Means and k-means: dimension free PAC-Bayesian bounds for some robust estimators

Olivier Catoni

CREST, CNRS – UMR 9194,
Université Paris-Saclay

We will show, through results obtained in collaboratiton with Ilaria Giulini and Gautier Appert, how PAC-Bayesian bounds can be used as an alternative to other methods to prove concentration inequalities and complexity bounds. We will take the example of the robust estimation of the mean of a random vector. We will then discuss the problem of vector quantization according to a robust modification of the k-means loss function.

References

[AC21]   Gautier Appert and Olivier Catoni. "New bounds for k-means and information k-means." In: *arXiv preprint arXiv:2101.05728* (2021). URL: https://arxiv.org/abs/2101.05728.

[Cat12]   Olivier Catoni. "Challenging the empirical mean and empirical variance: A deviation study." In: *Ann. Inst. H. Poincaré Probab. Statist.* 48.4 (Nov. 2012), pp. 1148–1185. DOI: 10.1214/11-AIHP454. URL: https://doi.org/10.1214/11-AIHP454.

[CG17a]   Olivier Catoni and Ilaria Giulini. "Dimension free PAC-Bayesian bounds for the estimation of the mean of a random vector." In: *the Nips 2017 Workshop : (Almost) 50 shades of Bayesian learning : PAC-Bayesian trends and insights.* 2017, pp. 1–4. eprint: https://bguedj.github.io/nips2017/pdf/PAC-Bayes_2017_paper_1.pdf. URL: https://bguedj.github.io/nips2017/50shadesbayesian.html.

[CG17b]   Olivier Catoni and Ilaria Giulini. "Dimension-free PAC-Bayesian bounds for matrices, vectors, and linear least squares regression." In: *arXiv preprint arXiv:1712.02747* (2017). URL: https://arxiv.org/abs/1712.02747.