Suplementary Material of

GROS:

A General Robust Aggregation Strategy

Alejandro Cholaquidis [1], Emilien Joly[2], Leonardo Moreno [3]

[1] Centro de Matemáticas, Facultad de Ciencias, UDELAR, Uruguay.

[2] Centro de Investigación en Matemáticas, CIMAT, Guanajuato, México

[3] Instituto de Estadística, Departamento de Métodos Cuantitativos, Facultad de Ciencias Económicas y de Administración, UDELAR, Uruguay.

1. Heavy-tailed bandits

To get a logarithmic regret for the UCB (i.e., $R_T/\log(T) \to C$), it is usually assumed that the distributions P_i are sub-Gaussian. This can be weakened to require the distributions to have finite moment generating function, see Agrawal (1995). To overcome this limitation, instead of just estimating the mean at each step t of the UCB, in Bubeck, Cesa-Bianchi and Lugosi (2013) it is proposed to use a robust estimator, $\widehat{\mu}_{i,t}$, of $\mathbb{E}(P_i)$, that fulfills the following assumption:

Assumption 1 Let $\epsilon \in (0,1]$ be a positive parameter, and let c, v be a positive constant. Let $X_1, \ldots, X_n$ be i.i.d. random variables with finite mean μ. Suppose that for all $\delta \in (0,1)$ there exists an estimator $\widehat{\mu} = \widehat{\mu}(n, \delta)$ such that, with probability at least $1 - \delta$,

$$\widehat{\mu} \leq \mu + v^{1/(1+\epsilon)} \left(\frac{c \log(1/\delta)}{n} \right)^{\epsilon/(1+\epsilon)} \tag{1}$$

and also, with probability at least $1 - \delta$,

$$\mu \leq \widehat{\mu} + v^{1/(1+\epsilon)} \left(\frac{c \log(1/\delta)}{n} \right)^{\epsilon/(1+\epsilon)}. \tag{2}$$

In that case we say that $\widehat{\mu}$ fulfills Assumption 1.

Remark 1. From

$$\mathbb{P}\left(d(\mu^*, \mu) > 4\sqrt{\mathbb{E}d^2(\mu_1, \mu)} \right) \leq \delta. \tag{3}$$

our robust proposal μ^*, applied on each arm, and based on $K = \lceil 8 \log(1/\delta) \rceil$ groups, fulfills Assumption 1, for $\epsilon = 1$, $v = \mathbb{V}(P_i)$, ($\mathbb{V}(P_i)$ being the variance of the distribution P_i), and $c = 16$. In that case, for an arm $i = 1, \ldots, L$, $n = N_{t,i}/K$.

Bubeck, Cesa-Bianchi and Lugosi (2013) proposes the following robust variant of the UCB: given $\epsilon \in (0,1]$, for arm i, define $\widehat{\mu}_{i,s,t}$ as the estimator $\widehat{\mu}(s, t^{-2})$ based on the first s observed values $X_{i,1}, \ldots, X_{i,s}$ of the reward of

arm i. Define the index

$$B_{i,s,t} = \widehat{\mu}_{i,s,t} + v^{1/(1+\epsilon)} \left(\frac{c \log(t^2)}{s} \right)^{\epsilon/(1+\epsilon)}$$

for $s, t \geq 1$ and $B_{i,0,t} = +\infty$. Then, at time t draw an arm maximizing $B_{i,N_{i,t-1},t}$.

We propose the following algorithm. First we choose the arms at random from $t = 1, \ldots, t_0$, where t_0 guarantees that for all $i = 1, \ldots, L$, $N_{t_0,i}/\lceil 8 \log(t_0^2) \rceil \geq 1$. We compute, for each arm, the estimator given by Equation (1), denoted by $\mu^*_{t_0,i}$, where we split the $N_{t_0,i}$ observations into $K = \lceil 8 \log(t_0^2) \rceil$ groups, and compute the mean of each group. Define the index

$$\mathcal{B}_{i,N_{t_0,i},t_0} = \mu^*_{t_0,i} + 4\sqrt{\widehat{\mathbb{V}}(P_i)} \left(\frac{\log(t_0^2)}{N_{t_0,i}} \right)^{1/2},$$

where $\widehat{\mathbb{V}}(P_i)$ is any consistent estimation of $\mathbb{V}(P_i)$, or an upper bound. At time $t_0 + 1$, we choose the arm that maximize $\mathcal{B}_{i,N_{t_0,i},t_0}$, at time $t_0 + 2$ we choose the arm that maximize $\mathcal{B}_{i,N_{t_0+1,i},t_0+1}$, and so on.

