Working with Henry Wynn: Over 30 years of collaboration in a few slides (and photos)

Luc Pronzato (CNRS, Université Côte d'Azur, France)

and Anatoly Zhigljavsky (Cardiff University, UK)

June 2026

1 Introduction

A short extract from our contribution to the paper
"*Working with Henry Wynn: a challenge and an exceptional opportunity*",
with Alessandra Giovagnoli, Hugo Maruri-Aguilar and Eva Riccomagno

1 Introduction

A short extract from our contribution to the paper "*Working with Henry Wynn: a challenge and an exceptional opportunity*", with Alessandra Giovagnoli, Hugo Maruri-Aguilar and Eva Riccomagno

2 parts (not mentioning several European projects on "concurrent engineering")

1. Dynamical systems in search and optimisation ($\approx$ 1991–2008, $\rightarrow$ mODa 3 & 4)
2. Simplicial measures of multivariate dispersion ($\approx$ 2015–2024, $\rightarrow$ mODa 13)

First meetings (early 90's) in London, then in Antibes...

A pause between two (intensive) working sessions:

2 Dynamical systems in search and optimisation

Objective: search for a target $\mathbf{x}^*$ in a given region R_0

[for instance, $\mathbf{x}^* =$ the minimiser of a function $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$]

2 Dynamical systems in search and optimisation

Objective: search for a target $\mathbf{x}^*$ in a given region R_0

[for instance, $\mathbf{x}^* =$ the minimiser of a function $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$]

Consistency

→ "Observations" (evaluations of f) force restrictions on possible location of $\mathbf{x}^*$:

After n "observations" $y_1, \ldots, y_n$, we know that $\mathbf{x}^*$ belongs to a *consistency region* $R_n = \mathscr{X}(y_1, \ldots, y_n)$ (may have a complicated form)

2 Dynamical systems in search and optimisation

Objective: search for a target $\mathbf{x}^*$ in a given region R_0

[for instance, $\mathbf{x}^* =$ the minimiser of a function $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$]

Consistency

→ "Observations" (evaluations of f) force restrictions on possible location of $\mathbf{x}^*$:

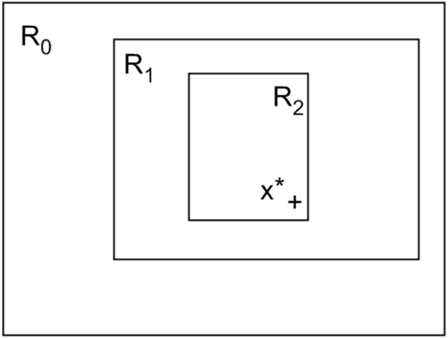
After n "observations" $y_1, \ldots, y_n$, we know that $\mathbf{x}^*$ belongs to a *consistency region* $R_n = \mathscr{X}(y_1, \ldots, y_n)$ (may have a complicated form)

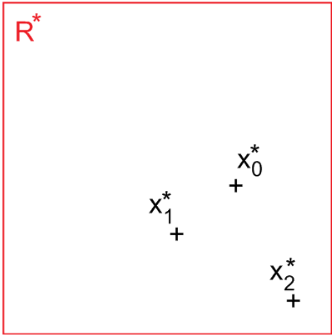
Renormalisation

Renormalise R_n (or a simpler outer approximation of it)

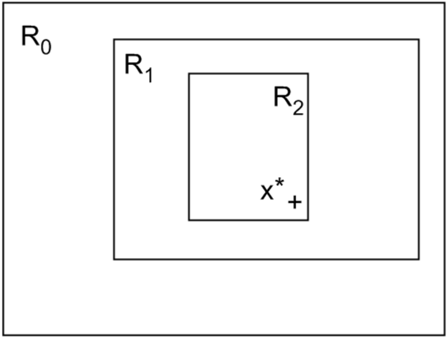
to a base region R^*

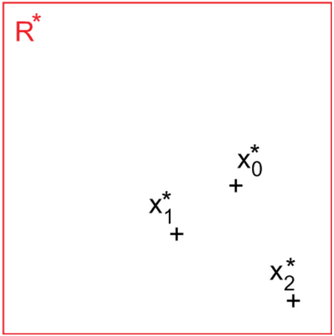
consistency ➠ renormalisation





consistency ➠ renormalisation





$\mathbf{x}^*$ fixed in R_0 ➠ $\boxed{\text{dynamical system } \mathbf{x}_0^* \to \mathbf{x}_1^* \to \mathbf{x}_2^* \to \cdots \text{ in } R^* \text{ fixed}}$

