Sampling to understand high dimensional functions

Part II: new material

Art B. Owen

Stanford University

This course

This is a two part course.

I) **Review** of material on:

a) Functional ANOVA

b) quasi-regression

c) quasi-Monte Carlo sampling

II) **New** material on:

a) applications to missing heritability problems

b) effective dimension of some Sobolev spaces

c) new Sobol' index quantities

Part II(a): missing heritability

Based on work in progress with Or Zuk of the Broad Institute

.

Real world data $\cdots$ that resembles a computer experiment.

Background on SNPs

SNPs are **S**ingle **N**ucleotide **P**olymorphisms

Base pairs are A, C, G or T

At most loci, everybody gets the same

At some loci, there is a minor allele, e.g. most are A but 10% are T

SNP chips, measure $100{,}000$s to $1{,}000{,}000$s of loci

Each person has 0, 1, or 2 copies of the minor allele

Typical sample sizes

1000s of people and $100{,}000$s of SNPs

Prototypical data

Subject	SNP_1	SNP_2	$\cdots$	$\text{SNP}_{100,000}$	Height	Diabetes
1	0	0	$\cdots$	1	1.7m	0
2	2	0	$\cdots$	0	2.1m	1
3	1	0	$\cdots$	1	1.5m	0
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
1000	0	1	$\cdots$	0	1.9m	1

We would like to predict based on SNPs.

Typically use p-value 10^{-8} to avoid false positives.

The few discoveries explain only a little of phenotype (eg height).

Eg: height may be ≈ 80% heritable but discovered genes explain only 5–10%.

Where is the missing heritability?

Quasi-regression for heritability

Genotype is $\boldsymbol{x} = (x_1, \dots, x_d)$. Phenotype is y.

Assume $\|\boldsymbol{x}\|$ and $|y|$ bounded.

Both normalized: $\mathbb{E}(y) = \mathbb{E}(x_j) = 0$ and $\mathbb{E}(y^2) = \mathbb{E}(x_j^2) = 1$.

Model

$$y_i = \sum_{j=1}^{d} \beta_j x_{ij} + \varepsilon_i$$

$$\mathbb{E}(x_j \varepsilon) = 0$$

ε includes environment, environment$\times$genes, genes$\times$genes

Linear heritability is

$$\sigma_{\mathrm{L}}^2 = \sum_{j=1}^{d} \beta_j^2$$

We assume (unrealistically) that x_j is independent of x_k for $j \neq k$.

Holds for most pairs j, k Will require later adjustments

Quasi-regression

$$\widetilde{\beta}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}y_i \qquad \hat{\sigma}^2_{\mathrm{L}} = \sum_{j=1}^{d}\widetilde{\beta}_j^2$$

Recall

$$\hat{\sigma}^2_{\mathrm{L,BC}} = \frac{n}{n-1}\bigl(\hat{\sigma}^2_{\mathrm{L}} - \widehat{B}\bigr), \quad \text{where}$$

$$\widehat{B} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{d} x_{ij}^2 y_i^2 = \frac{1}{n^2}\sum_{i=1}^{n}\|\boldsymbol{x}_i\|^2 y_i^2 = O_p\Bigl(\frac{d}{n}\Bigr)$$

Then

$$\mathbb{E}\Bigl(\bigl(\hat{\sigma}^2_{\mathrm{L,BC}} - \sigma^2_{\mathrm{L}}\bigr)^2\Bigr) = O\Bigl(\frac{1}{n} + \frac{d^2}{n^3}\Bigr) \qquad n \geqslant 2, \quad d \geqslant 1 \tag{1}$$

We need $n \gg d^{2/3}$ i.e. $d \gg n$ ok even if β is not sparse

$\exists$ lower bounds for estimation of non-sparse β

Candès & Davenport (2011), Raskutti, Wainwright & Yu (2009)

