Hilbert series and polynomial models for Smolyak-type sparse grid designs

Henry Wynn, London School of Economics

h.wynn@lse.ac.uk

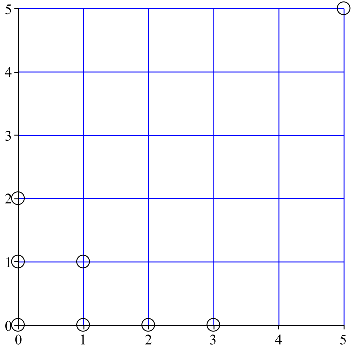
CIRM, May 2017
Joint with Hugo Maruri, QMUL

The background from stochastic FE

- Special grid designs are used in stochastic finite elemenst methods
- They are used to build interpolators and for quadrature for polynomials
- The input variable is a "parameter" which may eventually be random
- In that case they are sometimes called polynomial chaos expansions (PCE)
- Quadrature is wrt an input measure eg Gasussian
- Close link to Gaussian quadrature, in which case the gerid locations may b e zeros of orthogonal polynomials
- Nesting problem: zero of orthogonal polynomials interlace but do not nest, whereas nesting is useful for augmentation/sequential

The α-notation

A basic grid: Experimental design meets models.



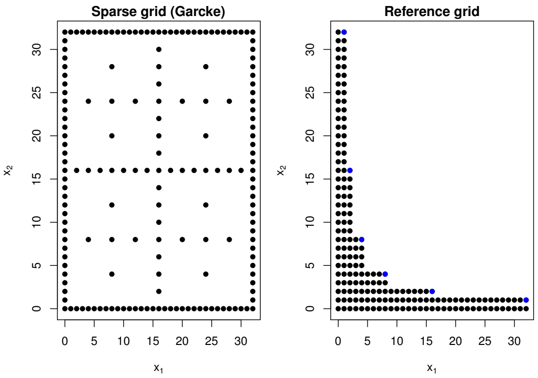
$\alpha = (\alpha_1, \alpha_2)$ is both a design point and monomilal in a polynomial model
Design: $D = \{(0,0),(1,0),(0,1),(2,0),(1,1),(0,2),(3,0)\}$
Model: $L = \{1, x_1, x_2, x_1^2, x_1x_2, x_2^2, x_1^3\}$

Inclusion-exclusion

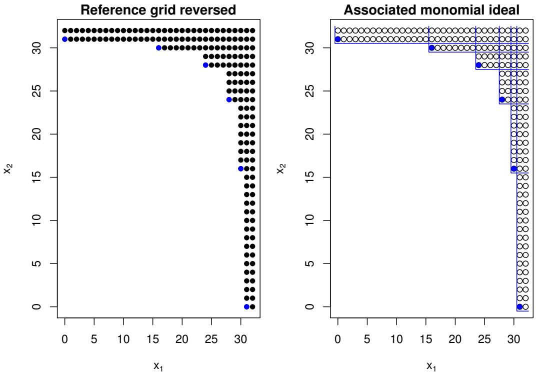
The grid and the model can be be written as a union of "tensor" grids, which in DOE we call "full factorial" design. Schematically we can say

$$D = \begin{bmatrix} * & * & * & * \end{bmatrix} + \begin{bmatrix} * & * \\ * & * \end{bmatrix} + \begin{bmatrix} * \\ * \\ * \end{bmatrix} - \begin{bmatrix} * & * \end{bmatrix} - \begin{bmatrix} * \\ * \end{bmatrix}$$

Sparse grid



Flipping the grid



Monomial ideals and Hilbert function

Suppose we have a set of integer vectors $\{\alpha, \alpha \in S\}$, then we can define the generating function of S as

$$G_S(x) = \sum_{\alpha \in S} x^{\alpha}.$$

1. $S = \{0, 1, 2, \ldots\}$

$$G_S(x) = \frac{1}{1-x}$$

2. $S = \{0, 1, 2, \ldots n-1\}$

$$G_S(x) = \frac{1-x^n}{1-x}$$

3. $S = \{x^n, x^{n+1}, \ldots\}$

$$G_S(x) = \frac{x^n}{1-x}$$

Monomials:

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d}$$

Monomial ideals:

$$I = < x^{\beta^{(1)}}, \ldots, x^{\beta^{(m)}} >$$

The upper orthants gives all the monomials in the corresponding monomial ideal $< x^\beta >$.
The union of upper gives all monomials in I. For two orthants we have

$$I = < x^\alpha, x^\beta >$$

and the generating function is

$$G_{Q(\alpha)\cup Q(\beta)} = G_{Q(\alpha)} + G_{Q(\beta)} - G_{Q(\alpha\wedge\beta)},$$

where

$$\alpha \wedge \beta = (\max(\alpha_1, \beta_1), \ldots, \max(\alpha_d, \beta_d)),$$

which corresponds to $\mathrm{LCM}(x^\alpha, x^\beta)$.

