Optimal-transport based conformal prediction

Claire Boyer, Kimia Nadjahi and Gauthier Thurin

Conformal prediction

Aim: to **quantify uncertainty** in predictive models

Conformal prediction

Aim: to **quantify uncertainty** in predictive models.

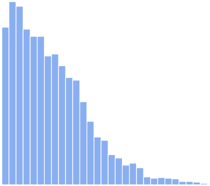
Desired output: interval that contains the ground truth with probability α.

Conformal prediction framework

Arbitrary predictor $\hat{f}$

$\hat{f}(X) \approx Y$

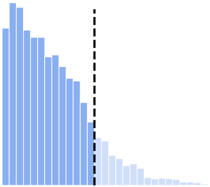
$\longrightarrow$



Scores

$s(X, Y)$

$\longrightarrow$



Quantiles

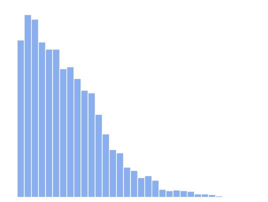
$\mathbb{P}\big(s(X, Y) \leq \hat{q}_\alpha\big) \geq \alpha$

Existing works

Quality of prediction sets depends on the chosen score, so its design has motivated a great deal of work [Angelopoulos and Bates, 2023, Sesia and Romano, 2021, Lei et al., 2018, Chernozhukov et al., 2021].

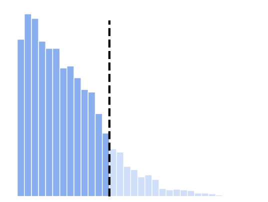
Arbitrary predictor $\hat{f}$ $\longrightarrow$

$\hat{f}(X) \approx Y$



Scores

$\longrightarrow$



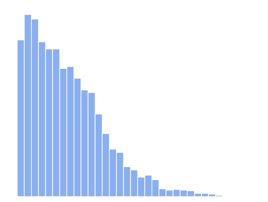
Quantiles

Our work

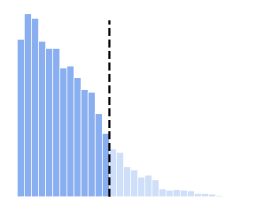
Instead, we focus on the other main ingredient of CP.

Arbitrary predictor $\hat{f}$ $\longrightarrow$

$\hat{f}(X) \approx Y$



$\longrightarrow$



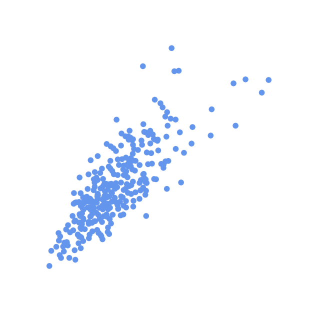
Scores

Quantiles

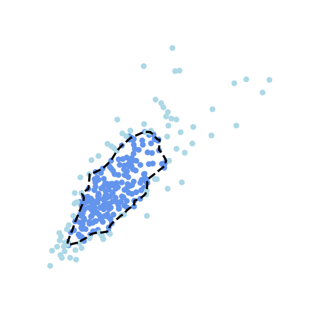
Multivariate quantiles in conformal prediction

General recipe for multivariate scores

From a black box $\hat{f}$, consider multiple scores $s(X, Y) \in \mathbb{R}^d$.



$\longrightarrow$



Scores

Quantiles

$$\mathbb{P}\big(s(X, Y) \in \widehat{\mathcal{Q}}_\alpha\big) \geq \alpha$$

Why multivariate scores ?

Motivations:

1. Multivariate regression $Y \in \mathbb{R}^d$
2. Classification with multiple labels $Y \in \{c_1, c_2, c_3\}$
3. Several notions of errors to be combined $\big(s_1(X, Y), s_2(X, Y)\big)$.

Multivariate quantiles ?

This requires a notion of multivariate quantiles, to produce regions $\widehat{\mathcal{Q}}_\alpha$ such that

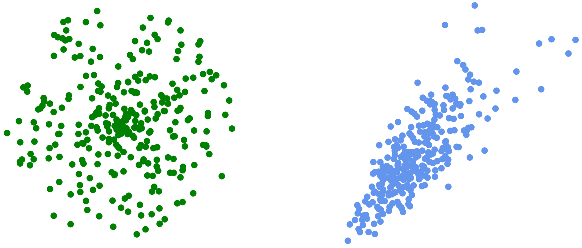
$$\mathbb{P}\big(s(X, Y) \in \widehat{\mathcal{Q}}_\alpha\big) \geq \alpha$$

Plenty of different definitions[1], but we focus on a recent one that uses tools from **measure transportation** [Chernozhukov et al., 2017, Hallin et al., 2021].

[1] [Tukey, 1975, Serfling, 2002]

Transport-based quantiles

For data points $(S_1, \cdots, S_n)$, define ranks $(U_1, \cdots, U_n)$ from a reference distribution (here supported on the unit ball).