Concentration of measure

$$\|\boldsymbol{x}_i\|^2 = \sum_{j=1}^{d} x_{ij}^2 \approx d$$

$$\widehat{B} = \frac{1}{n^2}\sum_{i=1}^{n} \|\boldsymbol{x}_i\|^2 y_i^2$$

$$\widetilde{B} = \frac{d}{n^2}\sum_{i=1}^{n} y_i^2$$

We can show

$$\mathbb{E}\bigl((\widehat{B}-\widetilde{B})^2\bigr) = O\Bigl(\frac{d}{n^2}\Bigr)$$

So replacing $\|\boldsymbol{x}_i\|^2$ by d makes a negligible change

Somewhat faster

O (2000) has an example with $d = 1{,}000{,}000$ and $n = 100{,}000$

Additive model

$$z_j = \frac{x_j^2 - \mathbb{E}(x_j^3)x_j - 1}{\sqrt{\mathbb{E}(x_j^4) - \mathbb{E}(x_j^3)^2 - 1}}$$

$$y_i = \sum_{j=1}^{d} \beta_j x_{ij} + \sum_{j=1}^{d} \gamma_j z_{ij} + \varepsilon_i$$

Captures effects of dominant vs. recessive genes

$x_1, \cdots, x_d,\ z_1, \cdots, z_d$ are **uncorrelated**

Additive heritability

$$\sigma_{\mathrm{A}}^2 = \sum_{j=1}^{d} \beta_j^2 + \sum_{j=1}^{d} \gamma_j^2$$

$$\widetilde{\gamma}_j = \frac{1}{n}\sum_{i=1}^{n} z_{ij} y_i$$

$$\hat{\sigma}_{\mathrm{A}}^2 = \sum_{j} (\widetilde{\beta}_j^2 + \widetilde{\gamma}_j^2)$$

$$\hat{\sigma}_{\mathrm{A,BC}}^2 = \frac{n}{n-1}\bigl(\hat{\sigma}_{\mathrm{A}}^2 - \widehat{B}_{\mathrm{A}}\bigr), \quad \text{where}$$

$$\widehat{B}_{\mathrm{A}} = \frac{1}{n^2}\sum_{i=1}^{n} (\|\boldsymbol{x}_i\|^2 + \|\boldsymbol{z}_i\|^2) y_i^2 \approx \frac{2d}{n^2}\sum_{i=1}^{n} y_i^2 .$$

Again

$$\mathbb{E}\Bigl(\bigl(\hat{\sigma}_{\mathrm{A,BC}}^2 - \sigma_{\mathrm{A}}^2\bigr)^2\Bigr) = O\Bigl(\frac{1}{n} + \frac{d^2}{n^3}\Bigr)$$

Quadratic model

$$y_i = \sum_{j=1}^{d} \beta_j x_{ij} + \sum_{j=1}^{d} \gamma_j z_{ij} + \sum_{j<k} \beta_{jk} x_{ij} x_{ik} \varepsilon_i$$

The model allows interactions.

There are $2d + d(d-1)/2$ parameters.

$$\sigma_{\mathrm{Q}}^2 = \sum_j \beta_j^2 + \sum_j \gamma_j^2 + \sum_{j<k} \beta_{jk}^2$$

For this model

Squared error $O\Big(\frac{1}{n} + \frac{d^4}{n^3}\Big)$ we need n of order at least $d^{4/3}$ (ouch)

It no longer works to use $\|\boldsymbol{x}_i\|^2 \approx d$

Could be an issue for bias adjusted quasi-regression estimation of Sobol' indices

Part II(b):

Effective dimension of some Sobolev spaces

Weighted spaces are used in QMC, starting with

Hickernell (1996) and Sloan & Wozniakowski (1998)

They downweight high order interactions and high index variables.

1) x_j ordinarily more important than x_{j+1}

2) $f_u(\boldsymbol{x})$ ordinarily more important than $f_v(\boldsymbol{x})$ for $|v| > |u|$

QMC can integrate functions in some such spaces without any curse of dimensionality

Those spaces are dominated by low order components (as follows)

Effective dimension

A function has effective dimension s in the **truncation** sense if

$$\sum_{u\subseteq\{1,2,\dots,s\}} \sigma_u^2 \geqslant (1-\varepsilon)\sigma^2$$

A function has effective dimension s in the **superposition** sense if

$$\sum_{|u|\leqslant s} \sigma_u^2 \geqslant (1-\varepsilon)\sigma^2$$

Commonly $\varepsilon = 0.01$

QMC often succeeds on functions of low effective dimension in the superposition sense

Caflisch, Morokoff & O (1997)

Weighted spaces

Pick weights $1 \geqslant \gamma_1 > \gamma_2 > \cdots > \gamma_d > 0$. Let $\gamma_u = \prod_{j \in u} \gamma_j$.

