end{cases}$$

- For f polynomial there exists unique r in

$$\operatorname{Span}\left(H_0, H_1(x), H_1(y), H_1(x)H_1(y), H_2(y)\right) = \operatorname{Span}\left(1, x, y, xy, y^2\right)$$

  s.t. $f = \sum q_i g_i + r$.

- If

$$\begin{aligned} q_1(x,y) &= a_o + a_1 H_1(x) + a_2 H_1(y) + a_3 H_1(x) H_1(y), \\ q_2(x,y) &= \theta_1(x) + \theta_2(x) H_1(y) + \theta_3(x) H_2(y), \\ q_3(x,y) &= a_4 + a_5 H_1(y) \end{aligned}$$

  then

$$\mathrm{E}(f(Z_1, Z_2)) = \mathrm{E}(r(Z_1, Z_2))$$

Gaussian quadrature

- For $k = 1, \dots, n$ and $x_1, \dots, x_n \in \mathbb{R}$ pairwise distinct, define the Lagrange polynomials $l_k(x) = \prod_{i:i\neq k} \frac{x-x_i}{x_k-x_i}$. These are indicator polynomial functions of degree $n-1$, namely $l_k(x_i) = \delta_{ik}$, and form a $\mathbb{R}$-vector space basis of the set of polynomials of degree at most $(n-1)$, $\mathbb{P}_{n-1}$.
- If r has degree smaller than n then $r(x) = \sum_{k=1}^n r(x_k) l_k(x)$ and for $\lambda_k = \mathrm{E}(l_k(Z))$ by linearity $\mathrm{E}(r(Z)) = \sum_{k=1}^n r(x_k)\,\mathrm{E}(l_k(Z)) = \sum_{k=1}^n r(x_k)\lambda_k$.
- Putting all together, on $\mathcal{D}_n = \{x : H_n(x) = 0\} = \{x_1, \dots, x_n\}$ and for f polynomial of degree at most $(2n-1)$ or s.t. $c_n(\frac{f-r}{H_n}) = 0$,

$$\mathrm{E}\left(f(Z)\right) = \mathrm{E}\left(r(Z)\right) = \sum_{k=1}^n r(x_k)\,\mathrm{E}\left(l_k(Z)\right)$$

$$= \sum_{k=1}^n f(x_k)\,\mathrm{E}\left(l_k(Z)\right) = \mathrm{E_n}\left(f(X)\right),$$

where $\mathrm{P_n}\left(X = x_k\right) = \mathrm{E}\left(l_k(Z)\right) = \lambda_k$ is a probability on $\mathcal{D}$.

Algebraic computation of the weights λ_k

Theorem

Let λ *be the polynomial of degree* $n-1$ *such that* $\lambda(x_k) = \lambda_k$ *then*

$$\lambda(x) H_{n-1}^2(x) = \frac{(n-1)!}{n} \qquad \text{on } H_n = 0.$$

- E.g. for $n = 3$

$$0 = H_3(x) = x^3 - 3x$$
$$2/3 = \lambda(x) H_2^2 = (\theta_0 + \theta_1 x + \theta_2 x^2)(x^2 - 1)^2$$

reduce degree using $x^3 = 3x$ and equate coefficients to obtain

$$\lambda(x) = \frac{2}{3} - \frac{x^2}{6}$$

Evaluate to find $\lambda_{-\sqrt{3}} = \lambda(-\sqrt{3}) = \frac{1}{6} = \lambda_{\sqrt{3}}$ and $\lambda_0 = \lambda(0) = \frac{2}{3}$.

- The roots of H_n, $n > 2$, are not in $\mathbb{Q}$. Computer algebra systems work with rational fields. Working with algebraic extensions of fields could be slow.
- Sometimes there is no need to compute explicitly the weights.

A CoCoA code for the weighing polynomial

```
N:=4;                                    -- number of nodes
Use R::=Q[w,h[1..(N-1)]], Elim(w);       -- setting up the ring
Eqs:=[h[2]-h[1]*h[1]+1];                 -- the Hermite polynomials
For I:=3 To N-1 Do
  Append(Eqs,h[I]-h[1]*h[I-1]+(I-1)*h[I-2]) EndFor;
  Append(Eqs,h[1]*h[N-1]-(N-1)*h[N-2]); -- the nodes
Set Indentation;
Append(Eqs,N*w*h[N-1]^2-Fact(N-1));      -- the weighing polynomial
J:=Ideal(Eqs); GB_J:=GBasis(J);          -- the game
Last(GB_J);
```