Monomial ideals

- An ideal: $\langle gl(x), \ldots, g_m(x) \rangle$ is the set of all polynomials:

$$s_1(x)g(x) + \cdots s_m(x)g_m(x)$$

- A monomial ideal: all the $g_j(x)$ are monomials.

$$\begin{array}{ccccccc}
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
x_2^5 & x_1x_2^5 & x_1^2x_2^5 & x_1^3x_2^5 & x_1^4x_2^5 & x_1^5x_2^5 & \ldots \\
x_2^4 & x_1x_2^4 & x_1^2x_2^4 & x_1^3x_2^4 & x_1^4x_2^4 & x_1^5x_2^4 & \ldots \\
x_2^3 & x_1x_2^3 & x_1^2x_2^3 & x_1^3x_2^3 & x_1^4x_2^3 & x_1^5x_2^3 & \ldots \\
x_2^2 & x_1x_2^2 & x_1^2x_2^2 & x_1^3x_2^2 & x_1^4x_2^2 & x_1^5x_2^2 & \ldots \\
x_2 & x_1x_2 & x_1^2x_2 & x_1^3x_2 & x_1^4x_2 & x_1^5x_2 & \ldots \\
1 & x_1 & x_1^2 & x_1^3 & x_1^4 & x_1^5 & \ldots
\end{array}$$

- : $< x_1x_2^4, x_1^3x_2^2, x_1^5 >$

- The multigraded Hilbert series of I in terms of some resolution of I.

$$H = \frac{\sum(-1)^i \gamma_{i,\alpha} x^\alpha}{\prod_i (1 - x_i)}$$

- This method generalizes the classical inclusion-exclusion approach (corresponding to Taylor resolution)
- The minimal free resolution uses the multi-graded Betti numbers $\beta_{i,\mu}$ and gives *tighter* bounds than for *any* resolution:

$$H = \frac{\sum(-1)^i \beta_{i,\alpha} x^\alpha}{\prod_i (1 - x_i)}$$

Example contd.

j	0	1
31	2	-
46	2	2
52	2	-
53	-	2
55	-	1
tot	6	5

$$t[1]^{31} + t[2]^{31} + t[1]^{30}t[2]^{16} + t[1]^{16}t[2]^{30} + t[1]^{28}t[2]^{24} + t[1]^{24}t[2]^{28} - t[1]^{28}t[2]^{28} - t[1]^{30}t[2]^{24} - t[1]^{24}t[2]^{30} - t[1]^{31}t[2]^{16} - t[1]^{16}t[2]^{31}$$

Notation

Notation:

- D: basic nested design (defined later)
- $G^{(D)}$: reference grid
- $M^{(D)}$: reference model
- $P_M^{(D)}(y, x)$ polynomial interpolator of y on D using $M^{(D)}$
- G_α: tensor grid with "corner" α: $\{\beta : 0 \leq \beta \leq \alpha\}$
- M_α: tensor model with "corner" α: $\{x^\beta : 0 \leq \beta \leq \alpha\}$
- $P_\alpha(y_\alpha, x)$: polynomial interpolator of data y_α on G_α using model M_α

Inclusion-exclusion: G

$\tilde{G}$ is the indicator function for G.

$$\tilde{G} = \tilde{\bigcup_{\alpha \in A}} G_\alpha = \sum_j (-1)^{j-1} \sum_{\alpha \in A_j} \beta_{i,\alpha} \tilde{G}_{j,\alpha}$$

The decomposition of the polynomial interpolator on G is exactly the same as for the design:

$$P_G(y,x) = \sum_j (-1)^{j-1} \sum_{\alpha \in A_j} \beta_{i,\alpha} P_{G_\alpha, j, \alpha}(y_{j,\alpha}, x)$$

Inclusion-exclusion: D

$\tilde{D}$ is the indicator function for D.