Proposition 1 in Bubeck, Cesa-Bianchi and Lugosi (2013) proves that this algorithm attains logaritmic regret. More precisely,

$$R_T \leq \sum_{i:\Delta_i > 0} \left(32 \left(\frac{\mathbb{V}(P_i)}{\Delta_i} \right) \log(T) + 5\Delta_i \right),$$

where $\Delta_i = \mathbb{E}(P_{i^*}) - \mathbb{E}(P_i)$.

1.0.1 Simulations

As a toy example, let us consider the classical two-armed bandit problem and rewards given by $X_t|\mathcal{A}_t = j \sim \mu_j + S(3)$, for $j = 1, 2$, where $S(3)$ is a random variable following Student's distribution with 3 degrees of freedom, $\mu_1 = 7$, and $\mu_2 = 8$. In our algorithm, indicated by RUCB in Figure 1, the first $t_0 = 40$ arms are chosen at random.

The results are shown in Figure 1, where the red dotted horizontal line ($y = 8$) is the maximum expected gain. The dashed lines corresponds to the mean rewards of the UCB (orange) and the RUCB (blue) respectively. As can be seen, it takes the RUCB algorithm about 120 steps to outperform the UCB algorithm, and the difference grows larger as the number of steps increases.

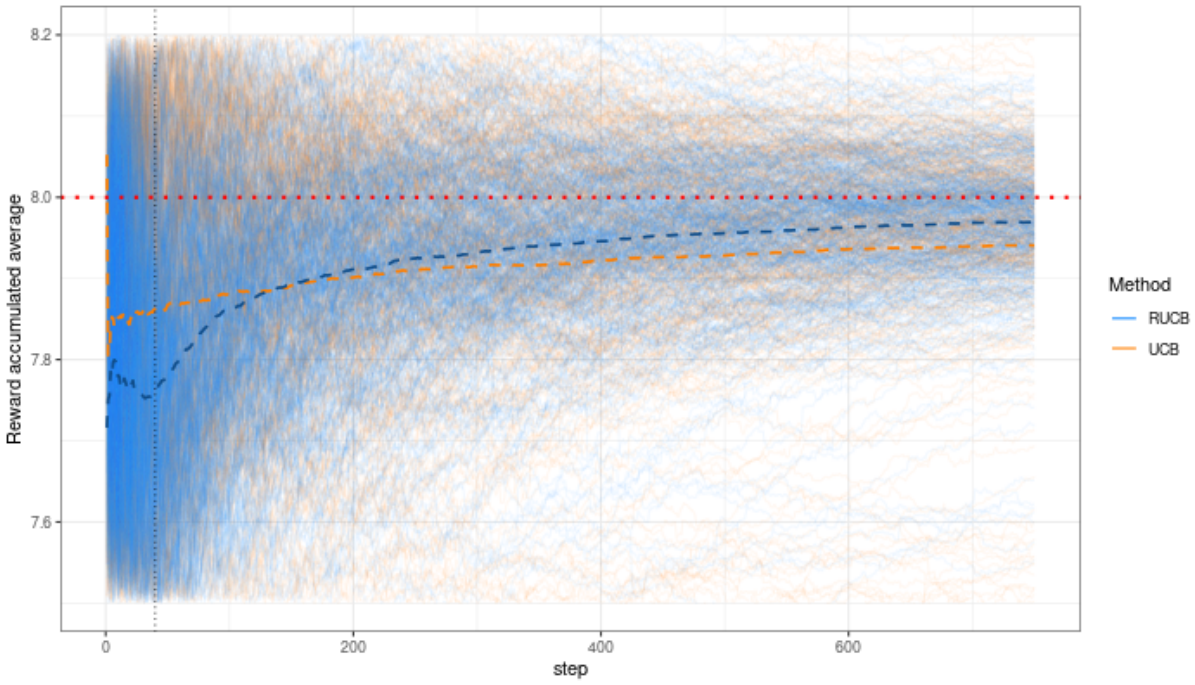


Figure 1: Cumulative gains over 500 replications, for $t = 1, \ldots, 750$. The red dotted horizontal line ($y = 8$) is the maximum expected gain. The black dotted vertical line ($x = 40$) indicates the number of random warm-up runs in the RUCB algorithm. The dashed lines depict the mean reward of the UCB (orange) and RUCB (blue) algorithms.

