Beyond univariate Bayes spaces

DDA Meeting, Girona, 3 June 2024

Karel Hron

Department of Mathematical Analysis and Applications of Mathematics

Faculty of Science – Palacký University, Olomouc, Czech Republic

Outline

How to decompose multivariate densities?

- Decomposition of multivariate probability density functions (PDFs) in order to reveal dependence structure among random variables belongs to key tasks in probability theory and multivariate statistics
- One well-established approach is represented by copula theory (Sklar, 1959), but other(s) are searched for in context of functional data analysis of PDFs as data objects
- Could Bayes spaces (Egozcue et al., 2006; van den Boogaart, 2014) be of use here?

How to decompose multivariate densities?

Review paper by Petersen, Zhang and Kokoszka (2022):

Looking beyond univariate densities, one often encounters multivariate distributions... Extension of the models and methods discussed in the review to this setting is nontrivial at best... The Bayes space representation (...) provides a sound theoretical base for multivariate densities, although its practical implementation and utility has only been given limited, if any, consideration...

ELSEVIER

Econometrics and Statistics
Volume 21, January 2022, Pages 159-178

Modeling Probability Density Functions as Data Objects

Alexander Petersen [a b], Chao Zhang [b], Piotr Kokoszka [c]

Some inspiration and key ingrediences

- An orthogonal decomposition was constructed for compositional cubes as multi-factorial generalization of compositional data (Fačevicová et al., 2022) – aka *multivariate discrete distributions* – next step is to proceed to the general case and provide a solid theoretical framework

Some inspiration and key ingrediences

- An orthogonal decomposition was constructed for compositional cubes as multi-factorial generalization of compositional data (Fačevicová et al., 2022) – aka *multivariate discrete distributions* – next step is to proceed to the general case and provide a solid theoretical framework
- This is achieved with the help of the Hoeffding-Sobol decomposition formula (Hoeffding, 1948), known from copula theory
- A general framework for orthogonal decomposition of PDFs is developed in Genest et al. (2023) by following previous results in Mercadier et al. (2022)

Bivariate PDFs

- The domain Ω is a Cartesian product of two domains Ω_X and Ω_Y, i.e., $\Omega = \Omega_X \times \Omega_Y$. The reference measure λ can be decomposed as a product measure $\lambda = \lambda_X \times \lambda_Y$ which implies again the *Hilbert space structure* of the Bayes space $\mathcal{B}^2(\lambda)$

Bivariate PDFs

- The domain Ω is a Cartesian product of two domains Ω_X and Ω_Y, i.e., $\Omega = \Omega_X \times \Omega_Y$. The reference measure λ can be decomposed as a product measure $\lambda = \lambda_X \times \lambda_Y$ which implies again the *Hilbert space structure* of the Bayes space $\mathcal{B}^2(\lambda)$

$\Rightarrow$ *Perturbation* and *powering* operations: $f, g \in \mathcal{B}^2(\lambda)$, $\alpha \in \mathbb{R}$

$$(f \oplus g)(x, y) =_{\mathcal{B}^2(\lambda)} f(x, y) \cdot g(x, y), \; (\alpha \odot f)(x, y) =_{\mathcal{B}^2(\lambda)} f(x, y)^{\alpha}$$

Bivariate PDFs

- The domain Ω is a Cartesian product of two domains Ω_X and Ω_Y, i.e., $\Omega = \Omega_X \times \Omega_Y$. The reference measure λ can be decomposed as a product measure $\lambda = \lambda_X \times \lambda_Y$ which implies again the *Hilbert space structure* of the Bayes space $\mathcal{B}^2(\lambda)$

$\Rightarrow$ *Perturbation* and *powering* operations: $f, g \in \mathcal{B}^2(\lambda)$, $\alpha \in \mathbb{R}$

$$(f \oplus g)(x,y) =_{\mathcal{B}^2(\lambda)} f(x,y) \cdot g(x,y), \; (\alpha \odot f)(x,y) =_{\mathcal{B}^2(\lambda)} f(x,y)^\alpha$$

- the *Bayes inner product*:

$$\langle f, g \rangle_{\mathcal{B}^2(\lambda)} = \frac{1}{2\lambda(\Omega)} \iint_\Omega \iint_\Omega \ln \frac{f(x,y)}{f(s,t)} \ln \frac{g(x,y)}{g(s,t)} \, d\lambda(x,y) \, d\lambda(s,t)$$

$\Rightarrow$ $||f||_{\mathcal{B}^2(\lambda)} = \sqrt{\langle f, f \rangle_{\mathcal{B}^2(\lambda)}}, \quad d_{\mathcal{B}^2(\lambda)}(f,g) = ||f \ominus g||_{\mathcal{B}^2(\lambda)}$,
where $f \ominus g = f \oplus [(-1) \odot g]$

Bayes spaces: centred logratio transformation

- Also the *centred logratio (clr) transformation* is easily extendable to the bivariate case

Bayes spaces: centred logratio transformation

- Also the *centred logratio (clr) transformation* is easily extendable to the bivariate case
- The clr transformation of a bivariate density $f(x,y) \in \mathcal{B}^2(\lambda)$ is a real function $f^c : \Omega \to \mathbb{R}$, $f^c \in L_0^2(\lambda)$:

$$\begin{aligned} f^c(x,y) &= \ln f(x,y) - \frac{1}{\lambda(\Omega)} \iint_{\Omega} \ln f(x,y)\, d\lambda = \\ &= \ln f(x,y) - \frac{1}{\lambda(\Omega)} \int_{\Omega_X} \int_{\Omega_Y} \ln f(x,y)\, d\lambda_X d\lambda_Y \end{aligned}$$

- The *zero integral constraint* is imposed on $f^c(x,y)$ ... $L_0^2(\lambda)$

Orthogonal decomposition of bivariate PDFs

- One of the key goals in probability theory is to study dependence structure between two random variables

Orthogonal decomposition of bivariate PDFs

- One of the key goals in probability theory is to study dependence structure between two random variables
- The well-known **Sklar's theorem** in theory of *copulas* provides a decomposition of any PDF into its interactive and independent parts, the latter being built as the product of the respective marginal PDFs

