Optimal designs for stochastic processes whose covariance is a function of the mean

Jesus.LopezFidalgo@uclm.es
M. Amo-Salas
E. Porcu
University of Castilla-La Mancha

OUTLINE

1. Justification of the model.
2. Conditions for positive definite functions.
3. Optimal designs.
4. Illustrative examples.

Justification

- Accident in facilities that handle radioactive materials.
- Modeling the retention in lungs of radioisotope particles:
    - Large sum of exponentials.
- Usual assumption: variance as a function of the mean.
- In the same way, the covariance to be a function of the mean.
- Conditions for the covariance function to be positive definite.

Model

- Observations $z(t_1), \ldots, z(t_n)$ from a stochastic process:

$$Z(t) = \eta(t;\theta) + \varepsilon(t), t \in \mathbb{T}, \theta \in \Theta \subseteq \mathbb{R}^m.$$

- $\varepsilon(\cdot)$, Gaussian process with weakly stationary correlated observations:

$$\mathrm{Cov}[\varepsilon(t), \varepsilon(s)] = \mathrm{C}(h),\ h = |t - s|.$$

- **Remark:** The variance of $\varepsilon(\cdot)$, $\mathrm{C}(0)$, will be constant.
- **Objective:** Modeling the covariance structure through the mean.

The covariance

$$\mathrm{C}(h) = \mathrm{Cov}[Z(t), Z(s)] \quad = \quad \varphi[\eta(h;\theta)],$$

Positive definite function

$\mathrm{C} : \mathbb{R} \to \mathbb{R}$ continuous and even such that for any $\{t_i\}_{i=1}^n \in \mathbb{R}$,

$$\{\mathrm{C}(t_i - t_j)\}_{ij} \geq 0.$$

Aim: Conditions ensuring $h \mapsto \varphi[\eta(h;\theta)]$ is positive definite on $\mathbb{R}$.

Previous results

Completely monotonic function:

$\varphi : [0, \infty) \to (0, \infty)$ with $(-1)^n \varphi^{(n)}(t) \geq 0$ for $0 \leq t$.

Bernstein's theorem

If φ is a completely monotonic function then it is the Laplace transform of a positive and bounded measure μ:

$$\varphi(t) = \int_{(0,\infty)} \exp(-rt)\mu(\mathrm{d}r).$$

Previous results

Polya's type criterion

Let φ continuous and convex, such that $\varphi(0) = 1$, $\varphi(-t) = \varphi(t)$ and $\lim_{t\to\infty} \varphi(t) = 0$. Then there exists a continuous cdf F such that F' exists, is even, continuous everywhere except possibly at $t = 0$ with

$$\varphi(t) = \int_{(-\infty,\infty)} \exp(itx)\mathrm{d}F(x).$$

Bochner's theorem

$\mathrm{C}(\cdot)$ is positive definite if and only if it is the Fourier transform of some non–negative and bounded measure μ:

$$\mathrm{C}(h) = \int_{\mathbb{R}} \mathrm{e}^{ih\omega} \mu(\mathrm{d}\omega).$$

Thus, if φ satisfies Polya's conditions then it is positive definite.

Lemma

Let $t \mapsto \varphi(t) = e^{-t}$, $t \geq 0$ and $\eta(\cdot;\theta) : \mathbb{R} \to (0,\infty)$ continuous, increasing and such that $\lim_{t\to\infty} \eta(t;\theta) = \infty$ and $\varphi \circ \eta(\cdot;\theta)$ convex for $t \geq 0$. Then the composition $\psi(h) = \frac{1}{\varphi[\eta(0;\theta)]}\varphi \circ \eta(\cdot;\theta)$ is positive definite on $\mathbb{R}$.

Proof

$\varphi \circ \eta(\cdot;\theta)$ is nonnegative, decreasing and convex on the positive real line and vanishes at infinity.

$\psi(0) = 1$ and extending the definition of η to the negative real line:

$$\eta(-t;\theta) \equiv \eta(t;\theta),\ t > 0,$$

it meets all the requirements of the Pólya type criterion.