$$\tilde{D} = \tilde{\bigcup_{\alpha \in A}} D_\alpha = \sum_j (-1)^{j-1} \sum_{\alpha \in A_j} \beta_{i,\alpha} \tilde{D}_{j,\alpha}$$

On D

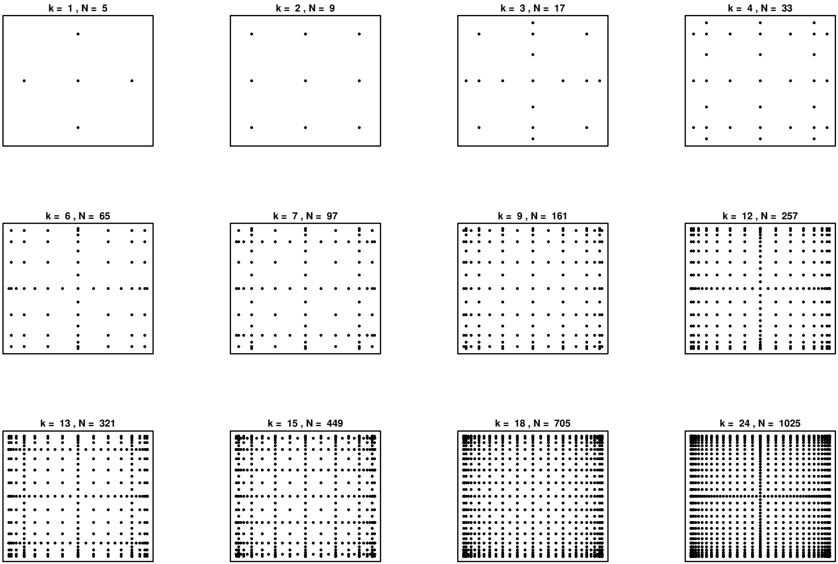
$$P_\alpha(y, x) = \sum_j (-1)^{j-1} \sum_{\alpha \in A_j} \beta_{i,\alpha} P_{D_\alpha, j, \alpha}(y_{j,\alpha}, x),$$

where $D_{j,\alpha}$ is the inverse image of $G_{j,\alpha}$ under the mapping $D \to G$

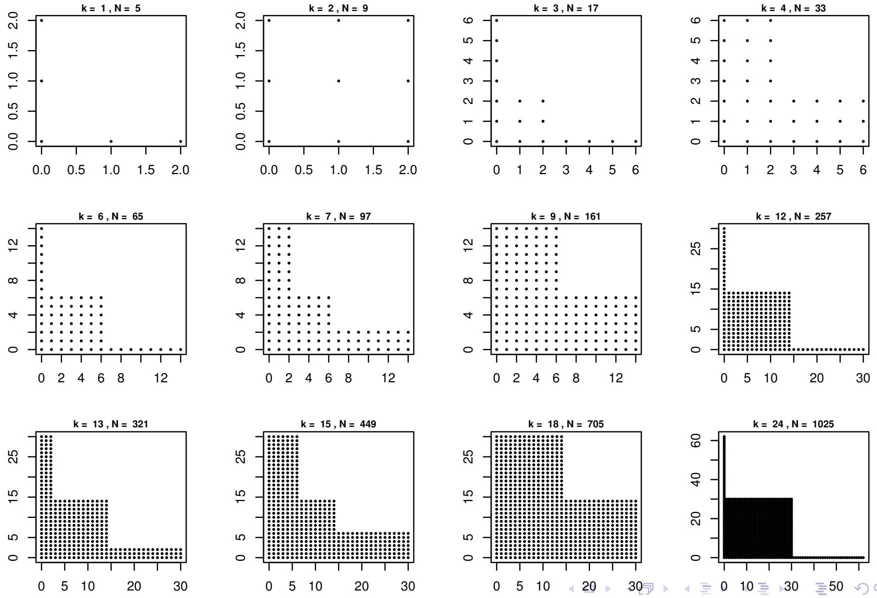
Summary

$$
\begin{tikzcd}
D \arrow[r] \arrow[d] & G^{(D)} \arrow[r] \arrow[d] & I^{(G)} \arrow[d] \\
\{D_{j,\alpha}\} \arrow[d] & G^{(D)}_{j,\alpha} \arrow[l] & \{I_{j,\alpha}\} \arrow[l] \\
P_{D_\alpha,j,\alpha}(y_{j,\alpha}, x) & &
\end{tikzcd}
$$