Inner product and norm

$$\langle f, g \rangle_{W_{d,\gamma}} = \sum_{u \subseteq 1:d} \frac{1}{\gamma_u} \int \frac{\partial^{|u|} f(\boldsymbol{x})}{\partial \boldsymbol{x}_u} \frac{\partial^{|u|} \bar{g}(\boldsymbol{x})}{\partial \boldsymbol{x}_u} \, \mathrm{d}\boldsymbol{x}$$

$$\|f\|^2_{W_{d,\gamma}} = \langle f, f \rangle_{W_{d,\gamma}}.$$

Function class $=$ ball

$$f \in \mathcal{B}(d, \gamma, \rho) \equiv \{ f \mid \|f\|_{W_{d,\gamma}} \leqslant \rho \}$$

How it works

Large $|u| \implies$ small $\gamma_u \implies$ large penalty factor $1/\gamma_u$

$\implies$ only small $\dfrac{\partial^{|u|} f}{\partial \boldsymbol{x}_u}$ are allowed

So $\boldsymbol{x}_u$ not important.

Tractability over $\mathcal{B}(d, \gamma, \rho)$

Quadrature is **tractable** if the worst case error is $O(n^{-a} d^{b})$

Quadrature is **strongly tractable** if the worst case error is $O(n^{-a})$, uniformly in d

(see monographs Novak & Woźniakowski (2008,2010))

$$\sum_{j=1}^{\infty} \gamma_j < \infty \implies \text{strong tractability, with } a = 1/2$$

$$\sum_{j=1}^{\infty} \gamma_j^{1/2} < \infty \implies \text{strong tractability, with } a = 1 - \epsilon$$

Rapid weight decay $\implies$ no dimension effect

What sort of functions are in $\mathcal{B}(d, \gamma, \rho)$?

Strong properties $\cdots$ often come from strong assumptions

Tractability and effective dimension

When quadrature is strongly tractable,

then the functions in $\mathcal{B}(d,\gamma,\rho)$ do not have large σ_u^2 for large $|u|$.

To describe this

Find ρ^*, the smallest ρ that allows $f \in \mathcal{B}(d,\gamma,\rho)$ with $\mathrm{Var}(f) = 1$.

Then find the largest σ_u^2 for $f \in \mathcal{B}(d,\gamma,\rho^*)$ and $|u| \geqslant s$.

If this $\sigma_u^2 \leqslant 0.01$ then $\mathcal{B}(d,\gamma,\rho)$ has effective dimension at most s

An inequality

Let f have continuous derivative $f' \in L^2[0,1]$ with $\int_0^1 f(x)\,\mathrm{d}x = 0$. Then

$$\sigma^2 = \int_0^1 f(x)^2\,\mathrm{d}x \leqslant \frac{1}{\pi^2}\int_0^1 |f'(x)|^2\,\mathrm{d}x.$$

Sobol' (1963) appeals to calculus of variations.

Fourier methods generalize nicely to higher dimensional analogues.

We can bound σ_u^2 by a multiple of

$$\int \Bigl(\frac{\partial^{|u|} f(\boldsymbol{x})}{\partial \boldsymbol{x}_u}\Bigr)^2\,\mathrm{d}\boldsymbol{x}.$$

There is no reverse inequality, bounding $\int |f'|^2$ by σ^2.

