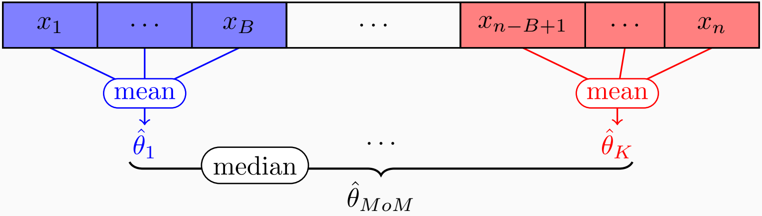
On Medians of (Randomized) Pairwise Means

Pierre Laforgue[1], Stephan Clémençon[1], Patrice Bertail[2]

[1] LTCI, Télécom Paris, Institut Polytechnique de Paris, France
[2] Modal'X, UPL, Université Paris-Nanterre, France

The Median of Means

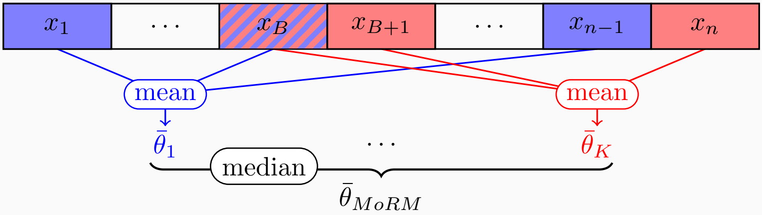


$x_1, \ldots, x_n$ i.i.d. realizations of r.v. X s.t. $\mathbb{E}[X] = \theta$, $\mathrm{Var}(X) = \sigma^2$.

$\forall \delta \in [e^{1-\frac{n}{2}}, 1[$, set $K := \lceil \log(1/\delta) \rceil$, it holds [Devroye et al., 2016]:

$$\mathbb{P}\left\{ \left| \hat{\theta}_{\mathrm{MoM}} - \theta \right| > 2\sqrt{2} e \sigma \sqrt{\frac{1 + \log(1/\delta)}{n}} \right\} \leq \delta.$$

The Median of Randomized Means (1st contribution)



With blocks formed by SWoR, $\forall\ \tau \in]0, 1/2[,\ \forall\ \delta \in [2e^{-\frac{8\tau^2 n}{9}}, 1[$, set $K := \left\lceil \frac{\log(2/\delta)}{2(1/2-\tau)^2} \right\rceil$, and $B := \left\lfloor \frac{8\tau^2 n}{9 \log(2/\delta)} \right\rfloor$, it holds:

$$\mathbb{P}\left\{ \left| \bar{\theta}_{\text{MoRM}} - \theta \right| > \frac{3\sqrt{3}\ \sigma}{2\ \tau^{3/2}} \sqrt{\frac{\log(2/\delta)}{n}} \right\} \leq \delta.$$

U-statistics & Pairwise Learning

Estimate $\mathbb{E}[h(X_1, X_2)]$ from an i.i.d. sample $x_1, \ldots, x_n$:

$$U_n(h) = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} h(x_i, x_j).$$

Ex: the empirical variance when $h(x, x') = \frac{(x-x')^2}{2}$.

Encountered *e.g.* in *pairwise ranking* or in *metric learning*:

$$\widehat{\mathcal{R}}_n(r) = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} \mathbb{1}\left\{ r(x_i, x_j) \cdot (y_i - y_j) \le 0 \right\}.$$

The Median of (Randomized) U-statistics (2nd contribution)

Blocks are formed either by partitioning or by SWoR. Medians of the (randomized) U-statistics verify $\forall \tau \in]0, 1/2[$ w.p.a.l. $1 - \delta$:

$$\left|\hat{\theta}_{\mathrm{MoU}} - \theta(h)\right| \leq \sqrt{\frac{C_1 \log \frac{1}{\delta}}{n} + \frac{C_2 \log^2(\frac{1}{\delta})}{n\left(2n - 9 \log \frac{1}{\delta}\right)}},$$

$$\left|\bar{\theta}_{\mathrm{MoRU}} - \theta(h)\right| \leq \sqrt{\frac{C_1(\tau) \log \frac{2}{\delta}}{n} + \frac{C_2(\tau) \log^2(\frac{2}{\delta})}{n\left(8n - 9 \log \frac{2}{\delta}\right)}},$$

with $C_1(\tau) \xrightarrow[\tau \to \frac{1}{2}]{} C_1$ and $C_2(\tau) \xrightarrow[\tau \to \frac{1}{2}]{} C_2$.

The Pairwise Tournament Procedure (3rd contribution)

Adapted from [Lugosi and Mendelson, 2016].

We want to find $f^* \in \underset{f \in \mathcal{F}}{\operatorname{argmin}}\ \mathcal{R}(f) = \mathbb{E}[\ell(f,(X,X'))]$.

For $f \in \mathcal{F}$, let $H_f := \sqrt{\ell(f,X,X')}$. For any pair $(f,g) \in \mathcal{F}^2$:

1) Compute the MoU estimate of $\|H_f - H_g\|_{L_1}$

$$\Phi_{\mathcal{S}}(f,g) = \text{median}\left(\hat{U}_1|H_f - H_g|, \ldots, \hat{U}_K|H_f - H_g|\right).$$

2) If it is *large enough*, compute the *match*

$$\Psi_{\mathcal{S}'}(f,g) = \text{median}\left(\hat{U}_1(H_f^2 - H_g^2), \ldots, \hat{U}_{K'}(H_f^2 - H_g^2)\right).$$

$\hat{f}$ winning all its matches verify w.p.a.l. $1 - \exp(c_0 n \min\{1, r^2\})$

$$\mathcal{R}(\hat{f}) - \mathcal{R}(f^*) \le cr.$$

Conclusion

- MoM exhibits good guarantees with few assumptions

Conclusion

- MoM exhibits good guarantees with few assumptions
- **1st contrib.** Guarantees preserved through randomization

Conclusion

- MoM exhibits good guarantees with few assumptions
- **1$^{\text{st}}$ contrib.** Guarantees preserved through randomization
- **2$^{\text{nd}}$ contrib.** Extension to (randomized) U-statistics

Conclusion

- MoM exhibits good guarantees with few assumptions
- **1st contrib.** Guarantees preserved through randomization
- **2nd contrib.** Extension to (randomized) U-statistics
- **3rd contrib.** Pairwise tournament procedure

References

- Devroye, L., Lerasle, M., Lugosi, G., Oliveira, R. I., et al. (2016).
  Sub-gaussian mean estimators.
  The Annals of Statistics, 44(6):2695–2725.
- Lugosi, G. and Mendelson, S. (2016).
  Risk minimization by median-of-means tournaments.
  arXiv preprint arXiv:1608.00757.
- Minsker, S. et al. (2015).
  Geometric Median and Robust Estimation in Banach Spaces.
  Bernoulli, 21(4):2308–2335.