2. Robust regression

A current topic of interest in statistics is that of regression models in the presence of noise. It is known that a small fraction of outliers can cause severe biases in classical regression estimators. These classical models gen-

erally assume additive residuals in the model with finite second moment (e.g., Gaussian). An extensive review of robust outlier estimation methods for nonparametric regression models is provided in Salibian-Barrera (2023).

We tackle the problem of estimating the function $m : \mathcal{X} \to \mathbb{R}$, from an i.i.d. sample of a random element $(X, Y) \in \mathcal{X} \times \mathbb{R}$, that satisfies the model

$$Y = m(X) + \epsilon$$

where ϵ is a noise term such that $\mathbb{E}[\epsilon|X] = 0$. To keep the following discussion fairly simple, we will only cover the case $\mathcal{X} = \mathbb{R}$. In the context where very little information is known about m, a general estimator is often given by the kernel estimator

$$\widehat{m}(x) = \frac{\sum_{i=1}^{n} K_h(X_i - x)Y_i}{\sum_{i=1}^{n} K_h(X_i - x)}, \tag{4}$$

where $K_h(x) = h^{-1}\phi(h^{-1}x)$ and $h > 0$ is some bandwidth parameter. The function ϕ is non-negative and such that $\int_{\mathbb{R}} \phi(x)dx = 1$. In general, the continuity of the function m is enough to get the consistency of the kernel estimator $\widehat{m}$ and a light tailed behavior for ϵ gives sub-exponential deviation bounds for the estimator around its mean value m. Nevertheless, in various concrete settings, one can face heavy tailed distributions for ϵ in such a way that the estimator proposed in (4) becomes highly unstable. Indeed, the

presence of (virtually) one outlier is enough to drive the estimator towards values very distant from m. One natural way to measure the quality of the estimator $\widehat{m}$ is to consider the L_2 distance

$$d_2(\widehat{m}, m) = \mathbb{E}\left[|\widehat{m}(X) - m(X)|^2\right]^{1/2}.$$

In this context, it is possible to introduce the robust version of the estimator by considering GROS with the distance d_2. In practice, the distance d_2 is actually intractable. To overcome this difficulty, it is common to use a discretization on a mesh for the space $\mathcal{X}$ and approximate the integral by a Riemann sum.

This estimator verifies a bound as in Theorem 2 for the associated distance. The estimator m^* is, in the sense of Equation (1), the best candidate in the class of the estimator $m_1, \ldots, m_K$ constructed on the disjoint groups $G_1, \ldots, G_K$ of data points. These estimators $m_1, \ldots, m_K$ are defined by

$$m_j(x) = \frac{\sum_{i \in G_j} K_h(X_i - x) Y_i}{\sum_{i \in G_j} K_h(X_i - x)},$$

so that these estimators are independent. Following subsection 2.1, the following estimator for $\widehat{m}$ can be proposed:

$$\widehat{m} = \underset{m \in \widetilde{\mathcal{M}}}{\operatorname{argmin}} \min_{I : |I| > \frac{K}{2}} \max_{j \in I} d_2(m_j, m),$$

where the $(m_j)_{j=1}^K$ are the kernel estimates of m on each of the K groups, and the set $\widetilde{\mathcal{M}}$ denotes the set of functions that are piece-wise equal to

one of the m_i. The difference from the naive definition is that the set of minimization does not end with an estimator of the form m_{i^*}, which allows avoiding choosing a function that may be a good fit in some regions of the space but sensitive to outliers in other parts.

