

On rank regression, minimization of U -processes and some probabilistic inequalities

Wojciech Rejchel

*Nicolaus Copernicus University, Toruń, Poland
University of Warmia and Mazury, Olsztyn, Poland*

Abstract

We consider the problem of rank regression, that is basing on some features of the objects we want to predict (guess) the order between these objects. In many algorithms the $0 - 1$ loss function is replaced by its convex surrogate - this trick allows to find very effective procedures (support vector machines [Vapnik, 1998] or boosting [Freund, 2004]). We are interesting in statistical properties of rank regression estimators. First, we show that in the linear model one can prove the strong consistency and the asymptotical normality of estimates [Niemiro, 2009]. Moreover, we present another very popular approach to the problem, namely, we look for probabilistic inequalities concerning the risk or the excess risk (i.e. comparing to the best one in the class) of estimators [Clemencon, 2008]. Finally, we show how one can obtain significant improvements in such inequalities [Rejchel, 2009]. The methods that we use come from the theory of empirical processes and U -processes.

Keywords

Rank regression, Convex risk, Empirical process, U -process.

References

- Clemencon, S., Lugosi, S., and Vayatis, N. (2008). Ranking and empirical minimization of U-statistics. *Ann. Statist.* 36, 844–874.
- Freund, Y., Iyer, R., Schapire, R.E., and Singer, Y. (2004). An efficient boosting algorithm for combining preferences. *J. Mach. Learn. Res.* 4, 933–969.

Niemiro, W. and Rejchel, W. (2009). Rank correlation estimators and their limiting distributions. *Statist. Papers* 50, 887–893.

Rejchel, W. (2009). Ranking - convex risk minimization. *Proc. of WASET* 56, 172–178.

Vapnik, V.N. (1998). *Statistical learning theory*. Wiley, New York.