

On Weighted Multivariate Sign Functions

Subhabrata Majumdar and Snigdhasu Chatterjee

Response to Comments by the Reviewer

We thank the reviewer for their considered comments, and apologize for the delay in preparing the response. We had some personal emergencies, followed by the pandemic, which greatly delayed the revision of our paper. Again, we apologize. The comments by the reviewer led to a substantial revision of the manuscript, primarily in Section 2 of the paper. We address the reviewer’s comments in detail below.

Reviewer comment: *The paper introduces generalizations to the spatial median that employ data-dependent weights, and applies proposed methods to robust scatter estimation of elliptically symmetric distributions as well as closely related problems. Presented ideas certainly deserve attention, however, I feel that the manuscript requires significant work. In particular, it was not clear to me that proposed estimators can be actually computed without knowledge of parameters of the distribution that are not available. I suggest that authors clarify this point early on.*

Author response: We thank the reviewer for finding that the ideas presented in this paper deserve attention. Based on the comments by the reviewer, we made significant changes in the paper, primarily in Section 2.

The proposed estimators can indeed be estimated without knowing the parameters, otherwise this would have been a meaningless exercise. It was our fault in not making the writing of the previous version of the paper clear enough: with the help of the reviewer’s comments we hope that the present version is very transparent on this matter. In the previous version of this paper, there were three significant confusing issues, which we have now fixed.

First, as the reviewer commented in her/his detailed comment (a), there was the question of whether the weights depend on q , the parameter we wish to estimate in Section 2 of the paper. We make a detailed response to this comment below, but in short, we have now restricted the parameter \mathbb{F} (the distribution parameter in the weight function) to match \mathbb{F}_X (the distribution from where the data is generated), and we substantially revised Theorem 2.1, Corollary 2.2, Corollary 2.4, to make sure that our results are valid even if the weights depend on q . More details on this are given in response to the reviewer’s comment (a) below.

Second, there was the issue of how to compute the weights since \mathbb{F} is unknown. We have used the reviewer’s suggestion in their detailed comment (i), and introduced the

computation of weights based on the empirical distribution function \mathbb{F}_n early on, in Section 1. We hope this step will greatly alleviate the reviewer’s concerns about the usability and computation of the proposed estimators without knowing the underlying distribution (or its parameters).

Third, we have now changed completely the technical conditions and proofs of Corollary 2.2 and Corollary 2.4, so that these no longer have any part that assumes that the value of q_0 is known. Our previous manuscript was partially written keeping hypothesis tests in mind (and hence the presence of a known value in the earlier incarnation of the corollaries), but these were both incompletely developed as well as poorly presented. In the revised version of the manuscript, the new corollaries have a different set of technical conditions that broadens their appeal, and completely different proofs.

We trust these changes will address the major concern of the reviewer. We address their specific detailed and minor comments point-by-point below.

Reviewer detailed comment (a): *Do I understand correctly that your weights are not allowed to depend on q ? Then your claim on top of page 3 that this is the same object that has been studied before (e.g., in [Koltchinskii \(1997\)](#)) is incorrect. If it is allowed to depend on q , then please explain how your results are different from the existing literature.*

Author response: The framework of our paper is somewhat different than that of [Koltchinskii \(1997\)](#) (and related papers). In principle, the idea of replacing an observation with a weighted version of its (generalized) sign can be used in numerous ways, only some of which may need the theoretical machinery developed in [Koltchinskii \(1997\)](#) or related papers. For example, in a testing context (not presented in the current manuscript), under the null hypothesis, \mathbb{F} may be entirely known, hence the weights may not depend at all on \mathbb{F}_X or its related parameter q . However, in many other contexts, there may be implicit dependence between the weights $W(\cdot, \mathbb{F})$ and a multivariate quantile q . For example, if $\mathbb{F} = \mathbb{F}_X = N(\mu, \Sigma)$, both the quantile and the weight depend on the parameters μ and Σ , so there may be implicit dependence between the two.

We thought deeply about this issue and decided that the best course of action for the present paper is restrict our attention to the case of $\mathbb{F} = \mathbb{F}_X$, and make explicit the fact that the weights $W(\cdot, \mathbb{F})$ may implicitly depend on q . That way, we are not only able to stay within the paradigm developed in the elegant paper by Koltchinskii ([Koltchinskii \(1997\)](#) and elsewhere) and leverage the theoretical understandings developed in these papers and related literature, we are actually able to bring those results into the application domain. Also, with the assumption that $\mathbb{F} = \mathbb{F}_X$, there is no room for ambiguity whether $W(\cdot, \mathbb{F})$ and q may be related: they are indeed related (except for perhaps simple instances).

With this assumption, our Theorem 2.1 does indeed strongly relate to existing literature, and we present a brief sketch of how that result is established following [Niemiro \(1992\)](#) and [Haberman \(1989\)](#). **However, all other results in our manuscript are new**, and relate to issues that have not been discussed in the literature before, to the best of our knowledge. Following Theorem 2.1, we present two corollaries that show the efficiencies that may be achieved by careful choice of weights. Section 3 onwards, the discussion is on the robust dispersion measures, eigenvalues and principal components, in-

fluence functions in the context of robust estimators of dispersion, and numerical results. To the best of our knowledge, the theoretical results from these sections are completely new. Also, the methodological parts of these sections typically compare favorably to the *ad hoc* or otherwise robust dispersion estimators that may be found in statistics literature, as demonstrated by the numerical results in our paper.