Consequences

Rate of convergence = rate of contraction of the R_i
= rate of expansion of the dynamical system
(Lyapunov exponents, entropies)

Ergodic behaviour, similar for almost any initialisation of the algorithm
→ ergodically, worst-cases events have measure zero
(we can improve upon algorithms considered as worst-case optimal)

The book (P, W & Z, 2000)* tells all the story...

* P, W & Z, 2000. *Dynamical Search*. Chapman & Hall/CRC, Boca Raton

$d = 1$: Second-order line search

Objective: minimise f having a unique minimiser x^* in an interval $[A_0, B_0)$

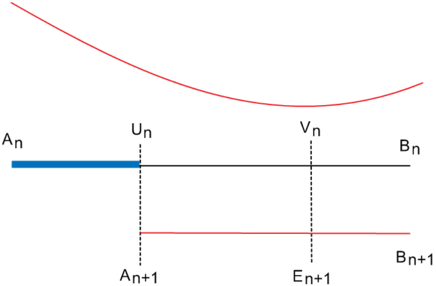
$\rightarrow$ compare $f(\cdot)$ at U_n and V_n in $[A_n, B_n)$, $U_n < V_n$

$d = 1$: Second-order line search

Objective: minimise f having a unique minimiser x^* in an interval $[A_0, B_0)$

→ compare $f(\cdot)$ at U_n and V_n in $[A_n, B_n)$, $U_n < V_n$

If $f(U_n) \geq f(V_n)$, delete $[A_n, U_n)$ (Left deletion)

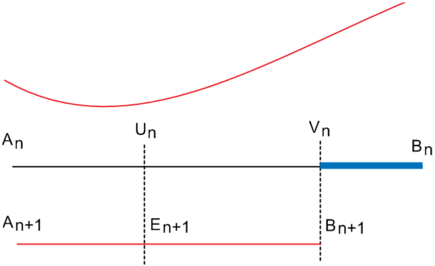


$d = 1$: Second-order line search

Objective: minimise f having a unique minimiser x^* in an interval $[A_0, B_0)$

➞ compare $f(\cdot)$ at U_n and V_n in $[A_n, B_n)$, $U_n < V_n$

If $f(U_n) < f(V_n)$, delete $[V_n, B_n)$ (Right deletion)

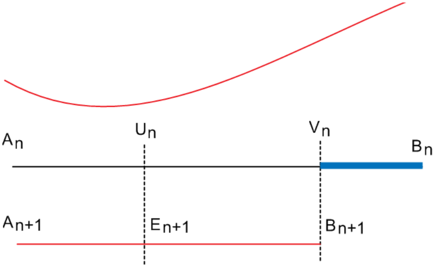


$d = 1$: Second-order line search

Objective: minimise f having a unique minimiser x^* in an interval $[A_0, B_0)$

➞ compare $f(\cdot)$ at U_n and V_n in $[A_n, B_n)$, $U_n < V_n$

If $f(U_n) < f(V_n)$, delete $[V_n, B_n)$ (Right deletion)



Next interval $[A_{n+1}, B_{n+1})$ contains U_n or V_n,

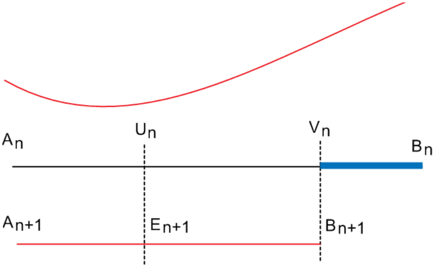
➞ let E_{n+1} denote this point; choose a new point E'_{n+1} in $[A_{n+1}, B_{n+1})$; repeat...

