1 Gervini-Yohai d-variate filter

In this Section we are going to show that the filters introduced in are a special case of our approach, using the following Gervini-Yohai depth

$$d_{GY}(\mathbf{t}, \mathbf{F}, \mathbf{G}) = \mathbf{1} - \mathbf{G}(\boldsymbol{\Delta}(\mathbf{t}, \mu(\mathbf{F}), \boldsymbol{\Sigma}(\mathbf{F}))),$$

where G is a continuous distribution function, $\mu(\mathbf{F})$ and $\boldsymbol{\Sigma}(\mathbf{F})$ are the location and scatter matrix functionals and $\Delta(t, F) = \Delta(\mathbf{t}, \mu(\mathbf{F}), \boldsymbol{\Sigma}(\mathbf{F})) = (\mathbf{t} - \mu(\mathbf{F}))^\top \boldsymbol{\Sigma}(\mathbf{F})^{-\mathbf{1}}(\mathbf{t} - \mu(\mathbf{F}))$ is the squared Mahalanobis distance. Appendix shows that this is a statistical data depth function. Let $\{G_n\}_{n=1}^\infty$ be a sequence of discrete distribution functions that might depends on $\hat{F}_n$ and such that $\sup_t |G_n(t) - G(t)| \overset{a.s.}{\to} 0$, we might define the finite sample version of the Gervini-Yohai depth as

$$d_{GY}(\mathbf{t}, \hat{\mathbf{F}}_\mathbf{n}, \mathbf{G}_\mathbf{n}) = \mathbf{1} - \mathbf{G}_\mathbf{n}(\boldsymbol{\Delta}(\mathbf{t}, \mu(\hat{\mathbf{F}}_\mathbf{n}), \boldsymbol{\Sigma}(\hat{\mathbf{F}}_\mathbf{n}))) \ ,$$

however for filtering purpose we will use two alternative definitions later on. The use of G_n, that might depend on the data, instead of G makes this sample depth semiparametric. We notice that the Mahalanobis depth, which is completely parametric, cannot be used for the purpose of defining a filter in a similar fashion. Let $1 \leq d \leq p$, $j_1, \ldots, j_d$ be an d-tuple of the integer numbers $1, \ldots, p$ and, for easy of presentation, let $\mathbf{Y}_\mathbf{i} = (\mathbf{X}_{\mathbf{ij_1}}, \ldots, \mathbf{X}_{\mathbf{ij_d}})$ be a subvector of dimension d of $\mathbf{X}_\mathbf{i}$. Consider a pair of initial location and scatter estimators

$$\mathbf{T}_{\mathbf{0n}}^{(\mathbf{d})} = \begin{pmatrix} T_{0n,j_1} \\ \ldots \\ T_{0n,j_d} \end{pmatrix} \quad \text{and} \quad \mathbf{C}_{\mathbf{0n}}^{(\mathbf{d})} = \begin{pmatrix} C_{0n,j_1j_1} \ldots C_{0n,j_1j_d} \\ \ldots\ldots\ldots \\ C_{0n,j_dj_1} \ldots C_{0n,j_dj_d} \end{pmatrix} .$$

Now, define the squared Mahalanobis distance for a data point $\mathbf{Y}_\mathbf{i}$ by $\Delta_i = \Delta(\mathbf{Y}_\mathbf{i}, \hat{\mathbf{F}}_\mathbf{n}) = \boldsymbol{\Delta}(\mathbf{Y}_\mathbf{i}, \mathbf{T}_{\mathbf{0n}}^{(\mathbf{d})}, \mathbf{C}_{\mathbf{0n}}^{(\mathbf{d})})$. Consider G the distribution function of a χ^2_d, H the distribution function of $\Delta = \Delta(\cdot, F)$ and let $\hat{H}_n$ be the empirical distribution function of Δ_i $(1 \leq i \leq n)$. We consider two finite sample version of the Gervini-Yohai depth, i.e.,

$$d_{GY}(\mathbf{t}, \hat{\mathbf{F}}_\mathbf{n}, \mathbf{G}) = \mathbf{1} - \mathbf{G}(\boldsymbol{\Delta}(\mathbf{t}, \hat{\mathbf{F}}_\mathbf{n})),$$

and

$$d_{GY}(\mathbf{t}, \hat{\mathbf{F}}_\mathbf{n}, \hat{\mathbf{H}}_\mathbf{n}) = \mathbf{1} - \hat{\mathbf{H}}_\mathbf{n}(\boldsymbol{\Delta}(\mathbf{t}, \hat{\mathbf{F}}_\mathbf{n})).$$

The proportion of flagged d-variate outliers is defined by

$$d_n = \sup_{\mathbf{t} \in \mathbf{A}} \{d_{GY}(\mathbf{t}, \hat{\mathbf{F}}_\mathbf{n}, \hat{\mathbf{H}}_\mathbf{n}) - \mathbf{d}_{\mathbf{GY}}(\mathbf{t}, \hat{\mathbf{F}}_\mathbf{n}, \mathbf{G})\}^+.$$