Orthogonal decomposition of bivariate PDFs

- One of the key goals in probability theory is to study dependence structure between two random variables
- The well-known **Sklar's theorem** in theory of *copulas* provides a decomposition of any PDF into its interactive and independent parts, the latter being built as the product of the respective marginal PDFs
- Relying on the Bayes space methodology allows one to provide a similar decomposition which is now, however, *orthogonal*

Orthogonal decomposition of bivariate PDFs

- One of the key goals in probability theory is to study dependence structure between two random variables
- The well-known **Sklar's theorem** in theory of *copulas* provides a decomposition of any PDF into its interactive and independent parts, the latter being built as the product of the respective marginal PDFs
- Relying on the Bayes space methodology allows one to provide a similar decomposition which is now, however, *orthogonal*
- This enables for an elegant *geometrical representation of the decomposition*, and for a powerful *probabilistic interpretation* with direct *consequences from the statistical viewpoint*

Orthogonal decomposition of bivariate PDFs

⇒ A novel definition of marginals, named *geometric marginals*, built upon marginalizing the bivariate clr transformation:

$$\begin{aligned} f^c_X(x) &= \frac{1}{\lambda_Y(\Omega_Y)} \int_{\Omega_Y} f^c(x,y)\, d\lambda_Y, \\ f^c_Y(y) &= \frac{1}{\lambda_X(\Omega_X)} \int_{\Omega_X} f^c(x,y)\, d\lambda_X; \end{aligned}$$

$f^c_X \in L^2_0(\lambda_X)$ and $f^c_Y \in L^2_0(\lambda_Y)$ (*clr-marginals*)

Orthogonal decomposition of bivariate PDFs

- *Geometric marginals* $f_X \in \mathcal{B}^2(\lambda_X)$ and $f_Y \in \mathcal{B}^2(\lambda_Y)$ are elements of $\mathcal{B}^2(\lambda_X)$ and $\mathcal{B}^2(\lambda_Y)$ associated with the clr-marginals:

$$
\begin{aligned}
f_X(x) &=_{\mathcal{B}^2(\lambda_X)} \exp\left\{\frac{1}{\lambda_Y(\Omega_Y)}\int_{\Omega_Y} \ln f(x,y)\, d\lambda_Y\right\},\\
f_Y(y) &=_{\mathcal{B}^2(\lambda_Y)} \exp\left\{\frac{1}{\lambda_X(\Omega_X)}\int_{\Omega_X} \ln f(x,y)\, d\lambda_X\right\}
\end{aligned}
$$

- The normalizing terms guarantee that the scale of λ (λ_X, λ_Y) does not matter for the construction of geometric marginals

Orthogonal decomposition of bivariate PDFs

- *Geometric marginals* $f_X \in \mathcal{B}^2(\lambda_X)$ and $f_Y \in \mathcal{B}^2(\lambda_Y)$ are elements of $\mathcal{B}^2(\lambda_X)$ and $\mathcal{B}^2(\lambda_Y)$ associated with the clr-marginals:

$$\begin{aligned} f_X(x) &=_{\mathcal{B}^2(\lambda_X)} \exp\left\{\frac{1}{\lambda_Y(\Omega_Y)}\int_{\Omega_Y} \ln f(x,y)\, d\lambda_Y\right\}, \\ f_Y(y) &=_{\mathcal{B}^2(\lambda_Y)} \exp\left\{\frac{1}{\lambda_X(\Omega_X)}\int_{\Omega_X} \ln f(x,y)\, d\lambda_X\right\} \end{aligned}$$

- The normalizing terms guarantee that the scale of λ (λ_X, λ_Y) does not matter for the construction of geometric marginals
- Geometric marginals are **orthogonal projections** of the bivariate density $f(x,y)$ (Genest et al., 2023)

Orthogonal decomposition of bivariate PDFs

- In probability theory, independence of random variables corresponds to the possibility of expressing a joint density as a product of its *marginals* – here *geometric marginals*.

Orthogonal decomposition of bivariate PDFs

- In probability theory, independence of random variables corresponds to the possibility of expressing a joint density as a product of its *marginals* – here *geometric marginals*.
- **Independence PDF**: $f_{\text{ind}}(x,y) = f_X(x)f_Y(y),\ (x,y) \in \Omega$

Orthogonal decomposition of bivariate PDFs

- In probability theory, independence of random variables corresponds to the possibility of expressing a joint density as a product of its *marginals* – here *geometric marginals*.
- **Independence PDF**: $f_{\mathrm{ind}}(x,y) = f_X(x)f_Y(y),\ (x,y) \in \Omega$
- **Interaction PDF**: $f_{\mathrm{int}}(x,y) = \frac{f(x,y)}{f_X(x)f_Y(y)} = f(x,y) \ominus f_{\mathrm{ind}}(x,y)$,

Orthogonal decomposition of bivariate PDFs

- In probability theory, independence of random variables corresponds to the possibility of expressing a joint density as a product of its *marginals* – here *geometric marginals*.
- **Independence PDF**: $f_{\text{ind}}(x, y) = f_X(x) f_Y(y), \ (x, y) \in \Omega$
- **Interaction PDF**: $f_{\text{int}}(x, y) = \frac{f(x,y)}{f_X(x) f_Y(y)} = f(x, y) \ominus f_{\text{ind}}(x, y)$,

$\Rightarrow$ The original bivariate PDF can be decomposed as

$$f(x, y) = f_{\text{ind}}(x, y) \oplus f_{\text{int}}(x, y)$$

Example: a truncated Gaussian density

- We consider a zero-mean bivariate Gaussian density $\mathcal{N}_2(\mu, \Sigma)$ with respect to the (product) Lebesgue measure $\lambda[I] = \lambda[I_1] \times \lambda[I_2]$, truncated on a rectangular domain $I = I_1 \times I_2 \subset \mathbb{R}^2$, with $I_1 = I_2 = [-T, T]$, $T = 5$

Example: a truncated Gaussian density

- We consider a zero-mean bivariate Gaussian density $\mathcal{N}_2(\mu, \Sigma)$ with respect to the (product) Lebesgue measure $\lambda[I] = \lambda[I_1] \times \lambda[I_2]$, truncated on a rectangular domain $I = I_1 \times I_2 \subset \mathbb{R}^2$, with $I_1 = I_2 = [-T, T]$, $T = 5$
- The PDF is defined, for $\mathbf{x} = (x, y) \in I$, as

$$\begin{aligned} f(x,y) &=_{\mathcal{B}^2(\lambda)} \exp\left\{\mathbf{x}^\top \Sigma^{-1} \mathbf{x}\right\} \\ &= \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{x^2}{\sigma_1^2} - 2\rho\frac{xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right]\right\}, \end{aligned}$$

with $\sigma_i^2 = \Sigma_{ii}$ and $\rho \in [0, 1]$ being the correlation coefficient