Related

Lamboni, Iooss, Popelin & Gamboa (2012)

bound Sobol' indices by integrated derivatives, for Boltzmann measures

Fourier representation

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \lambda_{\boldsymbol{k}} e^{2\pi i \boldsymbol{k}^\mathsf{T}\boldsymbol{x}}$$

$$\lambda_{\boldsymbol{k}} = \int_{[0,1]^d} f(\boldsymbol{x}) e^{-2\pi i \boldsymbol{k}^\mathsf{T}\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x}$$

$$\mathrm{Var}(f) = \sum_{\boldsymbol{k}\neq\boldsymbol{0}} |\lambda_{\boldsymbol{k}}|^2$$

After some algebra

$$\frac{\partial^{|u|} f(\boldsymbol{x})}{\partial \boldsymbol{x}_u} = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \lambda_{\boldsymbol{k}} \Big((2\pi i)^{|u|} \prod_{j\in u} k_j \Big) e^{2\pi i \boldsymbol{k}^\mathsf{T}\boldsymbol{x}}, \quad \text{and}$$

$$\|f\|^2_{W_{d,\gamma}} = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} |\lambda_{\boldsymbol{k}}|^2 \prod_{j=1}^{d} \left(1 + \frac{4\pi^2 k_j^2}{\gamma_j}\right)$$

Finding the dimension

maximize $\quad \mathrm{Var}(f) = \sum_{\boldsymbol{k}\neq\boldsymbol{0}} |\lambda_{\boldsymbol{k}}|^2$

subject to $\quad \|f\|^2 = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} |\lambda_{\boldsymbol{k}}|^2 \prod_{j=1}^{d}\Bigl(1+\frac{4\pi^2k_j^2}{\gamma_j}\Bigr) \leqslant \rho^2$

Maximum happens when $\lambda_{\boldsymbol{k}} = 0$ except for $\boldsymbol{k} = (\pm 1, 0, 0, \dots, 0)$.

We get $\|f\|^2 = \mathrm{Var}(f)(1+4\pi^2/\gamma_1),$ so $\rho^{*,2} = 1 + 4\pi^2/\gamma_1$.

$$\sigma_u^2 = \sum_{\boldsymbol{k}_u\in(\mathbb{Z}-\{0\})^{|u|}} |\lambda_{\boldsymbol{k}_u:\boldsymbol{0}_{-u}}|^2$$

Similarly $\quad \max_{f\in\mathcal{B}(d,\gamma,\rho^*)} \sigma_u^2 = \dfrac{1+4\pi^2/\gamma_1}{\prod_{j\in u}(1+4\pi^2/\gamma_j)}$

If $|u|$ is large and γ_j decay rapidly $\cdots$ then σ_u^2 must be small

Effective dimensions

For $\gamma_j = j^{-1.01}$

Just barely strongly tractable

Truncation dimension is 97

Superposition dimension is 2

For $\gamma_j = j^{-2}$

Almost $O(n^{-1+\epsilon})$ uniformly in d

Truncation dimension is 10

Superposition dimension is 1

These use $\varepsilon = 0.01$

At $\varepsilon = 0.0001$ superposition dimensions rise to 3 and 2

Part II(c): new Sobol' indices

Recall

$$\underline{\tau}_u^2 = \sum_{v \subseteq u} \sigma_v^2$$

$$\overline{\tau}_u^2 = \sum_{v \cap u \neq \varnothing} \sigma_v^2$$

$$\underline{\tau}_u^2 = \iint f(\boldsymbol{x})\bigl(f(\boldsymbol{x}_u{:}\boldsymbol{z}_{-u}) - f(\boldsymbol{z})\bigr)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z}$$

What other quantities can we get?

In principal we can get any σ_u^2 from

$$\sigma_u^2 = \sum_{v \subseteq u} (-1)^{|u-v|} \underline{\tau}_v^2$$

we prefer

1) unbiased estimates
2) low variance
3) averages of squared differences to differences of squares
4) to avoid $O(2^d)$ integrals, or even $O(d^2)$ integrals