$d = 1$: Second-order line search

Objective: minimise f having a unique minimiser x^* in an interval $[A_0, B_0)$

→ compare $f(\cdot)$ at U_n and V_n in $[A_n, B_n)$, $U_n < V_n$

If $f(U_n) < f(V_n)$, delete $[V_n, B_n)$ (Right deletion)



Next interval $[A_{n+1}, B_{n+1})$ contains U_n or V_n,

→ let E_{n+1} denote this point; choose a new point E'_{n+1} in $[A_{n+1}, B_{n+1})$; repeat...

Renormalise $[A_n, B_n)$ to $R^* = [0, 1)$: $A_n \to 0$, $B_n \to 1$, $E_n \to e_n$, $E'_n \to e'_n$

Golden Section algorithm (Kiefer, 1957)*:

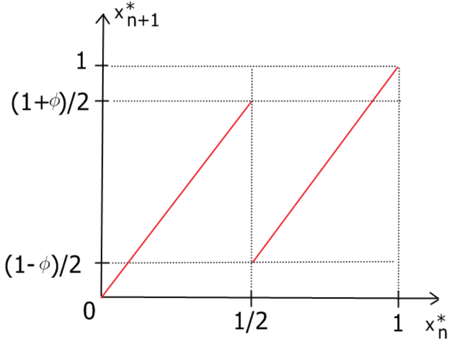
$$x^*_{n+1} = \begin{cases} (1+\phi)x^*_n & \text{if } x^*_n < 1/2 \\ (1+\phi)x^*_n - \phi & \text{if } x^*_n \geq 1/2 \end{cases}$$

* Kiefer, J., 1957. Optimum sequential search and approximation methods under minimum regularity assumptions. *J. Soc. Indust. Appl. Math.* **5** (3), 105–136

Golden Section algorithm (Kiefer, 1957)*:

$$x^*_{n+1} = \begin{cases} (1+\phi)x^*_n & \text{if } x^*_n < 1/2 \\ (1+\phi)x^*_n - \phi & \text{if } x^*_n \geq 1/2 \end{cases}$$

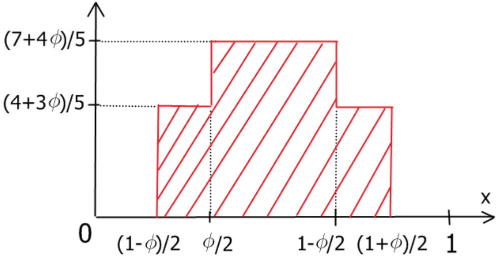
GS iteration



Golden Section algorithm (Kiefer, 1957)*:

$$x_{n+1}^* = \begin{cases} (1+\phi)x_n^* & \text{if } x_n^* < 1/2 \\ (1+\phi)x_n^* - \phi & \text{if } x_n^* \geq 1/2 \end{cases}$$

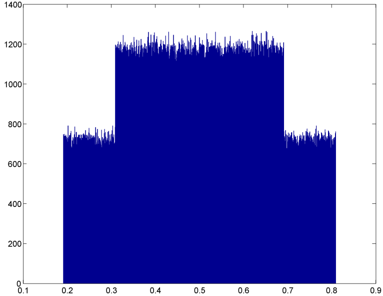
invariant density for the GS algorithm



Golden Section algorithm (Kiefer, 1957)*:

$$x^*_{n+1} = \begin{cases} (1+\phi)x^*_n & \text{if } x^*_n < 1/2 \\ (1+\phi)x^*_n - \phi & \text{if } x^*_n \geq 1/2 \end{cases}$$

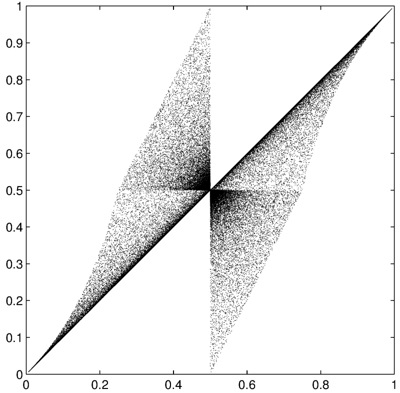
invariant density for the GS algorithm



Other second-order line-search algorithms?