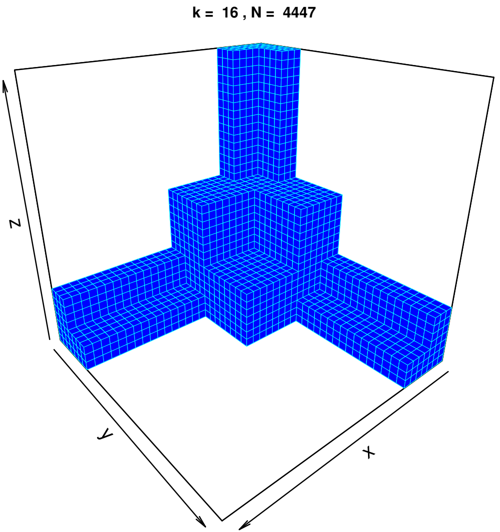
More examples



Reference grids



3-d example



Betti for a 3-d example

j	0	1	2
52	6	-	-
55	-	3	-
56	3	-	-
63	-	3	-
67	-	6	-
70	-	-	4
tot	9	12	4

Hilbert Series for 3-d example

$$
\begin{gathered}
+t[1]^2 8 \ast t[2]^2 4 + t[1]^2 4 \ast t[2]^2 8 + t[1]^2 8 \ast t[3]^2 4 + t[2]^2 8 \ast t[3]^2 4 + t[1]^2 4 \ast \\
t[3]^2 8 + t[2]^2 4 \ast t[3]^2 8 + t[1]^2 4 \ast t[2]^1 6 \ast t[3]^1 6 + t[1]^1 6 \ast t[2]^2 4 \ast t[3]^1 6 + \\
t[1]^1 6 \ast t[2]^1 6 \ast t[3]^2 4 - t[1]^2 8 \ast t[3]^2 8 - t[2]^2 8 \ast t[3]^2 8 - t[1]^2 8 \ast t[2]^2 8 - \\
t[1]^2 4 \ast t[2]^2 4 \ast t[3]^1 6 - t[1]^2 4 \ast t[2]^1 6 \ast t[3]^2 4 - t[1]^1 6 \ast t[2]^2 4 \ast t[3]^2 4 - \\
t[1]^2 8 \ast t[2]^2 4 \ast t[3]^1 6 - t[1]^2 4 \ast t[2]^2 8 \ast t[3]^1 6 - t[1]^2 8 \ast t[2]^1 6 \ast t[3]^2 4 - t[1]^1 6 \ast \\
t[2]^2 8 \ast t[3]^2 4 - t[1]^2 4 \ast t[2]^1 6 \ast t[3]^2 8 - t[1]^1 6 \ast t[2]^2 4 \ast t[3]^2 8 + t[1]^2 8 \ast t[2]^2 8 \ast \\
t[3]^1 6 + t[1]^2 4 \ast t[2]^2 4 \ast t[3]^2 4 + t[1]^2 8 \ast t[2]^1 6 \ast t[3]^2 8 + t[1]^1 6 \ast t[2]^2 8 \ast t[3]^2 8
\end{gathered}
$$

Nested designs

We need to define the class of "nested" designs under which the above results hold.

- Define a grid G in the usual way, in which variable x_i has levels $1, 2, \ldots, n_i$
- Define new levels for the design D_0 in which variabel x_i has levels

$$z_{i,1} \leq \cdots \leq z_{i,n_i}$$

  Apart from the spacing this still has ths same "structure" as G
- Let be a permutaion of $\{1, \ldots, n_i\}$, $i = 1 \ldots n$ and write $\pi = (\pi_1, \ldots, \pi_n)$
- From D_0 construct D_π which permutes the levels, for each factor so that all design point with levels $z_{i,j}$ for factor i are assigned levels $z_{i,z_{i,\pi_i(j)}}$ for $j = 1, \ldots, n_i$, $i = 1, \ldots, n$.

Alexander duality

In [1] we define an operation which allows us to construct new designs D which are nested and for which we know easilly what the model is.
Let $\mathcal{N}(k_i) = \{1 \ldots, k_i\}$ and define a tensor grid

$$G(k_1, \ldots, k_n) = \times_{i=1}^{n} \mathcal{N}(n_i),$$

For a grid G with assciated model M with maximum levels $(n_1, \ldots, n_n)$ interger $k_i \geq n_i, \; i = 1, \ldots, n$, define the complementary design.

$$D' = G(k_1, \ldots, k_n) \setminus G.$$

D' is a nested dersign which has a reference grid G' and associated model M' which is the Alexander dual of M.