The set of possibilities

$$\sum_{u\subseteq\mathcal{D}}\sum_{v\subseteq\mathcal{D}}\lambda_{u,v}\iint f(\boldsymbol{x}_u{:}\boldsymbol{z}_{-u})f(\boldsymbol{x}_v{:}\boldsymbol{z}_{-v})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z}$$

where

$$\boldsymbol{y}=\boldsymbol{x}_u{:}\boldsymbol{z}_{-u}\implies y_j=\begin{cases}x_j, & j\in u\\ z_j, & j\notin u\end{cases}$$

and

$$\lambda_{u,v}\in\mathbb{R}$$

A 2^{2d} dimensional space of Sobol' quantities (overparameterized)

We only want $\sum_u \delta_u\sigma_u^2$ for $\delta\in\mathbb{R}^{2^d}$

Special subsets

1) Squares

$$\int\!\!\int \Bigl(\sum_u \lambda_u f(\boldsymbol{x}_u{:}\boldsymbol{z}_{-u})\Bigr)^2 \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z}$$

Nonnegative

2) Bilinear

$$\int\!\!\int \Bigl(\sum_u \lambda_u f(\boldsymbol{x}_u{:}\boldsymbol{z}_{-u})\Bigr)\Bigl(\sum_u \gamma_u f(\boldsymbol{x}_u{:}\boldsymbol{z}_{-u})\Bigr) \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z}$$

Fast

3) All for one

$$\int\!\!\int \Bigl(\sum_u \lambda_u f(\boldsymbol{x}_u{:}\boldsymbol{z}_{-u})\Bigr) f(\boldsymbol{z}) \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z}$$

Simple

4) Contrast

$$\sum_u \sum_v \lambda_{u,v} = 0$$

Free of $f_{\varnothing}^2$

The basis

$$\iint f(\boldsymbol{x}_u\!:\!\boldsymbol{z}_{-u})f(\boldsymbol{x}_v\!:\!\boldsymbol{z}_{-v})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z} = f_\varnothing^2 + \underline{\tau}^2_{\mathrm{NXOR}(u,v)}$$

Not exclusive or

$j \in u$	$j \in v$	$\mathrm{XOR}(u,v)$	$\mathrm{NXOR}(u,v)$
0	0	0	1
0	1	1	0
1	0	1	0
1	1	0	1

$j \in \mathrm{NXOR}(u,v) \iff j \in u \cap v$ or $j \in u^c \cap v^c$

Proof

$$
\begin{aligned}
&\iint f(\boldsymbol{x}_u{:}\boldsymbol{z}_{-u})f(\boldsymbol{x}_v{:}\boldsymbol{z}_{-v})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z}\\
&=\sum_{w\subseteq\mathcal{D}}\sum_{w'\subseteq\mathcal{D}}\iint f_w(\boldsymbol{x}_u{:}\boldsymbol{z}_{-u})f_{w'}(\boldsymbol{x}_v{:}\boldsymbol{z}_{-v})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z}\\
&=\sum_{w\subseteq\mathcal{D}}\iint f_w(\boldsymbol{x}_u{:}\boldsymbol{z}_{-u})f_w(\boldsymbol{x}_v{:}\boldsymbol{z}_{-v})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z}\\
&=\sum_{w}1_{w\subseteq(u\cap v)\cup(u^c\cap v^c)}\int f_w(\boldsymbol{x})^2\,\mathrm{d}\boldsymbol{x}\\
&=\sum_{w\subseteq\mathrm{NXOR}(u,v)}\int f_w(\boldsymbol{x})^2\,\mathrm{d}\boldsymbol{x}\\
&=f_\varnothing^2+\underline{\tau}^2_{\mathrm{NXOR}(u,v)}.
\end{aligned}
$$

All for one

$$\sum_v \lambda_v \iint f(\boldsymbol{x}) f(\boldsymbol{x}_v{:}\boldsymbol{z}_{-v})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z}$$
$$= \sum_v \lambda_v \Bigl(f_\varnothing^2 + \underline{\tau}^2_{\mathrm{NXOR}(\mathcal{D},v)} \Bigr)$$
$$= f_\varnothing^2 \sum_v \lambda_v + \sum_v \lambda_v \underline{\tau}_v^2$$

Combinations with $\sum_{v\subseteq\mathcal{D}} \lambda_v = 0$ are free of $f_\varnothing^2$