As mentioned in the introduction, our focus is primarily on robust dispersion and principal component analysis. We believe we have been able to build a bridge between the elegant theoretical development on M -quantiles and the methodological aspects of robust dispersion estimation.

Reviewer detailed comment (b): *Regarding the proof of Theorem 2.1: if you claim that this is a new result, the proof can not be omitted completely: please include it in the supplementary material. If it is not new please give a precise reference.*

Author response: We do not claim that this result is substantially new (see detailed comments above). Consequently, in the present manuscript we have included a proof sketch of the theorem, that follows the arguments presented in Niemiro (1992) and Haberman (1989). Please find this sketch of proof right after Theorem 2.1. If needed, we can present a complete proof of this in the supplementary material.

Reviewer detailed comment (c): *I could not understand Corollary 2.2: is q_0 still a minimizer of $\Psi(q; 0; \mathbb{F})$? If so, how can the weight depend on this unknown quantity? Even if q_0 is a fixed vector, this still does not make sense to me. Additionally, please show that conditions A1-A3 hold in your examples.*

Author response: We have now completely changed this corollary. Thus, this issue is no longer present in the current manuscript.

Reviewer detailed comment (d): *Page 3 states “We report only the case of the weighted spatial median for brevity..” In this case, it could be better to focus on the median everywhere and avoid mentioning general quantiles.*

Author response: We thank the reviewer for this suggestion, and have adopted it, and do not elaborate on weighted spatial quantiles in this paper. Following the referee’s suggestion, in the present manuscript we only mention their general definition in Section 2, and then discuss the weighted spatial median in detail.

Reviewer detailed comment (e): *Corollary 2.4 states “then $S(X; q_0)$ is uniform on the unit ball in \mathbb{R}^p ..” – do you mean $S(Z; q_0)$? If X has arbitrary elliptically symmetric distribution, this is not true. Moreover, why is it interesting in this case to consider the case when \mathbb{F} is the (unknown) distribution of X ?*

Author response: The reviewer is absolutely correct, the spatial sign is indeed not independent of the norm for non-spherical but elliptical distributions. We have now completely changed this corollary.

With a few weak conditions on Σ from Wang et al. (2015) we are able to prove that the ARE is larger than 1 for general elliptical distributions.

We thank the referee for this comment, which led to a rigorous revision in this part.

Reviewer detailed comment (f): *Page 5: “..we propose using $W(x; \mathbb{F}_n)$: please define what you mean by an empirical distribution function here, as X is a vector.*

Author response: Thank you for pointing this out, we have now put in the proper definition. Please see page 7 of the current manuscript.

Reviewer detailed comment (g): *The weight functions discussed in section 2.1 can only be computed if the distribution of X is known. Why would such functions be of any use in statistical context? If the unknown distribution is replaced by its empirical version, why would all the useful properties be preserved? I suggest explaining these points clearly in the beginning.*

Author response: This is an excellent suggestion and we are happy to accept it. We have now put a brief discussion on this in the Introduction, where we explain that in practice \mathbb{F} is replaced by its empirical version \mathbb{F}_n , and mention a set of conditions presented later in the paper, that allows such plug-in estimators to be used while retaining asymptotic properties. See the top part of page 2.

Reviewer detailed comment (h): *Equation (5): is \tilde{X} a statistic?*

Author response: No. The quantity \tilde{X} is a random variable that acts as a robust counterpart of $X - \mu$ in the population, and involves the unknown parameters μ and \mathbb{F} . However, with standard plug-in estimators for μ and \mathbb{F} , we can get a sample equivalent of \tilde{X} , and the theoretical aspects of using such plug-in estimators in place of the true parameters is considered in the results that follow, see for example Lemma 3.2 or Theorem 3.3.

To make the discussion on this transparent, in page 8 of the manuscript, we have now rewritten this part to say: *As a robust surrogate for $X - \mu$, we consider the following random variable*

$$\tilde{X} = W(X, \mathbb{F})S(X; \mu) \equiv W(Z, \mathbb{F}_Z)|\Lambda^{1/2}Z|^{-1}\Gamma\Lambda^{1/2}Z.$$

In samples, the equivalent for \tilde{X} is $\hat{\tilde{X}} = W(X, \mathbb{F}_n)S(X; \hat{\mu})$ for a suitable location estimator $\hat{\mu}$, for example, the weighted spatial median.

Reviewer detailed comment (i): *On page 7, it is stated that “We now discuss the properties of the sample version $\tilde{\Sigma}$ computed from X . In practice, we cannot obtain $W(x)?W(x; F)$, and consequently use $W(x; \mathbb{F}_n)$ instead..” – this is a good point, but I suggest discussing it earlier in the paper, as it affects many assumptions made prior to this statement.*

Author response: We thank the reviewer for this excellent suggestion, this greatly helps in bringing clarity to our paper. We have now stated this right upfront in the introduction, see page 2.

Reviewer minor comment (a): *Last paragraph of page 3: should \hat{q}_n be \hat{q}_{nW} ?*

Author response: Thank you for pointing, we have now corrected this typo.

References

- Haberman, S. J. (1989). Concavity and estimation. *The Annals of Statistics*, 17(4):1631 – 1661.
- Koltchinskii, V. I. (1997). M -estimation, convexity and quantiles. *The Annals of Statistics*, 25(2):435 – 477.
- Niemiro, W. (1992). Asymptotics for M -estimators defined by convex minimization. *The Annals of Statistics*, 20(3):1514–1533.
- Wang, L., Peng, B., and Li, R. (2015). A high-dimensional nonparametric multivariate test for mean vector. *Journal of the American Statistical Association*, 110(512):1658 – 1669.