Main result

Theorem

If $\varphi(\cdot)$ is a completely monotonic function with $\varphi(0) = 1$ and η satisfies the conditions given in the Lemma, then

$$\mathrm{C}(h) = \varphi\left[\eta(h, \theta)\right]$$

is a permissible covariance function on $\mathbb{R}$.

Proof

Bernstein's theorem implies φ is the Laplace transform of a positive and bounded measure μ:

$$\varphi(t) = \int_{(0,\infty)} \exp(-rt)\mathrm{d}\mu(r).$$

Therefore,

$$\mathrm{C}(h) = \int_{(0,\infty)} \exp\left\{-r\eta(h,\theta)\right\}\mathrm{d}\mu(r)$$

is a scale mixture of the convex and positive definite functions (Lemma):

$$\psi_r(h) = \exp\left\{-r\eta(h,\theta)\right\}.$$

Positive definite functions form a convex cone (weak convergence top.).

Thus, it is positive definite for any finite and positive measure μ.

Particular completely monotonic functions

(I) Matérn function

$$\varphi(t;\phi,\gamma)=(\phi t)^{\gamma}\mathcal{K}_{\gamma}(\phi t),\ t\geq 0;\ \phi>0,\ \gamma>0,$$

$\mathcal{K}_{\gamma}$ is the modified Bessel function of the second kind.

- Decreases with respect to ϕ.
- Increases for small values of t and decreases after some value of t with respect to γ.

Special cases:

- Exponential model: $\varphi(t;\phi,1/2)=\exp(-\phi t)$.
  Covariance function associated to the Ornstein–Uhlembeck process.
- Gaussian model: $\varphi(t;\phi,\infty)=\exp(-\phi t^2)$.
- Autoregressive process of the third order: $\varphi(t;\phi,3/2)$.

Particular completely monotonic functions

(II) Dagum function

$$\varphi(t;\phi,\gamma) = 1 - \left(\frac{t^{\phi}}{1+t^{\phi}}\right)^{\gamma}, \; t \geq 0;\; 0 < \phi \leq 2,\; 0 < \gamma \leq 1.$$

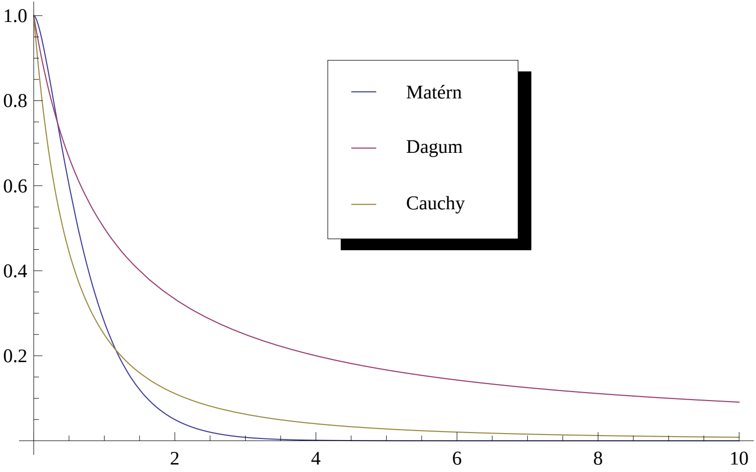
- Increases for small values of t and decreases after some value of t with respect to ϕ.
- Increases with respect to γ.

(III) Cauchy function

$$\varphi(t;\phi,\gamma) = (1+t^{\phi})^{-\gamma},\; t \geq 0;\; 0 < \phi \leq 2,\; \gamma > 0.$$

- Increases for small values of t and decreases after $t = 1$ with respect to ϕ.
- Decreases with respect to *gamma*.

Plots



Optimal designs

Exact design: $\xi = (t_1, \ldots, t_n) \equiv \mathbf{t}$ for a given n.