In order to fully use the result of Equation (4), we cite a result that gives an upper bound for the mean squared error for any of the estimators m_j previously defined. In chapter 5 of (Györfi, Kohler, Krzyzak and Walk, 2002, Theorem 5.2), we get that if m is λ-Lipschitz, and that $\mathrm{Var}(Y|X = x) \leq \sigma^2$ for all $x \in \mathcal{X}$, then

$$\mathbb{E}\left[\|m_1 - m\|^2\right] \leq c\left(\frac{\sigma^2 + \sup_{x\in\mathcal{X}} m(x)^2}{(n/K)h^d}\right) + \lambda^2 h^2.$$

This induces a choice of h of the form

$$h = c'\left(\frac{K}{n}\right)^{\frac{1}{d+2}}.$$

The upper bound then takes the form of

$$\mathbb{E}\left[\|m_1 - m\|^2\right] \leq c''\left(\sigma^2 + \sup_{x\in\mathcal{X}} m(x)^2\right)^{2/d+2} \times \left(\frac{n}{K}\right)^{-2/d+2},$$

which can be directly plugged into the main bound of the theorem.

2.0.1 Simulations

We compare, by means of simulations, the performance of GROS with some classical and robust regression alternatives proposed in the literature.

In this example we consider a sample $X_1, \ldots, X_{1000}$ with uniform distribution on $[0, 5]$. Let $m(x) = 4\sin(3x)$ and suppose that the centered noise follow the skew-normal Student distribution, see Fernández and Steel (1998); Azzalini (2013), whose density is defined as follows: denote by $t(x, \kappa, \nu)$ the density of the non-standardized Student's distribution with location κ and and ν degrees of freedom, and with cumulative distribution function $T(x, \kappa, \nu)$. Then, the density of the skew-normal Student is

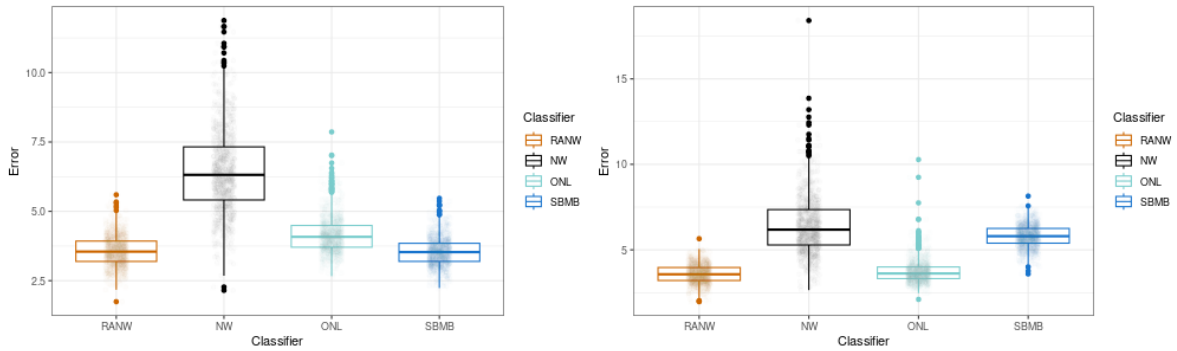
$$f(x; \kappa, \nu, \sigma, \xi) = \frac{1}{\sigma} t(x/\sigma, \kappa, \nu) T\left(\frac{\xi x}{\sigma}, \kappa, \nu\right),$$

where $\sigma > 0$ denotes the scale parameter. The slant (or skewness parameter) is ξ.