Symmetric algorithms: $\boxed{e'_n = 1 - e_n}$ for all n (with GS as a particular case)

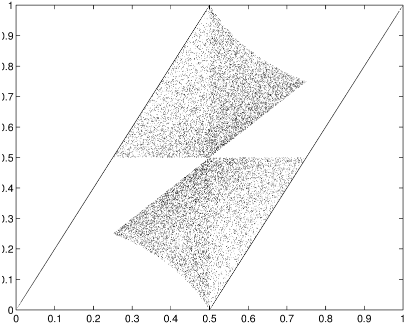
(x_n^*, x_{n+1}^*), $n = 1, \ldots, 100\,000$, for a typical symmetric algorithm



Midpoint algorithm: $\boxed{e'_n = 1/2}$ for all n

(assume $f(\cdot)$ locally symmetric at x^*)

(x^*_n, x^*_{n+1}), $n = 1, \ldots, 50\,000$

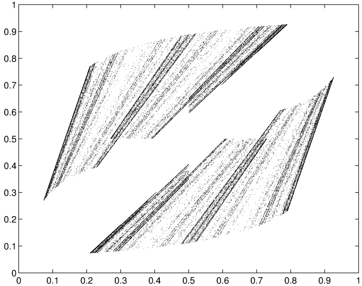


(faster than GS)

Window algorithm: $\boxed{e_n' = e_n + w \text{ if } e_n < \frac{1}{2} \text{ and } e_n' = e_n - w \text{ if } e_n \geq \frac{1}{2}}$

(assume $f(\cdot)$ locally symmetric at x^*)

(x_n^*, x_{n+1}^*), $n = 1, \ldots, 50\,000$ ($w = 1/8$)

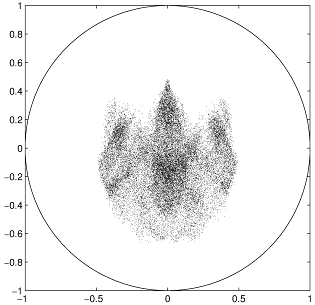


(the fastest we found)

$d > 1$: Ellipsoid algorithm?

Khachiyan (1979)* algorithm for linear programming:

Typical sequence $\mathbf{x}_n^*$



→ too complicated!

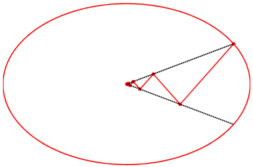
* Khachiyan, L., 1979. A polynomial algorithm in linear programming. *Doklady AkademïaNauk SSSR* **244**, 1093–1096, translated into English in *Soviet Mathematics Doklady*, **20**, 191–194

$d > 1$: Steepest descent to minimise a quadratic function!

SD: $\boxed{\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha_k \nabla_f(\mathbf{x}^{(k)})}$, with $\alpha_k = \arg\min_\alpha f[\mathbf{x}^{(k)} - \alpha \nabla_f(\mathbf{x}^{(k)})]$

Quadratic function: $f(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^\top \mathbf{A}(\mathbf{x} - \mathbf{x}^*)$, $\mathbf{A}$ pos. def.

Known result in $\mathbb{R}^d$ (Akaike, 1959)*
(Akaike's bird cage)

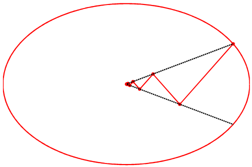


* Akaike, H., 1959. On a successive transformation of probability distribution and its application to the analysis of the optimum gradient method. *Ann. Inst. Statist. Math. Tokyo* **11**, 1–16

$d > 1$: Steepest descent to minimise a quadratic function!