Fisher information matrix:

$$I(\xi; \theta) = \mathbf{E}_\theta \left[\frac{\partial^2 \log f(z|\mathbf{t}, \theta)}{\partial \theta^2} \right],$$

where $z = (z(t_1), \ldots, z(t_n))'$, $\eta(\mathbf{t}, \theta) = (\eta(t_1, \theta), \ldots, \eta(t_n, \theta))'$.

Log-likelihood:

$$\begin{aligned} \log f(z|t, \theta) &= -\frac{1}{2}\{[z - \eta(t;\theta)]' \Sigma^{-1}(\theta)[z - \eta(t;\theta)] \\ &- \frac{1}{2} \log \det[\Sigma(\theta)] - \frac{n}{2} \log(2\pi)\}. \end{aligned}$$

Information matrix

$$\begin{aligned} I(\xi;\theta) &= \frac{\partial\eta'(t;\theta)}{\partial\theta}\Sigma^{-1}(\theta)\frac{\partial\eta(t;\theta)}{\partial\theta'} \\ &+ \frac{1}{2}tr\{\Sigma^{-1}(\theta)\frac{\partial\Sigma(\theta)}{\partial\theta}\Sigma^{-1}(\theta)\frac{\partial\Sigma(\theta)}{\partial\theta'}\}, \end{aligned}$$

where $\Sigma_{ij}(\theta) = \mathrm{C}(|t_j - t_i|;\theta)$.

Optimal designs

D-optimal design maximizes the determinant.

Design efficiency of ξ:

$$\left(\frac{\det[I(\xi;\theta)]}{\det[I(\xi^*;\theta)]}\right)^{1/m}.$$

Numerical algorithms with correlated observations (Brimkulov et al 1982, Ucinski–Atkinson, 2004).

Illustrative example

$$Z(t) = \eta(t; \alpha_1, \beta_1, \alpha_2, \beta_2) + \varepsilon(t) = \alpha_1 \exp(\beta_1 t) + \alpha_2 \exp(\beta_2 t) + \varepsilon(t),\ t \in \mathbb{T},$$

Design space: $\mathbb{T} = [0, 100]$.

Nominal values for the parameter vector (Locally D–optimal):

$$\alpha_1 = 4 \quad \beta_1 = 0.05 \quad \alpha_2 = 2 \quad \beta_2 = 0.02.$$

Numerically checked that $\exp[-\eta(t; \alpha_1, \beta_1, \alpha_2, \beta_2)]$ is convex on t for the values of the parameter used here.

Fourth point of the design was always *sup* of $\mathbb{T}$

First three points for D–optimal under Matérn

$\phi \setminus \gamma$	0.1			0.5			1			2		
0.1	0	57.9	92.1	0	59.5	92.9	0	64.8	93.0	0	67.5	90.1
0.5	0	57.0	88.6	0	55.2	89.0	0	55.1	88.3	0	56.0	89.5
1	0	54.6	87.3	0	54.8	87.3	0	54.7	87.3	0	54.7	87.3
2	0	54.7	87.3	0	54.7	87.3	0	54.7	87.3	0	54.7	87.3

Optimal design for independent observations: (0, 54.7, 87.3, 100).

Both *sup* and *inf* of $\mathbb{T}$ appear in all the designs.

First three points for D–optimal under Dagum–type

$\phi \setminus \gamma$	0.1			0.5			1		
0.1	5.1	62.8	92.3	5.1	63.0	92.4	7.1	63.6	92.4
0.5	0	61.5	92.3	0	61.7	92.2	5.1	62.3	92.3
1	0	58.8	91.6	0	58.8	91.6	0	59.0	91.7
2	0	55.9	89.6	0	56.8	89.7	0	56.4	89.7

Optimal design for independent observations: (0, 54.7, 87.3, 100).

Increasing values of γ is related to increasing strength of the correlation.

First three points for D–optimal under Cauchy

$\phi \setminus \gamma$			0.1			0.5			1			2
0.1	11.0	63.4	92.6	8.3	62.8	92.2	7.0	62.2	92.3	2.2	62.3	92.3
0.5	8.5	63.2	92.4	0	62.5	92.3	0	62.0	92.3	0	59.2	91.9
1	4.2	63.1	92.4	0	61.5	92.4	0	59.0	91.7	0	56.6	90.1
2	0	62.2	92.3	0	58.5	91.4	0	56.1	89.6	0	55.1	87.8

Optimal design for independent observations: (0, 54.7, 87.3, 100).