- Clr transformation of f:

$$f^c(x,y) = -\frac{1}{2(1-\rho^2)}\left[\frac{x^2}{\sigma_1^2} - 2\rho\frac{xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right] + \frac{T^2}{6(1-\rho^2)}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)$$

Example: a truncated Gaussian density

- Independence and interaction (clr-)PDFs in case of Lebesgue reference measure:

$$
\begin{aligned}
f^c_{\mathrm{ind}}(x,y) &= -\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2}+\frac{y^2}{\sigma_2^2}\right)+\frac{T^2}{6(1-\rho^2)}\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right);\\
f^c_{\mathrm{int}}(x,y) &= \frac{1}{(1-\rho^2)}\left(\rho\frac{xy}{\sigma_1\sigma_2}\right).
\end{aligned}
$$

Example: a truncated Gaussian density

- Independence and interaction (clr-)PDFs in case of Lebesgue reference measure:

$$
\begin{aligned}
f^{c}_{\mathrm{ind}}(x,y) &= -\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2}+\frac{y^2}{\sigma_2^2}\right)+\frac{T^2}{6(1-\rho^2)}\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right);\\
f^{c}_{\mathrm{int}}(x,y) &= \frac{1}{(1-\rho^2)}\left(\rho\frac{xy}{\sigma_1\sigma_2}\right).
\end{aligned}
$$

- The interaction PDF precisely captures the terms in f^c depending on the mixed polynomial xy (i.e., the interaction between x and y), its magnitude is controlled by the magnitude of ρ.

Example: a truncated Gaussian density

- Independence and interaction (clr-)PDFs in case of Lebesgue reference measure:

$$
\begin{aligned}
f^c_{\mathrm{ind}}(x,y) &= -\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2}+\frac{y^2}{\sigma_2^2}\right)+\frac{T^2}{6(1-\rho^2)}\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right);\\
f^c_{\mathrm{int}}(x,y) &= \frac{1}{(1-\rho^2)}\left(\rho\frac{xy}{\sigma_1\sigma_2}\right).
\end{aligned}
$$

- The interaction PDF precisely captures the terms in f^c depending on the mixed polynomial xy (i.e., the interaction between x and y), its magnitude is controlled by the magnitude of ρ.
- In case of independence ($\rho = 0$), f^c_{int} is null, and $f_{\mathrm{int}} = 0_\oplus$.

Example: a truncated Gaussian density

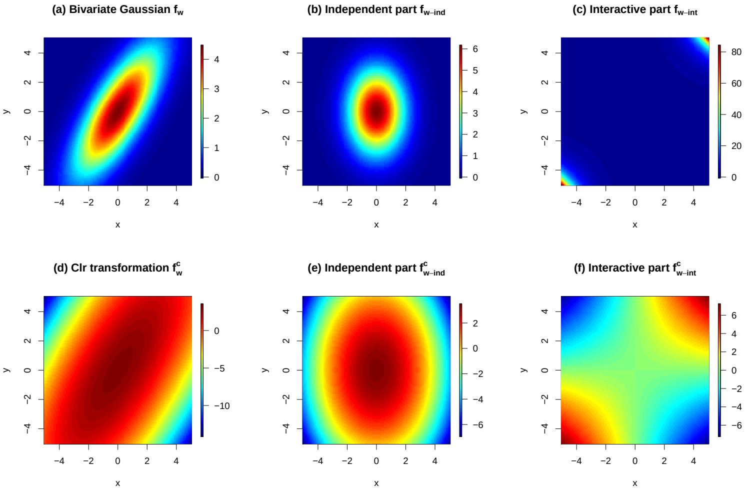
- Independence and interaction PDFs in case of Lebesgue reference measure:

$$f_X(x) =_{\mathcal{B}^2(\lambda_X)} \exp\left[-\frac{1}{2(1-\rho^2)}\frac{x^2}{\sigma_1^2}\right], \text{ (likewise } f_Y(y));$$

$$f_{\text{int}}(x,y) =_{\mathcal{B}^2(\lambda)} \exp\left[\frac{1}{(1-\rho^2)}\left(\rho\frac{xy}{\sigma_1\sigma_2}\right)\right].$$

- The interaction PDF precisely captures the terms in f depending on the mixed polynomial xy (i.e., the interaction between x and y), its magnitude is controlled by the magnitude of ρ.
- In case of independence ($\rho = 0$), f_{int}^c is null, and $f_{\text{int}} = 0_\oplus$.

Gaussian example



Conditional probabilities as space quotient (KGvdB)

- Conditional distributions differ by a marginal:

$$f_X(\cdot|y) = g_X(\cdot|y) \Leftrightarrow f(x,y) \ominus g(x,y) \in \mathcal{B}^2(\lambda_Y)$$

which is the definition of equivalence in quotient spaces, i.e.

$$f_X(\cdot|y) \in \frac{\mathcal{B}^2(\lambda_X \times \lambda_Y)}{\mathcal{B}^2(\lambda_Y)}$$

(*spaces of regression models*)

- **Remark**: The difference is the difference of marginal distributions:

$$f(x,y) \ominus g(x,y) =_{\mathcal{B}^2(\lambda_X \times \lambda_Y)} f_Y(y) \ominus g_Y(y)$$

Conditional subspaces (KGvdB)

- The other way around:

$$f_X(\cdot|y) = \frac{f_X(x) \oplus f_Y(y) \oplus f_{\mathrm{int}}(x, y)}{f_Y(y)},$$

i.e., the conditional distribution just misses the marginal part

Conditional subspaces (KGvdB)

- The other way around:

$$f_X(\cdot|y) = \frac{f_X(x) \oplus f_Y(y) \oplus f_{\mathrm{int}}(x,y)}{f_Y(y)},$$

  i.e., the conditional distribution just misses the marginal part

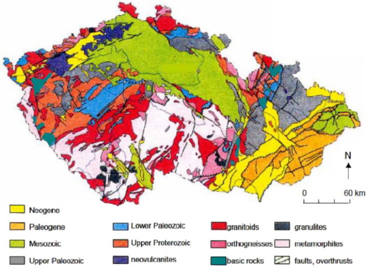
- Conditional distributions are represented by the projection on the orthogonal complement of the conditioning space: $\mathcal{B}^2(\lambda_Y)^{\perp}$