SD: $\boxed{\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha_k \nabla_f(\mathbf{x}^{(k)})}$, with $\alpha_k = \arg\min_\alpha f[\mathbf{x}^{(k)} - \alpha \nabla_f(\mathbf{x}^{(k)})]$

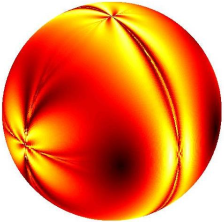
Quadratic function: $f(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^\top \mathbf{A}(\mathbf{x} - \mathbf{x}^*)$, $\mathbf{A}$ pos. def.



Known result in $\mathbb{R}^d$ (Akaike, 1959)*
(Akaike's bird cage)

...but detailed behaviour is rather complicated even in finite (and small) dimension

$d = 3$, $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 4$
the colour (of starting point on $\mathbb{S}_2(0,1)$) depends on limiting p



* Akaike, H., 1959. On a successive transformation of probability distribution and its application to the analysis of the optimum gradient method. *Ann. Inst. Statist. Math. Tokyo* **11**, 1–16

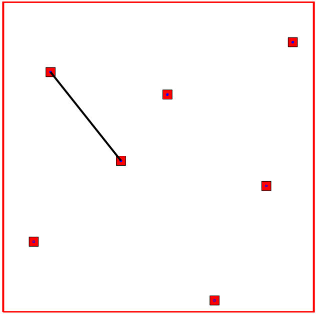
Another pause between two (intensive) working sessions:

3 Simplicial measures of multivariate dispersion

Take n points x_i i.i.d. (prob. measure μ) in $\mathbb{R}^d$
Compute squared distances between all pairs :

$$\frac{2}{n(n-1)} \sum_{i<j} \|x_i - x_j\|^2 = 2\,\mathrm{trace}[\widehat{\mathbf{V}}_n]$$

with $\widehat{\mathbf{V}}_n$ the empirical covariance matrix

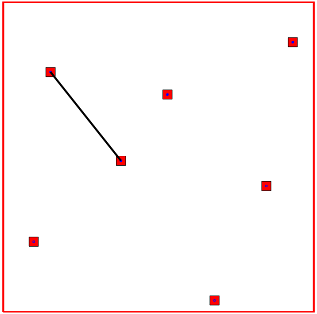


3 Simplicial measures of multivariate dispersion

Take n points x_i i.i.d. (prob. measure μ) in $\mathbb{R}^d$
Compute squared distances between all pairs :

$$\frac{2}{n(n-1)} \sum_{i<j} \|x_i - x_j\|^2 = 2\,\mathrm{trace}[\widehat{\mathbf{V}}_n]$$

with $\widehat{\mathbf{V}}_n$ the empirical covariance matrix

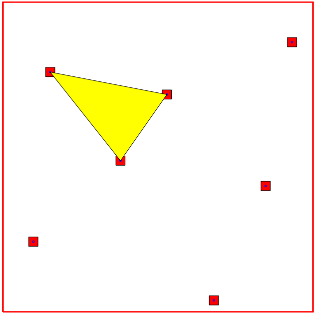


$$\int\int \|x_1 - x_2\|^2 \,\mu(\mathrm{d}x_1)\mu(\mathrm{d}x_2) = 2\,\mathrm{trace}(\mathbf{V}_\mu)$$

Take n points x_i i.i.d. (prob. measure μ) in $\mathbb{R}^d$

Compute squared area of triangles for all triples :

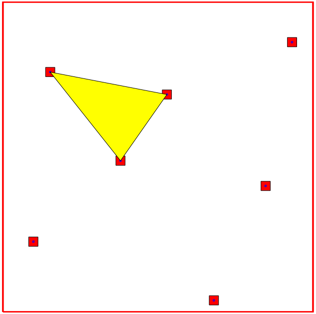
$$\frac{6}{n(n-1)(n-2)} \sum_{i<j<k} \mathscr{V}_2^2(x_i, x_j, x_k) = ?$$