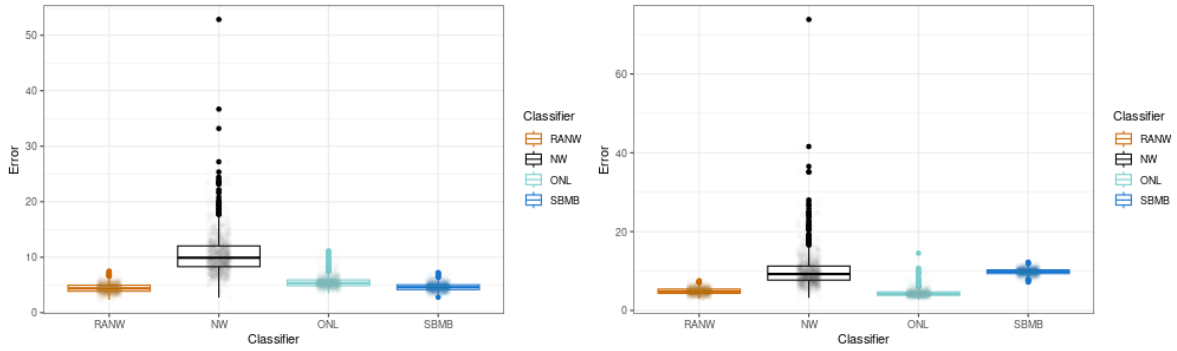
We will write NW for the non-parametric Nadaraya–Watson kernel regression estimator (4), see Nadaraya (1964); Watson (1964). As robust estimators, we consider the proposals developed in Oh, Nychka and Lee (2007) (ONL estimator) and Boente, Martínez and Salibián-Barrera (2017) (SBMB estimator). Both estimators are implemented in the **R** language: the first in the **fields** library and the other in the **RBF** library. Our estimator in this context will be called RANW (Robust Aggregation for Naradaya–Watson). For the latter, $K = 12$ was considered. Figure 3 shows the estimated regressions in each scenario and in red the true function f. It is clear that the NW estimator performs poorly in all cases.

The performance of each estimator is measured by the average distance $d_2(\widehat{m}, m)$ over 1000 replicates, considering $\kappa = 0$, $\nu = 3$ (note that the noises have heavy tails) and varying the parameters σ and ξ. For the bandwidth parameter h, we choose 0.2 in all cases.

Figure 2 shows box plots of the errors for four choices of the parameters. As can be seen, when the noise is asymmetric (i.e., $\xi \neq 1$), the best-performing estimators are RANW and ONL, depending on the value of the scale parameter. Note that when the distribution of the noise is symmetric (i.e., $\xi = 1$), SBMB and RANW perform the best.