For $\underline{\tau}_u^2$ take

$$\lambda_v = \begin{cases} (-1)^{|u-v|} & u \subseteq v \\ 0 & \text{else.} \end{cases}$$

(Lots of alternation)

Mean squares

If $\sum_v \lambda_v = 0$ then the following is a nonnegative estimate of a linear combination of σ_u^2

$$\iint \Bigl(\sum_v \lambda_v f(\boldsymbol{x}_v{:}\boldsymbol{z}_{-v})\Bigr)^2 \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z}$$

$$= \sum_v \sum_w \lambda_v \lambda_w \iint f(\boldsymbol{x}_v{:}\boldsymbol{z}_{-v}) f(\boldsymbol{x}_w{:}\boldsymbol{z}_{-w}) \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z}$$

$$= \sum_v \sum_w \lambda_v \lambda_w \underline{\tau}^2_{\mathrm{NXOR}(w,v)}$$

Combinations with $\sum_{v\subseteq\mathcal{D}} \lambda_v = 0$ are free of $f_\varnothing^2$

Q: What can we get as a square?

A: $\overline{\tau}_u^2$ and Υ_u^2 (but what else?)

Cannot get $\underline{\tau}_u^2$ as a square

$$\iint \Bigl(\lambda_0 f(\boldsymbol{z}) + \lambda_1 f(\boldsymbol{x}_u{:}\boldsymbol{z}_{-u}) + \lambda_d f(\boldsymbol{x})\Bigr)^2 \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z}$$

with $\lambda_0 = A,\quad \lambda_d = B,\quad \lambda_1 = -A-B,$ WLOG $f_\varnothing = 0$

$$\begin{array}{cccc} \mathrm{XOR} & \varnothing & u & \mathcal{D} \\ \varnothing & \varnothing & u & \mathcal{D} \\ u & u & \varnothing & -u \\ \mathcal{D} & \mathcal{D} & -u & \varnothing \end{array} \quad \to \quad \begin{array}{cccc} \mathrm{NXOR} & \varnothing & u & \mathcal{D} \\ \varnothing & \mathcal{D} & -u & \varnothing \\ u & -u & \mathcal{D} & u \\ \mathcal{D} & \varnothing & u & \mathcal{D} \end{array}$$

$$\to \quad \begin{array}{cccc} \underline{\tau}^2_{\mathrm{NXOR}} & \varnothing & u & \mathcal{D} \\ \varnothing & \sigma^2 & \underline{\tau}^2_{-u} & 0 \\ u & \underline{\tau}^2_{-u} & \sigma^2 & \underline{\tau}^2_u \\ \mathcal{D} & 0 & \underline{\tau}^2_u & \sigma^2 \end{array} \quad \to \quad \begin{array}{c} \sigma^2(A^2 + B^2 + (A+B)^2) \\ -\underline{\tau}^2_u B(A+B) - \underline{\tau}^2_{-u} A(A+B) \end{array}$$

$\underline{\tau}_u^2$ as a square

We get

$$\iint \Bigl(Af(\boldsymbol{z}) - (A+B)f(\boldsymbol{x}_u{:}\boldsymbol{z}_{-u}) + B\lambda_d f(\boldsymbol{z})\Bigr)^2 \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z}$$
$$= \sigma^2(A^2 + B^2 + (A+B)^2) - \underline{\tau}_u^2 B(A+B) - \underline{\tau}_{-u}^2 A(A+B)$$

To eliminate σ^2 and $\underline{\tau}_{-u}^2$ we need $A = B = 0$.

Substitution $\underline{\tau}_{-u}^2 = \sigma^2 - \overline{\tau}_u^2$ does not help.