Take n points x_i i.i.d. (prob. measure μ) in $\mathbb{R}^d$

Compute squared area of triangles for all triples :

$$\frac{6}{n(n-1)(n-2)} \sum_{i<j<k} \mathscr{V}_2^2(x_i, x_j, x_k) = \, ?$$



$$\psi_2(\mu) = \int\int\int \mathscr{V}_2^2(x_1, x_2, x_3)\, \mu(\mathrm{d}x_1)\, \mu(\mathrm{d}x_2)\, \mu(\mathrm{d}x_3) = ?$$

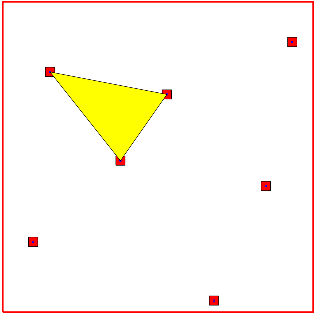
What about volumes of k-dimensional simplices in $\mathbb{R}^d$, $k \leq d$?

Which properties in terms of functionals of μ?

Take n points x_i i.i.d. (prob. measure μ) in $\mathbb{R}^d$

Compute squared area of triangles for all triples :

$$\frac{6}{n(n-1)(n-2)} \sum_{i<j<k} \mathscr{V}_2^2(x_i, x_j, x_k) = \ ?$$



$$\psi_2(\mu) = \int\int\int \mathscr{V}_2^2(x_1, x_2, x_3)\, \mu(\mathrm{d}x_1)\, \mu(\mathrm{d}x_2)\, \mu(\mathrm{d}x_3) = ?$$

What about volumes of k-dimensional simplices in $\mathbb{R}^d$, $k \leq d$?

Which properties in terms of functionals of μ?

What about using other powers δ:

$$\psi_{k,\delta}(\mu) = \int \cdots \int \mathscr{V}_k^\delta(x_1, \ldots, x_{k+1})\, \mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_{k+1}) \ ?$$

$\delta = 2$: $\mu \in \mathscr{M} \longrightarrow \psi_{k,2}(\mu) = \mathsf{E}_\mu\{\mathscr{V}_k^2(x_1, \ldots, x_{k+1})\}$

$k = 2$ ➠ triangles $\psi_2(\mu) = \int\int\int \mathscr{V}_2^2(x_1, x_2, x_3)\, \mu(\mathrm{d}x_1)\, \mu(\mathrm{d}x_2)\, \mu(\mathrm{d}x_3)$

$$\psi_{k,2}(\mu) = \Psi_{k,2}[\mathbf{V}_\mu] = \frac{k+1}{k!} \sum_{i_1 < i_2 < \cdots < i_k} \lambda_{i_1}[\mathbf{V}_\mu] \times \cdots \times \lambda_{i_k}[\mathbf{V}_\mu]\,,$$

$\psi_{k,2}^{1/k}(\cdot)$ is shift-invariant, homogeneous of degree 2 and concave
→ optimal design, equivalence theorem, etc.

(López-Fidalgo and Rodríguez-Díaz, 1998)[1]
(P, W & Z, 2017)[2], (P, W & Z, mODa 12)

[1] López-Fidalgo, J., Rodríguez-Díaz, J., 1998. Characteristic polynomial criteria in optimal experimental design. In *Proc. of MODA'5*, Marseilles, June 22–26, 1998. Physica Verlag, Heidelberg, pp. 31–38

[2] P, W & Z, 2017. Extended generalised variances, with applications. *Bernoulli* **23** (4A), 2617–2642, arXiv:1411.6428, hal-01308092

$\delta = 2$: $\mu \in \mathscr{M} \longrightarrow \psi_{k,2}(\mu) = \mathsf{E}_\mu\{\mathscr{V}_k^2(x_1,\ldots,x_{k+1})\}$

$k = 2$ ➠ triangles $\psi_2(\mu) = \int\int\int \mathscr{V}_2^2(x_1,x_2,x_3)\,\mu(\mathrm{d}x_1)\,\mu(\mathrm{d}x_2)\,\mu(\mathrm{d}x_3)$

$$\psi_{k,2}(\mu) = \Psi_{k,2}[\mathbf{V}_\mu] = \frac{k+1}{k!} \sum_{i_1<i_2<\cdots<i_k} \lambda_{i_1}[\mathbf{V}_\mu] \times \cdots \times \lambda_{i_k}[\mathbf{V}_\mu]\,,$$