- Any joint structure exists with every combination of marginals

Motivation example: PDFs of chemical elements

- Trivariate densities of (log-transformed) copper (Cu), lead (Pb) and zinc (Zn) soil concentration data in 77 districts of the Czech Republic from the Register of Contaminated Areas (registr kontaminovaných ploch) collected by the Department of Agriculture of the Czech Republic (eAGRI, 2020)
- At least higher hundreds of concentration values available in each district
- Districts are characterized by diverse geological origin and anthropogenic contamination

Motivation example: PDFs of chemical elements

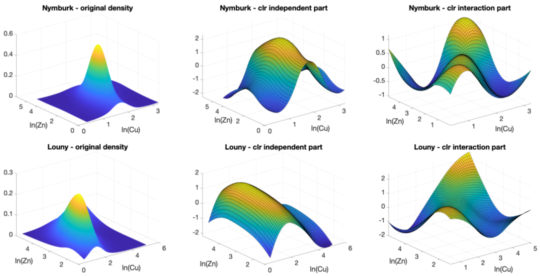


Source: Czech Geological Survey, Prague

Bivariate Cu-Zn densities are depicted:

- the first case corresponds to Nymburk County, where soil contamination is likely low
- the second case corresponds to Louny County, where heavy soil contamination is suspected due to abundant hops culture as well as ore mining and processing

Motivation example: PDFs of chemical elements



- What information can be obtained from trivariate densities?
- How can we capture interactions between elements?

Notation

- Let $(\Omega, \mathcal{A})$ be now a d-dimensional product space and let λ be a finite product reference measure. Specifically, suppose that for $i \in D$, $(\Omega_i, \mathcal{A}_i)$ is a measurable space and λ_i is a finite, positive, real-valued measure on it. Set

$$\Omega = \Omega_1 \times \cdots \times \Omega_d, \quad \mathcal{A} = \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_d, \quad \lambda = \lambda_1 \otimes \cdots \otimes \lambda_d.$$

- For arbitrary non-empty $I \subseteq D$, let

$$\Omega_I = \times_{i \in I} \Omega_i, \quad \mathcal{A}_I = \bigotimes_{i \in I} \mathcal{A}_i, \quad \lambda_I = \bigotimes_{i \in I} \lambda_i.$$

Notation

- Let $(\Omega, \mathcal{A})$ be now a d-dimensional product space and let λ be a finite product reference measure. Specifically, suppose that for $i \in D$, $(\Omega_i, \mathcal{A}_i)$ is a measurable space and λ_i is a finite, positive, real-valued measure on it. Set

$$\Omega = \Omega_1 \times \cdots \times \Omega_d, \quad \mathcal{A} = \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_d, \quad \lambda = \lambda_1 \otimes \cdots \otimes \lambda_d.$$

- For arbitrary non-empty $I \subseteq D$, let

$$\Omega_I = \times_{i \in I} \Omega_i, \quad \mathcal{A}_I = \bigotimes_{i \in I} \mathcal{A}_i, \quad \lambda_I = \bigotimes_{i \in I} \lambda_i.$$

♡ Necessity of having the product space is the only restriction we need to take into account.

Notation

- For arbitrary $I \subseteq D = \{1, \dots, d\}$, $I \neq \varnothing$, let $\mathcal{B}^2_I(\lambda)$ be the subspace of $\mathcal{B}^2(\lambda)$ given by

$$\mathcal{B}^2_I(\lambda) = \left\{ \mu \in \mathcal{B}^2(\lambda) \,\middle|\, \exists_{\mu_I \in \mathcal{B}^2(\lambda_I)} \forall_{A_i \in \mathcal{A}_i}\ \mu\left(\times_{i \in D} A_i\right) = \prod_{i \notin I} \lambda_i(A_i) \mu_I\left(\times_{i \in I} A_i\right) \right\}$$

$\Rightarrow$ Any λ-density in $\mathcal{B}^2_I(\lambda)$ is a function of arguments with indices in I only

Notation

- For arbitrary $I \subseteq D = \{1, \dots, d\}$, $I \neq \emptyset$, let $\mathcal{B}^2_I(\lambda)$ be the subspace of $\mathcal{B}^2(\lambda)$ given by

$$\mathcal{B}^2_I(\lambda) = \left\{ \mu \in \mathcal{B}^2(\lambda) \,\middle|\, \exists_{\mu_I \in \mathcal{B}^2(\lambda_I)} \forall_{A_i \in \mathcal{A}_i} \; \mu\left(\times_{i \in D} A_i\right) = \prod_{i \notin I} \lambda_i(A_i) \mu_I\left(\times_{i \in I} A_i\right) \right\}$$

$\Rightarrow$ Any λ-density in $\mathcal{B}^2_I(\lambda)$ is a function of arguments with indices in I only

- Note that when $D = I$, $\mathcal{B}^2_D(\lambda) = \mathcal{B}^2(\lambda)$
- By convention, $\mathcal{B}^2_\emptyset(\lambda)$ is the trivial subspace that contains only the neutral element λ

Geometric marginals

- For $I \subsetneq D$, the *I-th geometric marginal* is

$$f_I = \exp\left\{ \frac{1}{\lambda_{D\setminus I}(\Omega_{D\setminus I})} \int_{\Omega_{D\setminus I}} \ln(f) d\lambda_{D\setminus I} \right\}.$$

and its clr transformation

$$\mathrm{clr}(f_I) = \frac{1}{\lambda_{D\setminus I}(\Omega_{D\setminus I})} \int_{\Omega_{D\setminus I}} \mathrm{clr}(f) d\lambda_{D\setminus I}$$

Geometric marginals

- For $I \subsetneq D$, the *I-th geometric marginal* is

$$f_I = \exp\left\{ \frac{1}{\lambda_{D\setminus I}(\Omega_{D\setminus I})} \int_{\Omega_{D\setminus I}} \ln(f) d\lambda_{D\setminus I} \right\}.$$

and its clr transformation

$$\mathrm{clr}(f_I) = \frac{1}{\lambda_{D\setminus I}(\Omega_{D\setminus I})} \int_{\Omega_{D\setminus I}} \mathrm{clr}(f) d\lambda_{D\setminus I}$$