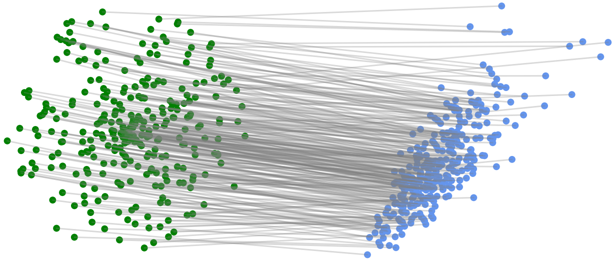


$U_i = \frac{i}{n}\varphi_i$ for φ_i uniformly sampled from the sphere $\{\varphi : \|\varphi\| = 1\}$,

$\Longrightarrow \mathbb{B}(0, \alpha)$ contains a proportion α of ranks

Transport-based quantiles

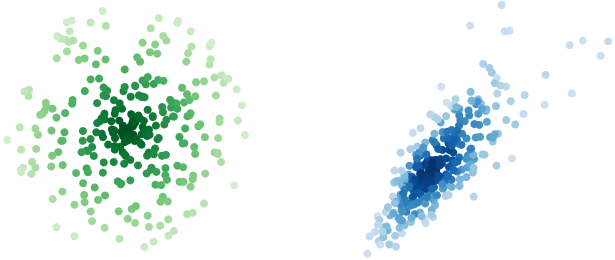
Solve optimal transport problem: $\sigma_n = \underset{\sigma \in P_n}{\operatorname{argmin}} \sum_{i=1}^{n} \|S_i - U_{\sigma(i)}\|^2$.



To each S_i is associated a multivariate rank $U_{\sigma(i)}$.

Center-outward ordering

From this one-to-one correspondance, centrality for one S_i is defined via the centrality of the associated rank $U_{\sigma(i)}$.



A quantile region: the image of $\mathbb{B}(0, \alpha)$ that must contain a proportion α of data S_i.

Example of the methodology for multiple-output regression

From $X \in \mathbb{R}$, a black box $\hat{f} \in \mathbb{R}^2$ to predict $\hat{f}(X) \approx Y \in \mathbb{R}^2$.

$\longrightarrow$

Scores

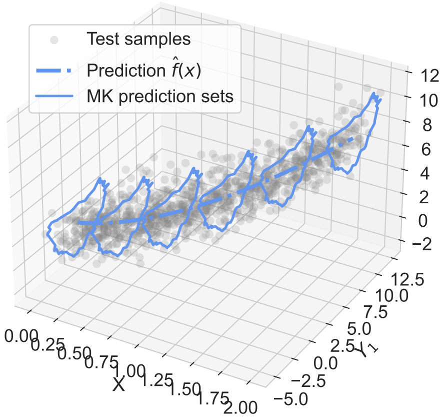
$$s(x, y) = y - \hat{f}(x) \in \mathbb{R}^2$$

Quantiles

$$\mathbb{P}\big(s(X, Y) \in \widehat{\mathcal{Q}}_\alpha\big) \geq \alpha$$

Example: multiple-output regression

$$y - \hat{f}(x) \in \widehat{\mathcal{Q}}_\alpha \Longrightarrow y \in \{\hat{f}(x)\} + \widehat{\mathcal{Q}}_\alpha$$



Angelopoulos, A. N. and Bates, S. (2023).
Conformal prediction: A gentle introduction.
Foundations and Trends® in Machine Learning, 16(4):494–591.

Chernozhukov, V., Galichon, A., Hallin, M., and Henry, M. (2017).
Monge–Kantorovich depth, quantiles, ranks and signs.
The Annals of Statistics, 45(1):223 – 256.

Chernozhukov, V., Wüthrich, K., and Zhu, Y. (2021).
Distributional conformal prediction.
Proceedings of the National Academy of Sciences, 118(48):e2107794118.

References ii

Hallin, M., del Barrio, E., Cuesta-Albertos, J., and Matrán, C. (2021).
Distribution and quantile functions, ranks and signs in dimension d: A measure transportation approach.
The Annals of Statistics, 49(2):1139 – 1165.

Lei, J., G'Sell, M., Rinaldo, A., Tibshirani, R. J., and Wasserman, L. (2018).
Distribution-free predictive inference for regression.
Journal of the American Statistical Association, 113(523):1094–1111.

Romano, Y., Patterson, E., and Candes, E. (2019).
Conformalized quantile regression.
Advances in neural information processing systems, 32.

Serfling, R. (2002).
Quantile functions for multivariate analysis: approaches and applications.
Statistica Neerlandica, 56(2):214–232.

Sesia, M. and Romano, Y. (2021).
Conformal prediction using conditional histograms.
Advances in Neural Information Processing Systems, 34:6304–6315.

Thurin, G., Nadjahi, K., and Boyer, C. (2025).
Optimal transport-based conformal prediction.
arXiv preprint arXiv:2501.18991.

Tukey, J. W. (1975).
Mathematics and the picturing of data.
Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), 2:523–531.