Introducing more terms inside the square means more terms that need to cancel (e.g. $\underline{\tau}_v^2$ for $v \neq u$)

Weighted sums of squares

$$\sum_{r=1}^{R} \alpha_r \iint \Bigl(\sum_{u \subseteq \mathcal{D}} \lambda_{r,u} f(\boldsymbol{x}_u{:}\boldsymbol{z}_{-u}) \Bigr)^2 \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{z}$$

The coefficient of $\sigma^2_{\mathcal{D}}$ is:

$$\sum_{r=1}^{R} \alpha_r \sum_{u \subseteq \mathcal{D}} \lambda^2_{r,u}$$

We cannot make this 0 without $\alpha_r < 0$ (or trivially $\lambda_{r,u} = 0$ or all $\alpha_r = 0$)

So $\cdots$ no unbiased nonnegative estimate of $\underline{\tau}^2_u$. (checkmate)

Bilinear, with $O(d)$ evaluations

Suppose $\lambda_u = 0$ for $|u| \notin \{0, 1, d-1, d\}$. Same for $\gamma_v = 0$.

Then the rule

$$\sum_u \sum_v \lambda_u \gamma_v \iint f(\boldsymbol{x}_u{:}\boldsymbol{z}_{-u}) f(\boldsymbol{x}_v{:}\boldsymbol{z}_{-v}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{z}$$

takes $O(d)$ computation $\cdots$ not $O(d^2)$.

For $j \neq k$, let j represent $\{j\}$ and $-j$ represent $-\{j\}$ etc.

$O(d)$ pairs, with $k \neq j$

XOR	$\varnothing$	j	k	$-j$	$-k$	$\mathcal{D}$
$\varnothing$	$\varnothing$	j	k	$-j$	$-k$	$\mathcal{D}$
j	j	$\varnothing$	$\{j,k\}$	$\mathcal{D}$	$-\{j,k\}$	$-j$
$-j$	$-j$	$\mathcal{D}$	$-\{j,k\}$	$\varnothing$	$\{j,k\}$	j
$\mathcal{D}$	$\mathcal{D}$	$-j$	$-k$	j	k	$\varnothing$

NXOR	$\varnothing$	j	k	$-j$	$-k$	$\mathcal{D}$
$\varnothing$	$\mathcal{D}$	$-j$	$-k$	j	k	$\varnothing$
j	$-j$	$\mathcal{D}$	$-\{j,k\}$	$\varnothing$	$\{j,k\}$	j
$-j$	j	$\varnothing$	$\{j,k\}$	$\mathcal{D}$	$-\{j,k\}$	$-j$
$\mathcal{D}$	$\varnothing$	j	k	$-j$	$-k$	$\mathcal{D}$

$\underline{\tau}^2_{\mathrm{NXOR}(u,v)}$

Assuming $f_\varnothing = 0$ (WLOG if $\sum_u \lambda_u = 0$)

$$
\begin{array}{c}
\underline{\tau}^2_{\mathrm{NXOR}} \\ \varnothing \\ j \\ -j \\ \mathcal{D}
\end{array}
\begin{array}{c}
\begin{array}{cccccc}
\varnothing & j & k & -j & -k & \mathcal{D}
\end{array} \\
\left[\begin{array}{cccccc}
\sigma^2 & \underline{\tau}^2_{-j} & \underline{\tau}^2_{-k} & \underline{\tau}^2_{j} & \underline{\tau}^2_{k} & 0 \\
\underline{\tau}^2_{-j} & \sigma^2 & \underline{\tau}^2_{-\{j,k\}} & 0 & \underline{\tau}^2_{\{j,k\}} & \underline{\tau}^2_{j} \\
\underline{\tau}^2_{j} & 0 & \underline{\tau}^2_{\{j,k\}} & \sigma^2 & \underline{\tau}^2_{-\{j,k\}} & \underline{\tau}^2_{-j} \\
0 & \underline{\tau}^2_{j} & \underline{\tau}^2_{k} & \underline{\tau}^2_{-j} & \underline{\tau}^2_{-k} & \sigma^2
\end{array}\right]
\end{array}
$$

$$\sum_j \underline{\tau}^2_j = \sum_j \sigma^2_j \qquad \sum_j \sum_{k \neq j} \underline{\tau}^2_{-\{j,k\}} = d(d-1)\sigma^2 - \sum_u (2(d-1) - |u|)|u|\sigma^2_u$$

$$\sum_j \underline{\tau}^2_{-j} = d\sigma^2 - \sum_u |u|\sigma^2_u \qquad \sum_j \sum_{k \neq j} \underline{\tau}^2_{\{j,k\}} = \sum_{u:|u|=2} \sigma^2_u + 2\sum_j \sigma^2_j$$