♡ It is the unique orthogonal projection of $f (\equiv f_\mu)$ onto $\mathcal{B}_I^2(\lambda)$

- The logarithm of f in f_I can be replaced by its clr transformation due to properties of the exponential (and proportionality of densities in the Bayes space framework)

Decomposition of f

Kuo et al. (2010) reformulated the Hoeffding-Sobol decomposition in terms of projections which was used to derive the following:

Theorem (Decomposition of multivariate densities)

For any $f \in \mathcal{B}^2(\lambda)$,

$$f = f_{\mathrm{ind}} \oplus \bigoplus_{I \subseteq D, |I| \geq 2} f_{I,\mathrm{int}}$$

where the so-called independence and interaction parts are respectively given by

$$f_{\mathrm{ind}} = \bigoplus_{i=1}^{d} f_i, \quad f_{I,\mathrm{int}} = \bigoplus_{J \subseteq I, J \neq \varnothing} \{(-1)^{|I \setminus J|}\} \odot f_J.$$

Decomposition of f $(d = 3)$

- $f_{\mathrm{ind}} = f_1 \oplus f_2 \oplus f_3$

Decomposition of f $(d = 3)$

- $f_{\text{ind}} = f_1 \oplus f_2 \oplus f_3$
- $f_{\{1,2\},\text{int}} = f_{\{1,2\}} \ominus f_1 \ominus f_2$
- $f_{\{1,3\},\text{int}} = f_{\{1,3\}} \ominus f_1 \ominus f_3$
- $f_{\{2,3\},\text{int}} = f_{\{2,3\}} \ominus f_2 \ominus f_3$

Decomposition of f $(d = 3)$

- $f_{\mathrm{ind}} = f_1 \oplus f_2 \oplus f_3$
- $f_{\{1,2\},\mathrm{int}} = f_{\{1,2\}} \ominus f_1 \ominus f_2$
- $f_{\{1,3\},\mathrm{int}} = f_{\{1,3\}} \ominus f_1 \ominus f_3$
- $f_{\{2,3\},\mathrm{int}} = f_{\{2,3\}} \ominus f_2 \ominus f_3$
- $f_{\{1,2,3\},\mathrm{int}} = f \ominus f_{\{1,2\}} \ominus f_{\{1,3\}} \ominus f_{\{2,3\}} \oplus f_1 \oplus f_2 \oplus f_3$

Decomposition of f $(d = 3)$

- $f_{\mathrm{ind}} = f_1 \oplus f_2 \oplus f_3$
- $f_{\{1,2\},\mathrm{int}} = f_{\{1,2\}} \ominus f_1 \ominus f_2$
- $f_{\{1,3\},\mathrm{int}} = f_{\{1,3\}} \ominus f_1 \ominus f_3$
- $f_{\{2,3\},\mathrm{int}} = f_{\{2,3\}} \ominus f_2 \ominus f_3$
- $f_{\{1,2,3\},\mathrm{int}} = f \ominus f_{\{1,2\}} \ominus f_{\{1,3\}} \ominus f_{\{2,3\}} \oplus f_1 \oplus f_2 \oplus f_3$

$\Rightarrow$ Explicit formula for the decomposition using directly marginal densities

Decomposition of f

Also recursive formula (cf. Fačevicová et al., 2022) possible:

Theorem (Decomposition of multivariate densities)

For any $f \in \mathcal{B}^2(\lambda)$,

$$f = f_{\mathrm{ind}} \oplus \bigoplus_{I \subseteq D, |I| \geq 2} f_{I,\mathrm{int}}$$

where the so-called independence and interaction parts are respectively given by

$$f_{\mathrm{ind}} = \bigoplus_{i=1}^{d} f_i, \quad f_{I,\mathrm{int}} = f_I \ominus \left[\bigoplus_{J \subsetneq I, |J| \geq 2} f_{J,\mathrm{int}} \bigoplus_{i \in I} f_i \right].$$

Decomposition of f $(d = 3)$

- $f_{\mathrm{ind}} = f_1 \oplus f_2 \oplus f_3$

- $f_{\{1,2\},\mathrm{int}} = f_{\{1,2\}} \ominus f_1 \ominus f_2$
- $f_{\{1,3\},\mathrm{int}} = f_{\{1,3\}} \ominus f_1 \ominus f_3$
- $f_{\{2,3\},\mathrm{int}} = f_{\{2,3\}} \ominus f_2 \ominus f_3$

- $f_{\{1,2,3\},\mathrm{int}} = f \ominus f_{\{1,2\}} \ominus f_{\{1,3\}} \ominus f_{\{2,3\}} \oplus f_1 \oplus f_2 \oplus f_3$

Decomposition of f $(d = 3)$

- $f_{\text{ind}} = f_1 \oplus f_2 \oplus f_3$
- $f_{\{1,2\},\text{int}} = f_{\{1,2\}} \ominus f_1 \ominus f_2$
- $f_{\{1,3\},\text{int}} = f_{\{1,3\}} \ominus f_1 \ominus f_3$
- $f_{\{2,3\},\text{int}} = f_{\{2,3\}} \ominus f_2 \ominus f_3$
- $f_{\{1,2,3\},\text{int}} = f \ominus f_{\{1,2\}} \ominus f_{\{1,3\}} \ominus f_{\{2,3\}} \oplus f_1 \oplus f_2 \oplus f_3$

$\rightarrow$ Recursive formula:

- $f_{\{1,2\},\text{int}}$, $f_{\{1,3\},\text{int}}$, $f_{\{2,3\},\text{int}}$: no change from before
- $f_{\{1,2,3\},\text{int}} = f \ominus f_1 \ominus f_2 \ominus f_3 \ominus f_{\{1,2\},\text{int}} \ominus f_{\{1,3\},\text{int}} \ominus f_{\{2,3\},\text{int}}$

Decomposition of f $(d = 3)$

- $f_{\mathrm{ind}} = f_1 \oplus f_2 \oplus f_3$

- $f_{\{1,2\},\mathrm{int}} = f_{\{1,2\}} \ominus f_1 \ominus f_2$
- $f_{\{1,3\},\mathrm{int}} = f_{\{1,3\}} \ominus f_1 \ominus f_3$
- $f_{\{2,3\},\mathrm{int}} = f_{\{2,3\}} \ominus f_2 \ominus f_3$