The output is `3w + 1/4h[2] - 5/4`. Hence, $w(x) = \frac{5-h2}{12} = \frac{6-x^2}{12}$ and for $H_4(x) = x^4 - 6x^2 + 3 = 0$,

x	$-\sqrt{3-\sqrt{6}}$	$-\sqrt{3+\sqrt{6}}$	$\sqrt{3-\sqrt{6}}$	$\sqrt{3+\sqrt{6}}$
$w(x)$	$\frac{3+\sqrt{6}}{12}$	$\frac{3-\sqrt{6}}{12}$	$\frac{3+\sqrt{6}}{12}$	$\frac{31\sqrt{6}}{12}$

NF and orthogonal projection

Remark

1. Let $f(x)$ be a polynomial and $f(x) = q(x)H_n(x) + r(x)$ where q, r are unique with r of degree less than n. Let $Z \sim \mathcal{N}(0,1)$. Then q is a polynomial such that $m \geq n$

$$\mathrm{E}\left((f(Z) - q(Z)H_n(Z))H_m(Z)\right) = 0$$

2. Can be generalized for general fractions, i.e.

$$f = \sum_i q_i g_i + \mathrm{NF}(f) \quad g_i \text{ Gröbner basis}$$

- r has degree at most $n-1$, then $r(x) \in \mathrm{Span}\left(H_0, H_1, \ldots, H_{n-1}\right)$. In particular r is orthogonal to H_m for all $m \geq n$.
- Let there exist q_1 and q_2 distinct such that $f - q_1 H_n \perp H_m$ and $f - q_2 H_n \perp H_m$ for all $m \geq n$. Now $(q_1 - q_2)H_n$ is 0 or has degree not smaller than n. Furthermore it is orthogonal to H_m for all $m \geq n$. Necessarily it is 0, equivalently $q_1 = q_2$.

Fractions: $\mathcal{F} \subset \mathcal{D}_n$, $\#\mathcal{F} = m < n$

- Let $1_{\mathcal{F}}(x)$ be the polynomial of degree n such that $1_{\mathcal{F}}(x) = 1$ if $x \in \mathcal{F}$ and 0 if $x \in \mathcal{D}_n \setminus \mathcal{F}$ and let f be polynomial of degree at most $n-1$ and let $Z \sim \mathcal{N}(0,1)$. Then for $\mathrm{P_n}(X = x_k) = \lambda_k$

$$\mathrm{E}((f1_{\mathcal{F}})(Z)) = \sum_{x_k \in \mathcal{F}} f(x_k)\lambda_k = \mathrm{E_n}\left(f(X)1_{\mathcal{F}}(X)\right) = \mathrm{E_n}\left(f(X)|X \in \mathcal{F}\right)\mathrm{P_n}(X \in \mathcal{F}$$

- Let $\omega_{\mathcal{F}}(x) = \prod_{x_k \in \mathcal{F}}(x - x_k) = \sum_{i=0}^{m} c_i H_i(x)$ and note $l_k^{\mathcal{F}}(x) = \prod_{i \in \mathcal{F}, i \neq k} \frac{x - x_i}{x_k - x_i}$ $= \mathrm{NF}(l_k(x), \mathrm{Ideal}(\omega_{\mathcal{F}}(x))$ are the Lagrange polynomials for $\mathcal{F}$. For f a polynomial of degree N, write $f(x) = q(x)\omega_{\mathcal{F}}(x) + r(x)$ with $f(x_i) = r(x_i)$ on $\mathcal{F}$ and $r(x) = \sum_{x_k \in \mathcal{F}} f(x_k) l_k^{\mathcal{F}}(x)$. Let $q(x) = \sum_{j=0}^{N-m} b_j H_j(x)$. Then

$$\mathrm{E}\left(f(Z)\right) = \mathrm{E}\left(\sum_{j=0}^{N-m} b_j H_j(Z) \sum_{i=0}^{m} c_i H_i(Z)\right) + \mathrm{E}\left(r(Z)\right)$$

$$= b_0 c_0 + b_1 c_1 + \ldots + ((N-m) \wedge m)! \, b_{(N-m)\wedge m} c_{(N-m)\wedge m} + \sum_{x_k \in \mathcal{F}} f(x_k)\lambda_k^{\mathcal{F}}$$

where $\lambda_k^{\mathcal{F}} = \mathrm{E}\left(\mathrm{NF}(l_k(x), \mathrm{Ideal}(\omega_{\mathcal{F}}(x)))\right)$.