Review for 'Faithful Variable Screening for High-dimensional Convex Regression'

This paper considers the problem of variable screening in convex regression. A two-stage procedure is developed where first one-dimensional convex additive functions are found and then the residuals are fitted using one-dimensional concave functions. A quadratic program is presented and a finite sample statistical analysis is provided. Section 3 provides a background for the method of using additive approximations to convex regression functions, showing that under a boundary flatness condition, the additive approximation is faithful and motivating the two-stage approach. Section 4 shows how the problem can be expressed as a two-stage quadratic programming method. Section 5 provides a finite sample analysis when the number of relevant variables s is much smaller than n and Section 6 provides a simulation study.

I believe the idea is very interesting and potentially useful, the paper is reasonably well-written and the proofs appear to be correct. However the whole basis of the method working seems to be the boundary flatness condition and I don't believe the authors provide a good enough intuition for how strong this condition and when it is likely to be satisfied. In my opinion, neither the examples in Section 3 nor the simulations in Section 6 provide strong evidence that there are many functions will satisfy this condition, especially those with a support outside $[0, 1]^p$. It also seemed a bit strange to me that the choice of simulation functions were multivaraite Gaussian which in practice, I don't see why one would use convex regression to do variable screening when other parametric approaches seem more appropriate (e.g. clean and screen methods by Ke et al. and others). This method and others were compared with and so it again raised the question of what class of interesting examples would this method be useful for.

Hence I would recommend a major revision in which the authors provide more intuition for the boundary flatness condition and a sense of how likely the additive faithfulness condition is to be satisfied. Further simulations with non-parametric functions rather than multi-variate Gaussian would also be usefu. This should lead to an understanding of what interesting classes of functions this method would be useful for and then I beleive the paper would suitable for publication.

1 Major Comments

- The basis of all the results seem to be the boundary flatness condition. This condition seems to be an artifact of the analysis and based on my reading, it is unclear how restrctive this assumption is. The authors claim it is weak just after stating it and then present several straightforward examples in which it doesn't hold (e.g. $f(x_1, x_2) = x_1.x_2$) and multivariate Gaussian examples. Hence it is unclear to me how likely a given function is likely to fit into this framework.

- Following on from the previous point, the simulation study also does not add any further insight to this issue since the simulations are for multivariate Gaussians where other parametric approaches that deal with covariance seem more suitable. It is puzzling to me that the authors did not consider non-parametric examples where the additive model framework seems more suitable.

2 Minor Comments

- I didn't think there was any need for Section 2 since it feels very repetitive and slightly confusing that parts of the text are repeated almost verbatim. I would suggest removing this section.

- Page 2, second last paragraph 'and and.'