- $f_{\{1,2,3\},\mathrm{int}} = f \ominus f_{\{1,2\}} \ominus f_{\{1,3\}} \ominus f_{\{2,3\}} \oplus f_1 \oplus f_2 \oplus f_3$

$\rightarrow$ Recursive formula:

- $f_{\{1,2\},\mathrm{int}}$, $f_{\{1,3\},\mathrm{int}}$, $f_{\{2,3\},\mathrm{int}}$: no change from before
- $f_{\{1,2,3\},\mathrm{int}} = f \ominus f_1 \ominus f_2 \ominus f_3 \ominus f_{\{1,2\},\mathrm{int}} \ominus f_{\{1,3\},\mathrm{int}} \ominus f_{\{2,3\},\mathrm{int}}$

- The recursive formula is important for further derivations

Orthogonality of decomposition

- Some more notation for $I \subseteq D$ is needed:

Orthogonality of decomposition

- Some more notation for $I \subseteq D$ is needed:
- $\mathcal{B}^2_{I,\mathrm{int}}(\lambda)$ is defined as the orthogonal complement of the direct sum $\mathcal{B}^2_{I,\oplus}(\lambda) = \bigoplus_{i\in I} \mathcal{B}^2_{\{i\}}(\lambda) \bigoplus_{J\subsetneq I, |J|\geq 2} \mathcal{B}^2_{J,\mathrm{int}}(\lambda)$ in $\mathcal{B}^2_I(\lambda)$

$\Rightarrow$ $\mathcal{B}^2_{I,\mathrm{int}}(\lambda) = \{f \in \mathcal{B}^2_I(\lambda) \mid \forall_{g\in\mathcal{B}^2_{I,\oplus}(\lambda)} \langle f, g\rangle_{\mathcal{B}^2(\lambda)} = 0\}$

- Independence space:

$$\mathcal{B}^2_{\mathrm{ind}}(\lambda) = \{\mu \in \mathcal{B}^2(\lambda) \mid \exists_{\mu_1\in\mathcal{B}^2(\lambda_1)} \cdots \exists_{\mu_d\in\mathcal{B}^2(\lambda_d)} \mu = \mu_1 \otimes \cdots \otimes \mu_d\}$$

- These subspaces span the corresponding densities from the decomposition of f

Orthogonality of decomposition

Theorem (Orthogonality of decomposition)

For any $f \in \mathcal{B}^2(\lambda)$ and any $I \subseteq D$ with $|I| \geq 2$, $f_{I,\mathrm{int}}$ is the unique orthogonal projection of f on $\mathcal{B}^2_{I,\mathrm{int}}(\lambda)$. Furthermore,

$$\mathcal{B}^2(\lambda) = \bigoplus_{i \in D} \mathcal{B}^2_{\{i\}}(\lambda) \bigoplus_{I \subseteq D, |I| \geq 2} \mathcal{B}^2_{I,\mathrm{int}}(\lambda)$$

and all spaces featuring in the direct sum are mutually orthogonal.

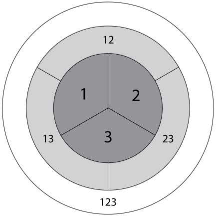
$\Rightarrow$ The independence space can be expressed as

$$\mathcal{B}^2_{\mathrm{ind}}(\lambda) = \mathcal{B}^2_{\{1\}}(\lambda) \oplus \cdots \oplus \mathcal{B}^2_{\{d\}}(\lambda)$$

and for any $f \in \mathcal{B}^2(\lambda)$, the independence part f_{ind} is the unique orthogonal projection of f on $\mathcal{B}^2_{\mathrm{ind}}(\lambda)$.

Orthogonality of decomposition ($d = 3$)

- The situation for $d = 3$ can be depicted as follows:



The sectors marked 1, 2, 3 correspond to the subspaces $\mathcal{B}^2_{\{1\}}(\lambda)$, $\mathcal{B}^2_{\{2\}}(\lambda)$, $\mathcal{B}^2_{\{3\}}(\lambda)$. In light grey the three orthogonal subspaces $\mathcal{B}^2_{I,\mathrm{int}}(\lambda)$ corresponding to the subsets $I = \{1,2\}$, $\{1,3\}$, and $\{2,3\}$ are depicted. The white annulus labeled 123 is $\mathcal{B}^2_{\{1,2,3\},\mathrm{int}}(\lambda)$, i.e., the orthogonal complement of the direct sum $\mathcal{B}^2_{\{1,2,3\},\oplus}(\lambda)$ of the grey zones.

Orthogonality of decomposition ($d = 3$)

- Summing up, one has

$$\mathcal{B}^2(\lambda) = \mathcal{B}^2_{\{1\}}(\lambda) \oplus \mathcal{B}^2_{\{2\}}(\lambda) \oplus \mathcal{B}^2_{\{3\}}(\lambda)$$
$$\oplus \mathcal{B}^2_{\{1,2\},\mathrm{int}}(\lambda) \oplus \mathcal{B}^2_{\{1,3\},\mathrm{int}}(\lambda) \oplus \mathcal{B}^2_{\{2,3\},\mathrm{int}}(\lambda) \oplus \mathcal{B}^2_{\{1,2,3\},\mathrm{int}}(\lambda),$$

$\Rightarrow$ The orthogonal decomposition of any $f \in \mathcal{B}^2(\lambda)$ takes the form

$$f = f_1 \oplus f_2 \oplus f_3 \oplus f_{\{1,2\},\mathrm{int}} \oplus f_{\{1,3\},\mathrm{int}} \oplus f_{\{2,3\},\mathrm{int}} \oplus f_{\{1,2,3\},\mathrm{int}}$$

Consequences of the decomposition (1)

(i) *Pythagoras' Theorem:* One always has

$$\|f_\mu\|^2_{\mathcal{B}^2(\lambda)} = \|f_{\mu,\mathrm{ind}}\|^2_{\mathcal{B}^2(\lambda)} + \sum_{I \subseteq D, |I| \geq 2} \|f_{\mu,I,\mathrm{int}}\|^2_{\mathcal{B}^2(\lambda)},$$

where $\|f_{\mu,\mathrm{ind}}\|^2_{\mathcal{B}^2(\lambda)} = \|f_{\mu,1}\|^2_{\mathcal{B}^2(\lambda)} + \cdots + \|f_{\mu,d}\|^2_{\mathcal{B}^2(\lambda)}$.
[Bayes norm = scalar measure of information (Egozcue and Pawlowsky-Glahn, 2018)]