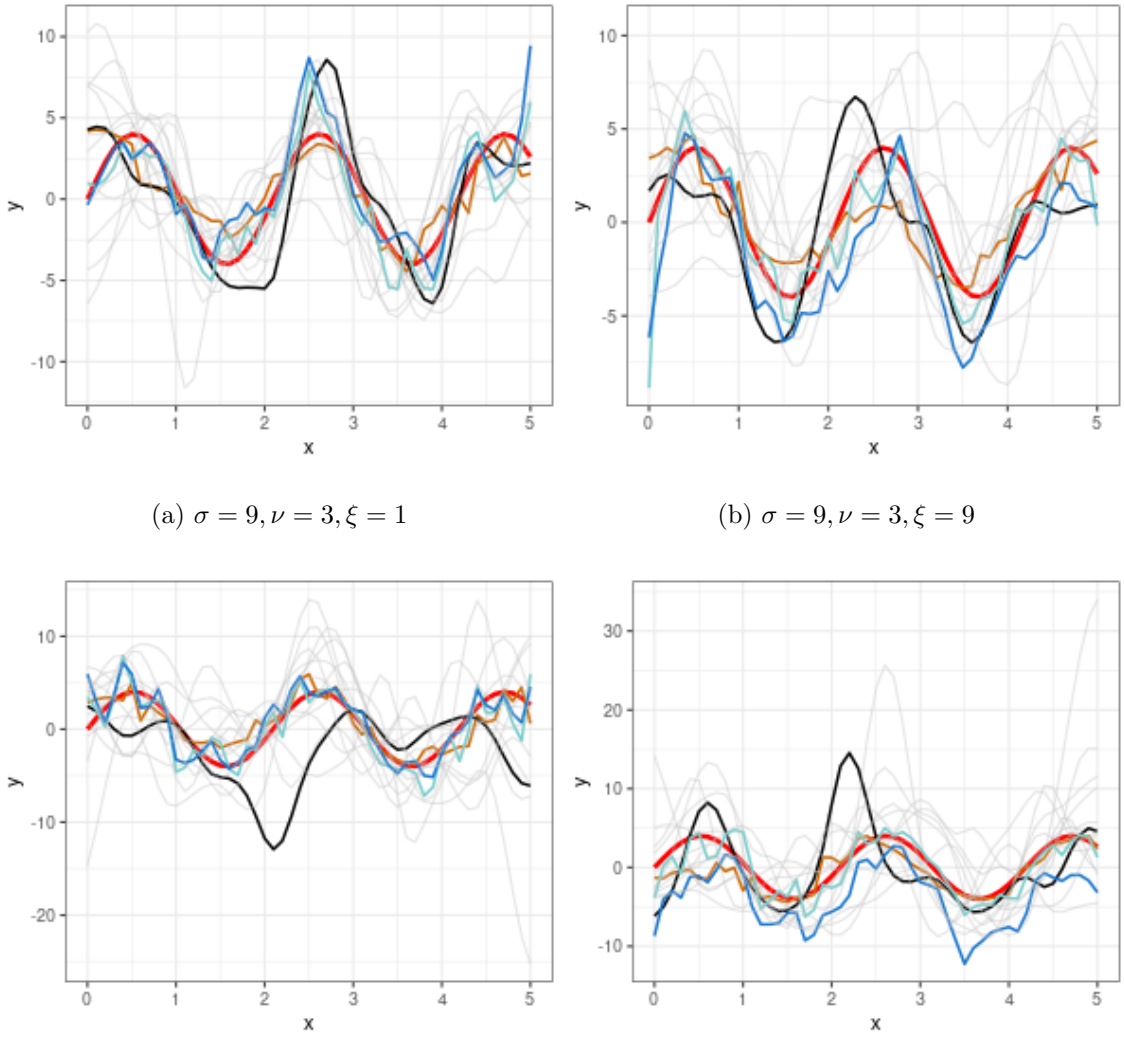


(a) $\sigma = 9, \nu = 3, \xi = 1$ (b) $\sigma = 9, \nu = 3, \xi = 9$



(c) $\sigma = 16, \nu = 3, \xi = 1$ (d) $\sigma = 16, \nu = 3, \xi = 9$

Figure 2: Box plot of classification errors (according to L2 distance) in 1000 replicates. The different scenarios are obtained in the skew-normal Student distribution with $\sigma \in \{9, 16\}$ and $\xi \in \{1, 9\}$, fixed $\nu = 3$ and $\kappa = 0$.



(a) $\sigma = 9, \nu = 3, \xi = 1$

(b) $\sigma = 9, \nu = 3, \xi = 9$

(c) $\sigma = 16, \nu = 3, \xi = 1$

(d) $\sigma = 16, \nu = 3, \xi = 9$

Figure 3: Regression functions estimated with the RANW (orange), NW (black), ONL (light blue) and SBMB (blue) in one replicate. The true function is shown in red. The different scenarios are obtained in the skew-normal Student distribution with $\sigma \in \{9, 16\}$ and $\xi \in \{1, 9\}$, fixed $\mu = 0$ and $\nu = 3$.

References

Agrawal, R. (1995). Sample mean based index policies by $o(\log n)$ regret for the multi-armed bandit problem. *Advances in Applied Probability* **27**(4), 1054–1078.

Bubeck, S., Cesa-Bianchi, N. and Lugosi, G. (2013). Bandits with heavy tail. *IEEE Transactions on Information Theory* **59**(11), 7711–7717.

Fernández, C. and Steel, M. F. J. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association* **93**(441), 359–371.

Azzalini, A. (2013). *The Skew-Normal and Related Families*. Institute of Mathematical Statistics Monographs. Cambridge University Press.

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications* **9**(1), 141–142.

Watson, G. S. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A* **26**(4), 359–372.

Oh, H.-S., Nychka, D. W. and Lee, T. C. M. (2007). The role of pseudo data for robust smoothing with application to wavelet regression. *Biometrika* **94**(4), 893–904.

Boente, G., Martínez, A. and Salibián-Barrera, M. (2017). Robust estimators for additive models using backfitting. *Journal of Nonparametric Statistics* **29**(4), 744–767.

Györfi, L., Kohler, M., Krzyzak, A. and Walk, H. (2002). *A Distribution-free Theory of Nonparametric Regression*. Springer.

Salibian-Barrera, M. (2023). Robust nonparametric regression: Review and practical considerations. *Econometrics and Statistics*.