Remarks

- The *sup* of the design space appears in all of the designs.
- Intermediate observations decrease slightly with the correlation.
- The *inf* appears when the strength of the correlation decreases.
- All the designs converge to the optimal design for independent observations when the correlation decreases:

$$(0,\ 54.7,\ 87.3,\ 100).$$

- The designs do not differ much for similar degrees of correlation.

Sensitivity against miss–specification of the parameters

Three values for each parameter: $\theta^{(0)}$, $\theta^{(0)} - \sigma_\theta$, $\theta^{(0)} + \sigma_\theta$.

σ_θ, estimated standard deviation obtained under the initial design.

- $3^4 \times 4 \times 4 = 1296$ combinations for Matérn and Cauchy.
- $3^4 \times 4 \times 3 = 972$ combinations for Dagum.

Fractional orthogonal design of size 9 for each combination of the covariance parameters:

- Computation of 144 $(= 9 \times 4 \times 4)$ designs for Matérn and Cauchy.
- Computation of 81 $(= 9 \times 4 \times 3)$ designs for the Dagum.

Results: relative efficiencies higher than 95%.

Miss–specification of the covariance structure parameters

- Matérn: around 50% in some cases.
- Dagum: all the cases more than 88%.
- Cauchy: State of nature $\phi = 0.1$, $\gamma = 0.1$ (very strong correlation) and nominal values $\phi_0 = 2$ and $\gamma_0 = 2 \Rightarrow$ relative efficiency around 82%.

  State of nature $\phi = \gamma = 2$ and nominal values $\phi_0 = 0.1$ and $\gamma_0 = 0.1 \Rightarrow$ relative efficiency around 66 %.

Miss–specification of the covariance structure

- Cauchy covariance ensures the highest level of robustness: For a true Dagum covariance efficiencies up to 90%.

  It decreases only in the extreme case assuming a double source of miss–specification: state of nature with a low correlation and nominal strong correlation (55–85%).
- Analogous considerations for Dagum and Matérn covariances.

Real case: radiation retention model in the human body

The observations need to be made on the same subject.

Exponential and triangular structures independent of the mean considered before.

Model,

$$Z(t;r,p) = \eta(t;r,p) + \epsilon(t) = r\sum_i \gamma_i e^{\alpha_i p + \beta_i t} + \epsilon(t),\ \ t \in \mathbb{T} = [0.5, 100]$$

68 summands with γ_i, α_i, β_i known and fixed.

r, quantity of absorbed radiation.

p, magnitude, shape and density of radioactive particles.

Radiation retention model in the human body

Nominal values: $r_0 = 1000$ and $p_0 = 5$.

Reparametrization of the model to be non–decreasing:

$$\eta'(t; r, p) = \eta(0; r, p) - \eta(t; r, p),$$

For computational reasons consider 4 summands:

$$\eta(t; r, p) = r(2.52e^{-0.93p-0.027t} + 3.66e^{-0.89p-0.001t} + 5.77e^{-174.97p-49.69t} + 3.77e^{-1.07p-4.14t}).$$

Numerically checked that $\exp[-\eta'(t; r, p)]$ is convex on t for the values of the parameter used here.