Consequences of the decomposition (1)

(i) *Pythagoras' Theorem:* One always has

$$\|f_\mu\|^2_{\mathcal{B}^2(\lambda)} = \|f_{\mu,\mathrm{ind}}\|^2_{\mathcal{B}^2(\lambda)} + \sum_{I\subseteq D, |I|\geq 2} \|f_{\mu,I,\mathrm{int}}\|^2_{\mathcal{B}^2(\lambda)},$$

where $\|f_{\mu,\mathrm{ind}}\|^2_{\mathcal{B}^2(\lambda)} = \|f_{\mu,1}\|^2_{\mathcal{B}^2(\lambda)} + \cdots + \|f_{\mu,d}\|^2_{\mathcal{B}^2(\lambda)}$.
[Bayes norm = scalar measure of information (Egozcue and Pawlowsky-Glahn, 2018)]

(ii) *Margin-free property:* For any given set $I \subseteq D$ with $|I| \geq 2$, let $f_\nu = f_{\mu,I,\mathrm{int}}$. Then for each $i \in D$, the i-th univariate geometric margin $f_{\nu,i}$ of f_ν satisfies $f_{\nu,i} \propto 1$. Moreover, for any set $J \subsetneq I$ with $|J| \geq 2$, the J-th geometric margin of f_ν satisfies $f_{\nu,J} \propto 1$.

Consequences of the decomposition (2)

(iii) *Independence:* If $f_\mu = f_{\mu_1} \cdots f_{\mu_d} \in \mathcal{B}^2_{\text{ind}}(\lambda)$, then for each $i \in D$, $f_{\mu,i} \propto f_{\mu_i}$ (arith. and geom. margins coincide) and

$$f_\mu \in \mathcal{B}^2_{\text{ind}}(\lambda) \Leftrightarrow f_\mu \propto f_{\mu,1} \oplus \cdots \oplus f_{\mu,d} \Leftrightarrow \forall_{I \subseteq D, |I| \geq 2}\, f_{\mu,I,\text{int}} \propto 1.$$

Consequences of the decomposition (2)

(iii) *Independence:* If $f_\mu = f_{\mu_1} \cdots f_{\mu_d} \in \mathcal{B}^2_{\text{ind}}(\lambda)$, then for each $i \in D$, $f_{\mu,i} \propto f_{\mu_i}$ (arith. and geom. margins coincide) and

$$f_\mu \in \mathcal{B}^2_{\text{ind}}(\lambda) \Leftrightarrow f_\mu \propto f_{\mu,1} \oplus \cdots \oplus f_{\mu,d} \Leftrightarrow \forall_{I \subseteq D, |I| \geq 2} \, f_{\mu,I,\text{int}} \propto 1.$$

(iv) *Yule perturbation:* For $g \propto g_1 \cdots g_d \in \mathcal{B}^2_{\text{ind}}(\lambda)$, set $h = f_\mu \oplus g$. Then the independence and interaction parts of h satisfy

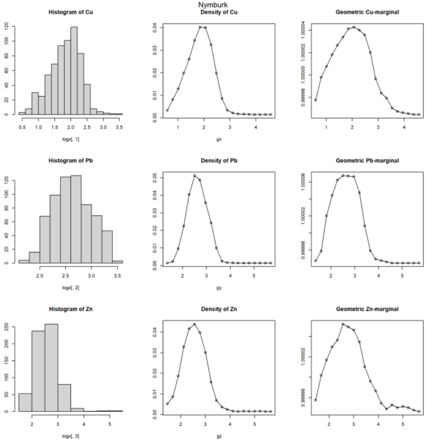
$$h_{\text{ind}} = f_{\mu,\text{ind}} \oplus g, \quad h_{\text{int}} = f_{\mu,\text{int}},$$

where

$$f_{\mu,\text{int}} = \bigoplus_{I \subseteq D, |I| \geq 2} f_{\mu,I,\text{int}}.$$

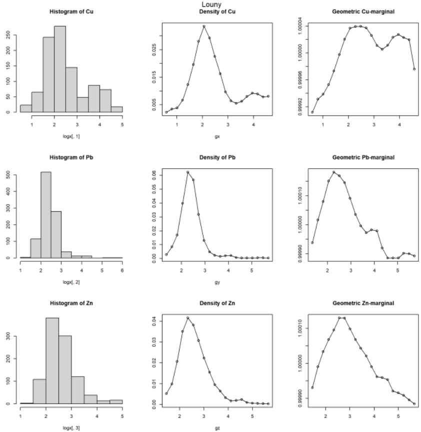
In particular, the geometric margins of h satisfy $h_i \propto f_{\mu,i} g_i$ for all $i \in D$.

PDFs of chemical elements: univariate analysis



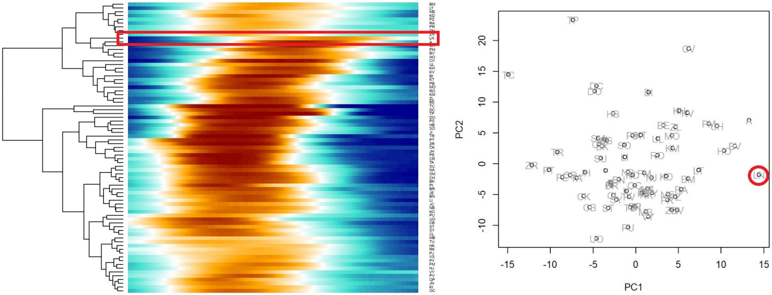
Geometric marginals capture also interelemental relationships!

PDFs of chemical elements: univariate analysis



Geometric marginals capture also interelemental relationships!