Example

$$
\begin{array}{cccc}
* & * & * & * \\
* & 0 & 0 & * \\
* & 0 & 0 & * \\
* & * & * & *
\end{array}
\longrightarrow
\begin{array}{cccc}
* & * & * & * \\
* & & & * \\
* & & & * \\
* & * & * & *
\end{array}
\longrightarrow
\begin{array}{cccc}
* & * & & \\
* & * & & \\
* & * & * & * \\
* & * & * & *
\end{array}
$$

$$
\begin{array}{cccccc}
* & * & * & * & * & * \\
* & 0 & 0 & 0 & 0 & * \\
* & 0 & * & * & 0 & * \\
* & 0 & * & * & 0 & * \\
* & 0 & 0 & 0 & 0 & * \\
* & * & * & * & * & *
\end{array}
\longrightarrow
\begin{array}{cccccc}
* & * & * & * & * & * \\
* & & & & & * \\
* & & * & * & & * \\
* & & * & * & & * \\
* & & & & & * \\
* & * & * & * & * & *
\end{array}
\longrightarrow
\begin{array}{cccccc}
* & * & & & & \\
* & * & & & & \\
* & * & * & * & & \\
* & * & * & * & & \\
* & * & * & * & * & * \\
* & * & * & * & * & *
\end{array}
$$

New types of design

$$
\begin{array}{cccccccccc}
\ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast \\
\ast & & & \ast & & & \ast & & & \ast \\
\ast & & & \ast & & & \ast & & & \ast \\
\ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast \\
\ast & & & \ast & & & \ast & & & \ast \\
\ast & & & \ast & & & \ast & & & \ast \\
\ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast \\
\ast & & & \ast & & & \ast & & & \ast \\
\ast & & & \ast & & & \ast & & & \ast \\
\ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast
\end{array}
\longrightarrow
\begin{array}{cccccccccc}
\ast & \ast & \ast & \ast & & & & & & \\
\ast & \ast & \ast & \ast & & & & & & \\
\ast & \ast & \ast & \ast & & & & & & \\
\ast & \ast & \ast & \ast & & & & & & \\
\ast & \ast & \ast & \ast & & & & & & \\
\ast & \ast & \ast & \ast & & & & & & \\
\ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast \\
\ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast \\
\ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast \\
\ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast
\end{array}
$$

Interpolation and quadrature error

In the approximation theory we assume we observe a function f on the design D. If D is in our class then inspection of the reference grid $G^{(D)}$ gives us our polynomial basis M and the interpolator

$$P_D(y, x)$$

- Quadrature
$$\left| \int (f) d\mu - \int (P_D(y, x)) \right| \leq B_1$$
- Interpolation
$$|| f - P_D(y, x)) || \leq B_2$$

Sketch of methods

- Lebesgue constants

$$\left\| f - \int P_D(y,x)) \right\| \leq \left\| \ f - P^* \right\| \times \Delta$$

  where P^* is the best polynomial approximation to f and Δ is a "Lebesque" constant which depends only on on the "structure" of D and the spacing.

- The inclusion-exclusion identity for the main tensor components is replaced by special disjoint components corresponding to the elementary cells, which are also tensors into which the reference grid and interpolator is divided. They take the form

$$\Delta_{1,i_1} \cdots \Delta_{n_{i_n}}.$$

Optimal spacing can decrease the errors in interpolation and quadrature.

References

1. Maruri-Aguilar, H., Senz-de-Cabezn, E. and Wynn, H.P., 2013. Alexander duality in experimental designs. Annals of the Institute of Statistical Mathematics, 65(4), pp.667-686.
2. Barthelmann, V., Novak, E. and Ritter, K., 2000. High dimensional polynomial interpolation on sparse grids. Advances in Computational Mathematics, 12(4), pp.273-288.
3. Ko, J. and Wynn, H.P., 2016. The algebraic method in quadrature for uncertainty quantification. SIAM/ASA Journal on Uncertainty Quantification, 4(1), pp.331-357.
4. Nobile, F., Tempone, R. and Webster, C.G., 2008. A sparse grid stochastic collocation method for partial differential equations with random input data. SIAM Journal on Numerical Analysis, 46(5), pp.2309-2345.
5. Plumlee, M., 2014. Fast prediction of deterministic functions using sparse grid experimental designs. Journal of the American Statistical Association, 109(508), pp.1581-1591.