Liu & O (2006) Theorem 2

Using $O(d)$ terms

We can estimate

$$\sum_u \sigma_u^2 1_{|u|=1}$$

$$\sum_u \sigma_u^2 1_{|u|=2} \qquad !$$

$$\sum_u |u|\sigma_u^2$$

$$\sum_u |u|^2\sigma_u^2$$

Another $O(d)$ quantity

Largest element in u:

$$\lceil u \rceil = \begin{cases} \max\{j \mid j \in u\}, & u \neq \varnothing \\ 0, & u = \varnothing. \end{cases}$$

Then

$$\begin{aligned} \sum_{j=1}^{d-1} \overline{\tau}^2_{\{1,2,\dots,j\}^c} &= \sum_u \sigma_u^2 \sum_{j=1}^{d-1} 1_{u \cap \{1,\dots,j\}^c \neq \varnothing} \\ &= \sum_u \sigma_u^2 \sum_{j=1}^{d-1} 1_{u \cap \{j+1,\dots,d\} \neq \varnothing} \\ &= \sum_u \sigma_u^2 (\lceil u \rceil - 1). \end{aligned}$$

A mean dimension in the truncation sense.

Easier to compute than effective dimension.

Optimal estimates

Sobol's estimates have been improved (!!) recently:

Kucherenko, Feil, Shah, Mauntz (2011), and

Janon, Klein, Lagnoux, Nodet & Prieur (2012) (Grenoble)

Let $\eta^2 = \sum_u \delta_u \sigma_u^2$.

We would like

$$\mathbb{E}(\hat{\eta}^2) = \eta^2 \quad \text{and,} \quad \mathrm{Var}(\hat{\eta}^2) = \text{minimum.}$$

Using variance components theory

Unfortunately $\mathrm{Var}(\hat{\eta}^2)$ depends on 4'th moments

Fortunately There is a theory of **MIN**imum **Q**uadratic **N**orm **UN**biased **E**stimates (MINQUE)*

Unfortunately They do not appear to be available for crossed random effects

Fortunately We can choose where to sample and our estimator.

*C. R. Rao (1970s)

Speculation

For all for one $\sum_u \lambda_u \underline{\tau}_v^2$

$$\begin{aligned} \text{minimize} \quad & \sum_u \lambda_u^2 \\ \text{subject to} \quad & \sum_u \lambda_u \underline{\tau}_u^2 = \sum_u \delta_u \sigma_u^2 \\ \text{and} \quad & \sum_u \lambda_u = 0. \end{aligned}$$

This ignores $\#$ of function evaluations. So instead

$$\text{minimize} \quad \Bigl(\sum_u \lambda_u^2\Bigr) \times \Bigl(\sum_u 1_{\lambda \neq 0}\Bigr) = \|\lambda\|_2^2 \times \|\lambda\|_0$$

Merci, la deuxieme fois

- GDR Coordinators: Clémentine Prieur, Bertrand Iooss, Fabien Mangeant
- Scientific and organizing committee
- Françoise Poggi
- National Science Foundation of the U.S. DMS-0906056

MCQMC 2014

Please come to Leuven for MCQMC 2014.

Sensitivity for extremes

Gary Tang asked about sensitivity measures that are more attuned to extreme values of $f(\boldsymbol{x})$.

Some joint work with Josef Dick:

1) Transform $f(\boldsymbol{x})$ (don't like)

2) Analysis of skewness $\int f(\boldsymbol{x})^3 \,\mathrm{d}\boldsymbol{x}$ (don't like either)

3) Analysis of fourth moment $\int f(\boldsymbol{x})^4 \,\mathrm{d}\boldsymbol{x}$ (don't like either)

4) Estimate $\int f_u(\boldsymbol{x})^4 \,\mathrm{d}\boldsymbol{x}$ (like much more, still testing!)