PDFs of chemical elements: “univariate” analysis

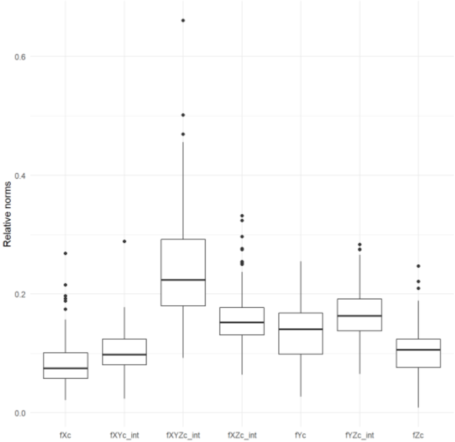


Cu geometric marginals, clr space

left: clustering, right: scores of functional principal component analysis

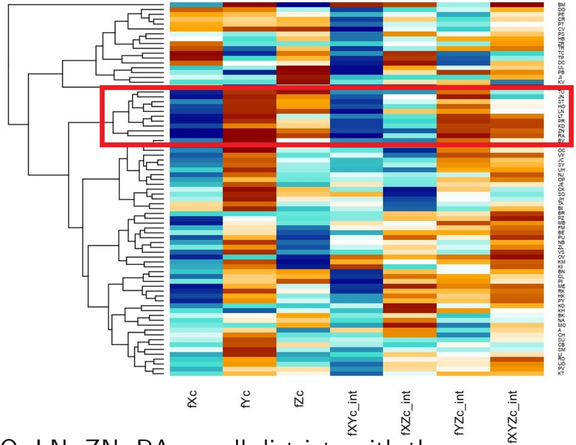
Louny . . . LN, Nymburk . . . NB

PDFs of chemical elements: decomposition



X . . . Cu, Y . . . Pb, Z . . . Zn

PDFs of chemical elements: decomposition (clr)



LN, LT, HO, LN, ZN, RA are all districts with the same source of contamination (agricultural spraying due to intensive land use)

Exploratory FDA of multivariate PDFs: R code

https://github.com/uradojic/EFDA-of-multivariate-densities.git

(with special thanks to Una Radojičić and Ivana Pavlů)

Outlook

- A connection between the presented approach and the copula methodology, and also its application for decomposition of copulas themselves were outlined in Genest et al. (2023)

Outlook

- A connection between the presented approach and the copula methodology, and also its application for decomposition of copulas themselves were outlined in Genest et al. (2023)
- Functional data analysis of multivariate densities (also related to possible sparsity of the decomposition) is currently under development, first concise study in Matys Grygar et al. (2024)

Outlook

- A connection between the presented approach and the copula methodology, and also its application for decomposition of copulas themselves were outlined in Genest et al. (2023)
- Functional data analysis of multivariate densities (also related to possible sparsity of the decomposition) is currently under development, first concise study in Matys Grygar et al. (2024)
- Related to this, also the spline representation in Bayes spaces (Machalová et al., 2021; Hron et al., 2022) can be extended to the multivariate case

Outlook

- A connection between the presented approach and the copula methodology, and also its application for decomposition of copulas themselves were outlined in Genest et al. (2023)
- Functional data analysis of multivariate densities (also related to possible sparsity of the decomposition) is currently under development, first concise study in Matys Grygar et al. (2024)
- Related to this, also the spline representation in Bayes spaces (Machalová et al., 2021; Hron et al., 2022) can be extended to the multivariate case
- A great potential of multivariate Bayes spaces further in Bayesian inference (Barfoot and D'Eleuterio, 2023; Wynne, 2023), graphical models, conditional distributions, generalized regression. . .

References

Aitchison, J.: *The statistical analysis of compositional data*. Chapman and Hall, London, 1986.

Barfoot, T.D., D'Eleuterio, G.M.T.: *Variational inference as iterative projection in a Bayesian Hilbert space*. Robotica 41(2), 632–667, 2023.

Egozcue, J.J., Díaz-Barrero, J.L., Pawlowsky-Glahn, V.: *Hilbert space of probability density functions based on Aitchison geometry*. Acta Mathematica Sinica 22, 1175-–1182, 2006.

Egozcue, J.J., Pawlowsky-Glahn, V.: *Evidence functions: A compositional approach to information*. Statistics and Operations Research Transactions 42, 1-–24, 2018.

Fačevicová, K., Filzmoser, P., Hron, K.: *Compositional cubes: A new concept for multi-factorial compositions*. Statistical Papers 64, 955–985, 2023.

Genest, C., Hron, K., Nešlehová, J.: Orthogonal decomposition of multivariate densities in Bayes spaces and its connection with copulas. Journal of Multivariate Analysis 198, 105228, 2023.

References

Hoeffding, W.: *A class of statistics with asymptotically normal distribution*. The Annals of Mathematical Statistics 19, 293–325, 1948.

Kuo, F., Sloan, I., Wasilkowski, G., Woźniakowski, H.: *On decompositions of multivariate functions*. Mathematics of Computation 79, 953–966, 2010.

Matys Grygar, T., Radojičić, U., Pavlů, I., Greven, S., Nešlehová, J.G., Tůmová, Š., Hron, K.: Exploratory functional data analysis of multivariate densities for the identification of agricultural soil contamination by risk elements. Journal of Geochemical Exploration 259, 107416, 2024.

Mercadier, C., Roussant, O., Genest, C.: *Linking the Hoeffding-Sobol and Möbius formulas through a decomposition of Kuo, Sloan, Wasilkowski, and Woźniakowski*. Statistics & Probability Letters 185, 109419, 2022.

Ministerstvo zemědělství České Republiky (eAGRI), Register of Contaminated Areas (in Czech), 2020.

Nelsen, R.B.: *An introduction to copulas*, second edition. Springer, New York, 2006.

References

Petersen, A., Zhang, C., Kokoszka, P.: *Modeling probability density functions as data objects*. Econometrics and Statistics 21, 159–178, 2022.

Sklar, A.: *Fonctions de répartition à n dimensions et leurs marges*. Publications de l'Institut de statistique de l'Université de Paris 8, 229–231, 1959.

Talská, R., Menafoglio, A., Hron, K., Egozcue, J.J., Palarea-Albaladejo, J.: *Weighting the domain of probability densities in functional data analysis*. Stat, 9(1), e283, 2020.

van den Boogaart, K.G., Egozcue, J.J., Pawlowsky-Glahn, V.: *Bayes Hilbert spaces*. Australian & New Zealand Journal of Statistics 56, 171–194, 2014.

Wynne, G.: *Bayes Hilbert spaces for posterior approximation*. arXiv:2304.09053, 2023.