Two–point D–optimal designs for the Cauchy case

$\phi \setminus \gamma$	0.1	0.5	1	2
0.1	0.5 84.61	0.5 82.40	0.5 81.36	0.5 82.10
0.5	0.5 81.16	0.5 74.13	0.5 66.41	0.5 50.57
1	0.5 80.85	0.5 64.21	1.5 2.5	1.5 2.5
2	0.5 81.51	1.5 2.5	1.5 2.5	1.5 2.5

Two–point D–optimal designs for the Dagum case

$\phi \setminus \gamma$	0.1		0.5		1	
0.1	0.5	85.17	0.5	82.75	0.5	81.36
0.5	0.5	84.14	0.5	70.18	0.5	66.41
1	0.5	67.10	1.5	2.5	1.5	2.5
2	1.5	2.5	1.5	2.5	1.5	2.5

Two–point D–optimal designs for the Matern case

$\phi \setminus \gamma$	0.1	0.5	1	2
0.1	1.5 2.92	0.5 86.53	0.5 91.25	0.5 89.77
0.5	1.5 2.5	1.5 2.5	0.5 67.62	0.5 84.85
1	1.5 2.5	1.5 2.5	1.5 2.5	1.5 2.5

Remarks

- When the strength of the correlation decreases, in all the cases the optimal design converges to $\xi = (1.5, 2.5)$.
- The optimal design is characterized by the presence of the *inf* and the other points move in a neighborhood of 70 days.
- Major strength of the correlation is related to a major distance between the points of the design.
- Sensitivity analysis:
  - All the designs are robust with respect to the nominal value of the parameter *p*.
  - Miss–specification induced by the covariance structure: around 85% when the strength of correlation is similar.
  - Very low efficiency is obtained if the miss–specification source is high in terms of strength of correlation.
  - Designs with non contiguous points exhibit higher efficiency (around 80%). Possible explanation: presence of the *inf*.

General remarks

- Analogue with a nonconstant variance (e.g. depending on the trend structure): nonstationary.
- The theorem is not true if φ is only convex instead of completely monotonic:

$$\varphi(t) = (1 - t)_+ \text{ is convex, but } \left(1 - \alpha \mathrm{e}^{\beta x}\right)_+ \text{ is not convex on } \mathbb{R}_+.$$

References

- Brimkulov, U.N., Krug, G.K. and Savanov, V.L., *Design of experiments in investigating random fields and processes*, Moskow, Nanko (in Russian), 1986.
- Dette, H, Wong, W.K. Optimal designs when the variance is a function of the mean. *Biometrics* 55 (3), 925 — —929, 1999.
- López–Fidalgo, J.F., Rodríguez–Díaz, J.M., Sánchez, G. and Santos–Martín, M.T., Optimal designs for compartmental models with correlated observations, *Journal of Applied Statistics* 32 (10), 1075 — —1088, 2005.
- López–Fidalgo, J.F. and J.F., Villarroel, J., Optimal designs for models of radiation retention with Poisson correlated response, *Statistics in Medicine* 26(9), 1999 — —2016, 2007.
- Amo–Salas, M., López–Fidalgo, J.F. and Rodríguez–Díaz, J.M., Optimizing the test power for a radiation retention model in the human body, *Pharmaceutical Statistics* 9 55 — —66, 2010.
- Pázman, A., Correlated optimum design with a parametrized covariance function, Report no.5, Dept. of Statistics and Mathematics, Wirtschaftsuniversität Wien, June 2004.
- Ucinski, D. and Atkinson, A.C., Experimental design for time-dependent models with correlated observations, *Studies in Nonlinear Dynamics & Econometrics*, 8 (2), 2004.
- Stehlík, M., Rodríguez–Díaz, J.M., Müller, W.G. and López Fidalgo, J.F., Optimal allocation of bioassays in the case of parametrized covariance functions: an application to lung's retention of radioactive particles, *Test*, 17, 56–68, 2008.
- Amo-Salas, M., López–Fidalgo, J.F. and López–Ríos V., Optimal designs for two nested pharmacokinetic models with correlated observations, *Journal of Statistical Planning and Inference*, in press.
- Pólya, G., Remarks on Characteristic Functions, *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 115–123, 1949.
- Zastavnyi, V. and Porcu, E., Compactly supported space-time covariance functions, *Bernoulli Journal*, 17(1), 456Ű465, 2011.
- Matérn, B., *Spatial Variation*, Springer, Berlin, 2nd edition, 1986.
- Berg, C., Mateu, J. and Porcu, E., The Dagum family of isotropic correlation functions, *Bernoulli*, 14(4), 1134–49, 2008.