General Sensitivity Indices for Hilbert Space-Valued Random Variables

Claude, Brayer
CEA, DES, IRESNE, DTN, Cadarache F-13108 Saint Paul-lez-Durance, France

Reda, Chhaibi
Institut de Mathématiques, UMR5219, Université de Toulouse, CNRS, UPS, F-31062 Toulouse Cedex 9, France

Fabrice, Gamboa
Institut de Mathématiques, UMR5219, Université de Toulouse, CNRS, UPS, F-31062 Toulouse Cedex 9, France

Léonid, Lopez
Institut Mathmatique de Toulouse, 31062 Toulouse, France
CEA, DES, IRESNE, DER, Cadarache F-13108 Saint-Paul-Lez-Durance, France

Clément, Pellegrini
Institut de Mathématiques, UMR5219, Université de Toulouse, CNRS, UPS, F-31062 Toulouse Cedex 9, France

This paper introduces a familly of generalized sensitivity index for Hilbert space-valued random variables, extending the framework of global sensitivity analysis (GSA) to accommodate more complex output spaces. The need for this work stems from the growing complexity of computer models in various scientific and engineering fields, where understanding the influence of inputs on outputs is crucial but often challenging due to computational constraints.

We build on recent advancements in dependence measures [1] and propose new sensitivity indices that quantifie the influence of a real-valued input X on a Hilbert space-valued output Y. This generalization allows for a more thorough analysis of complex systems, accommodating outputs that may be functional or high-dimensional.

The paper begins by establishing the mathematical framework, defining the conditional law and conditional expectation for Hilbert space-valued random variables. We then introduce the concept of equivalent random variables, which is central to the definition of our sensitivity index.

To address the crucial issue of estimation for this proposed generalized sensitivity indices, Λ_φ, we propose an estimation method based on rank statistics, inspired by the work of Gamboa et al. [3] and following the approach introduced by Chatterjee [2]. This method provides advantages over traditional estimation techniques, particularly in terms of computational efficiency and the capability to estimate multiple indices simultaneously.

The rank-based estimation approach uses the ranks of the input and output variables to approximate the sensitivity index. This is particularly advantageous for Hilbert space-valued outputs, as it avoids the complexities associated with direct estimation in high-dimensional spaces.

The estimation procedure involves computing the ranks of the X_i values from a sample of n observations (X_i, Y_i), ordering the Y_i values according to these ranks, and constructing the estimator for Λ_ϕ based on the differences between consecutive ordered $Y_{(R_i)}$ values. This rank-based approach is computationally efficient, allows simultaneous estimation of multiple sensitivity indices, and is robust to outliers and non-linear relationships. Under appropriate conditions, the estimator is consistent and asymptotically normal, providing a basis for constructing confidence intervals and hypothesis tests.

We provide a detailed analysis of the statistical properties of this estimator, including consistency and asymptotic normality results. These theoretical guarantees support the reliability of the proposed estimation method in practical applications.

Furthermore, we conduct numerical experiments to compare the performance of our rank-based estimator with traditional methods. These experiments demonstrate its efficiency and accuracy across a range of scenarios, including those with complex, high-dimensional outputs.
Our work contributes to the field of GSA by offering a rigorous and implementable tool for analyzing the sensitivity of complex, Hilbert space-valued outputs. The proposed index and its estimation method provide a means of understanding input-output relationships in high-dimensional and functional settings, with potential applications in various scientific and engineering disciplines.

References:

[1] Ansari, J., Langthaler, P. B., Fuchs, S., & Trutschnig, W. (2023). Quantifying and estimating dependence via sensitivity of conditional distributions. arXiv. doi:10.48550/ARXIV.2308.06168. Retrieved from https://arxiv.org/abs/2308.06168

[2] Chatterjee, S. (2021). A New Coefficient of Correlation. *Journal of the American Statistical Association*, 116(536), 2009–2022. doi:10.1080/01621459.2020.1758115. Retrieved from https://www.tandfonline.com/doi/full/10.1080/01621459.2020.1758115

[3] Gamboa, F., Gremaud, P., Klein, T., & Lagnoux, A. (2022). Global sensitivity analysis: A novel generation of mighty estimators based on rank statistics. *Bernoulli*, 28(4). doi:10.3150/21-BEJ1421. Retrieved from https://projecteuclid.org/journals/bernoulli/volume-28/issue-4/Global-sensitivity-analysis--A-novel-generation-of-mighty-estimators/10.3150/21-BEJ1421.full

[Claude Brayer; CEA, DES, IRESNE, DTN, Cadarache F-13108 Saint Paul-lez-Durance, France]
[claude.brayer@cea.fr –]

[Reda Chhaibi; Institut de Mathématiques, UMR5219, Université de Toulouse, CNRS, UPS, F-31062 Toulouse Cedex 9, France]
[reda.chhaibi@math.univ-toulouse.fr – https://www.math.univ-toulouse.fr/~rchhaibi/]

[Fabrice Gamboa; Institut de Mathématiques, UMR5219, Université de Toulouse, CNRS, UPS, F-31062 Toulouse Cedex 9, France]
[fabrice.gamboa@math.univ-toulouse.fr – https://www.math.univ-toulouse.fr/~gamboa/]

[Léonid Lopez; Institut Mathmatique de Toulouse, 31062 Toulouse, France
CEA, DES, IRESNE, DER, Cadarache F-13108 Saint-Paul-Lez-Durance, France]
[leonid.lopez@univ-tlse3.fr – https://sites.google.com/view/leonidpersonalwebpage/home]

[Clément Pellegrini; Institut de Mathématiques, UMR5219, Université de Toulouse, CNRS, UPS, F-31062 Toulouse Cedex 9, France]
[clement.pellegrini@math.univ-toulouse.fr – https://perso.math.univ-toulouse.fr/pellegrini/students/]