$\psi_{k,2}^{1/k}(\cdot)$ is shift-invariant, homogeneous of degree 2 and concave
→ optimal design, equivalence theorem, etc.
(López-Fidalgo and Rodríguez-Díaz, 1998)[1]
(P, W & Z, 2017)[2], (P, W & Z, mODa 12)

$\delta \neq 2$: $\mu \in \mathscr{M} \longrightarrow \psi_{k,\delta}(\mu) = \mathsf{E}_\mu\{\mathscr{V}_k^\delta(x_1,\ldots,x_{k+1})\}$

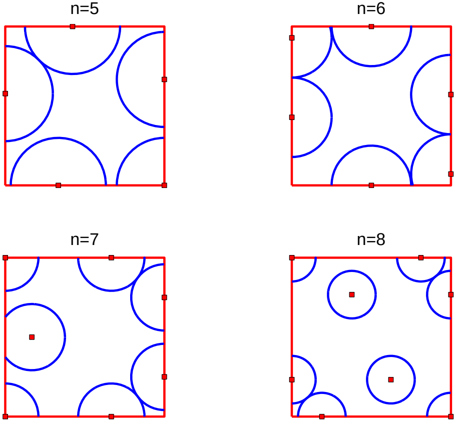
- $\delta > 0$ (maximise $\psi_{k,\delta}$): $k = 1$ in (Björck, 1956)[3], $k \geq 2$ ➠ some results. . .
- $\delta < 0$ (minimise $\psi_{k,\delta}$): $k = 1 \leftrightarrow$ potential theory, $k \geq 2$ largely open. . .

[3] Björck, G., 1956. Distributions of positive mass, which maximize a certain generalized energy integral. *Arkiv för Matematik* **3** (21), 255–269

Example: $\delta = -\infty$ ➞ maximise $\min_{\{x_1,\ldots,x_{k+1}\}\subset\mathbf{X}_n} \mathscr{V}_k(x_1,\ldots,x_{k+1})$

Example: $\delta = -\infty \rightarrow$ maximise $\min_{\{x_1,\ldots,x_{k+1}\}\subset\mathbf{X}_n} \mathscr{V}_k(x_1,\ldots,x_{k+1})$

$$k = d = 2:$$



(P, W & Z ProbaStat'2015)

References

Akaike, H., 1959. On a successive transformation of probability distribution and its application to the analysis of the optimum gradient method. Ann. Inst. Statist. Math. Tokyo 11, 1–16.

Björck, G., 1956. Distributions of positive mass, which maximize a certain generalized energy integral. Arkiv för Matematik 3 (21), 255–269.

Khachiyan, L., 1979. A polynomial algorithm in linear programming. Doklady AkademïaNauk SSSR 244, 1093–1096, translated into English in *Soviet Mathematics Doklady*, 20, 191–194.

Kiefer, J., 1957. Optimum sequential search and approximation methods under minimum regularity assumptions. J. Soc. Indust. Appl. Math. 5 (3), 105–136.

López-Fidalgo, J., Rodríguez-Díaz, J., 1998. Characteristic polynomial criteria in optimal experimental design. In: Atkinson, A., Pronzato, L., Wynn, H. (Eds.), Advances in Model–Oriented Data Analysis and Experimental Design, Proceedings of MODA'5, Marseilles, June 22–26, 1998. Physica Verlag, Heidelberg, pp. 31–38.

Pronzato, L., Wynn, H., Zhigljavsky, A., 2000. Dynamical Search. Chapman & Hall/CRC, Boca Raton.

Pronzato, L., Wynn, H., Zhigljavsky, A., 2017. Extended generalised variances, with applications. Bernoulli 23 (4A), 2617–2642, arXiv:1411.6428, hal-01308092.

Let's celebrate all these years: cheers, Henry!