Weighted Projection Quantiles Algorithm

October 7, 2014

Algorithm to calculate weighted projection quantile along the vector $\mathbf{u} \in \mathbb{R}^p$, given a set of observations $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n$:

1. ***Compute*** $Q_{proj}(\mathbf{u})$***, projection quantile along*** $\mathbf{u}$

   - Project each $\mathbf{X}_i$ along $\mathbf{u}$ to obtain $X_{\mathbf{u}i} = \frac{\langle \mathbf{X}_i, \mathbf{u} \rangle}{\|\mathbf{u}\|}$, for $i = 1, 2, ..., n$.
   - Find $\alpha = \frac{1+\|\mathbf{u}\|}{2}$-th quantile of $X_{\mathbf{u}1}, ..., X_{\mathbf{u}n}$, say $q_{\mathbf{u}}$..
   - $Q_{proj}(\mathbf{u}) = q_{\mathbf{u}} \mathbf{e}_{\mathbf{u}}$, $\mathbf{e}_{\mathbf{u}} = \mathbf{u}/\|\mathbf{u}\|$ being the unit vector along $\mathbf{u}$.

2. ***Compute Weights corresponding to this projection quantile*** $Q_{proj}(\mathbf{u})$

   - Compute global weights for the direction vector $\mathbf{u}$ by k-mean distance:
     - Compute k-mean distance corresponding to $Q_{proj}(\mathbf{u})$ using $d_k = \frac{1}{n} \sum_{i=1}^{n} d_i \mathbb{I}_{\{d_i < d_{(k)}\}}$, where d_i is the euclidean distace of $\mathbf{X}_i$ from $Q_{proj}(\mathbf{u})$ given by $\|\mathbf{X}_i - Q_{proj}(\mathbf{u})\|$. k is a tuning parameter.

     - Compute the weights corresponding to $\mathbf{u}$:

       $$w_{\mathbf{u}} = \exp(-a.d_k)$$

       where a is a tuning parameter.
   - Compute weights for each sample point $\mathbf{X}_i; i = 1, 2, ..., n$:
     - Compute the orthogonal Norms by $\|\mathbf{X}_{\mathbf{u}\perp i}\| = \|\mathbf{X}_i - X_{\mathbf{u}i} \mathbf{e}_{\mathbf{u}}\|$.
     - Compute weight of i^{th} sample:

       $$w_{2i} = \begin{cases} \exp\left[-b \frac{\|\mathbf{X}_{\mathbf{u}\perp i}\|}{\|\mathbf{X}_i\|}\right] & \text{if} \|\mathbf{X}_i\| \leq \epsilon \\ 0 & \text{otherwise} \end{cases}$$

       b, ϵ being tuning parameters.

3. ***Compute the weighted projection quantile***

   - Define $\tilde{X}_{\mathbf{u}i} = w_{\mathbf{u}} w_{2i} X_{\mathbf{u}i}$ for $i = 1, ..., n$.
   - Find $\alpha = \frac{1+\|\mathbf{u}\|}{2}$ th quantile of $\tilde{X}_{\mathbf{u}1}, ..., \tilde{X}_{\mathbf{u}n}$. Let it be $\tilde{q}_{\mathbf{u}}$.
   - Find the weighted projection quantile as $\tilde{Q}_{proj}(\mathbf{u}) = \tilde{q}_{\mathbf{u}} \mathbf{e}_{\mathbf{u}}$.

Algorithm for weighted projection quantile based data depth given data $\mathbf{X}_1, ..., \mathbf{X}_n$ and a new point $\mathbf{p}$

1. Compute Projection Quantiles for all n sample points along the direction of $\mathbf{p}$.

2. Find two closest points along the grid and their corresponding quantile values, say α_1 and α_2 corresponding to u_{p1} and u_{p2}.

3. Approximate the quantile value for p by interpolation $\alpha_p = \frac{(\alpha_1 - \alpha_2)(\|p\| - \|u_1\|}{\|u_1\| - \|u_2\|} + \alpha_1$.

4. Define **projection quantile depth** as $d_p = \exp(-\alpha_p)$.

Algorithm for weighted projection depth based classification scheme

1. Given data with k classes and a new observation p find the data depth for p for each class k.

2. Classify p in class j if $d_{pj} = min\{d_{pi}; i \in \{1, ..., k\}\}$