

Weighted Projection Quantiles Algorithm

October 21, 2014

Algorithm to calculate weighted projection quantile along the vector $\mathbf{u} \in \mathcal{B}_p$, given a set of observations $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$:

1. **Compute $Q_{proj}(\mathbf{u})$, the 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 $i = 1, 2, \dots, n$.
- Find $\alpha = \frac{1+\|\mathbf{u}\|}{2}$ -th quantile of $X_{\mathbf{u}1}, \dots, X_{\mathbf{u}n}$, 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 $\bar{d}_k = \frac{1}{n} \sum_{i=1}^n d_i \mathbb{I}_{\{d_i < d_{(k)}\}}$, where d_i is the euclidean distance 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 $\mathbf{X}_i; i = 1, 2, \dots, n$:
 - 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:

$$w_{2i} = \exp \left[-b \frac{\|\mathbf{X}_{\mathbf{u}\perp i}\|}{\|\mathbf{X}_i\|} \right] \mathbb{I}_{\{\|\mathbf{X}_{\mathbf{u}\perp i}\| \leq \epsilon\}}$$

b, ϵ being tuning parameters.

3. **Compute the weighted projection quantile**

- Suppose there are m observations with non-zero weights w_{2i} , with indices i_1, i_2, \dots, i_m . Define $\tilde{X}_{\mathbf{u}i_j} = w_{\mathbf{u}} w_{2i_j} X_{\mathbf{u}i_j}$.
- Find $\alpha = \frac{1+\|\mathbf{u}\|}{2}$ -th quantile of $\tilde{X}_{\mathbf{u}i_1}, \dots, \tilde{X}_{\mathbf{u}i_m}$. Let it be $\tilde{q}_{\mathbf{u}}$.
- Find the weighted projection quantile as $\tilde{Q}_{proj}(\mathbf{u}) = \tilde{q}_{\mathbf{u}}\mathbf{e}_{\mathbf{u}}$.

Definition Given a random vector $\mathbf{X} \in \mathbb{R}^p$ that follows a multivariate distribution F , and a point $\mathbf{p} \in \mathbb{R}^p$, find $\alpha_{\mathbf{p}}$ such that $\|\mathbf{p}\|$ is the $\alpha_{\mathbf{p}}$ -th quantile for the projection of \mathbf{X} on \mathbf{p} , say $X_{\mathbf{p}}$. Then the **Projection Quantile Depth** (PQD) at \mathbf{p} with respect to F is defined as

$$D(\mathbf{p}, F) = \exp(-\alpha_{\mathbf{p}})$$

Given data $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$, the PQD at a given \mathbf{p} can be estimated by finding the two nearest points on either side of $\|\mathbf{p}\|$ along \mathbf{p} , say $\mathbf{p}_1, \mathbf{p}_2$, obtain their corresponding quantiles, say α_1, α_2 respectively, then estimate $\alpha_{\mathbf{p}}$ by a linear approximation:

$$\hat{\alpha}_{\mathbf{p}} = \frac{(\alpha_1 - \alpha_2)(\|\mathbf{p}\| - \|\mathbf{p}_1\|)}{\|\mathbf{p}_1\| - \|\mathbf{p}_2\|} + \alpha_1$$

and plugging it in the above definition.

Algorithm 1 Algorithm for PQD-based classification

- 1: **procedure** PQDClassifier(training data $\mathbf{X}_i \in \mathbb{R}^{n_i \times p}$ with class labels i ; $i = 1, 2, \dots, k$, new data $\mathbf{x}_{new} \in \mathbb{R}^p$)
- 2: Set $i = 1$.
- 3: *top*:
- 4: Estimate from the sample the PQD of \mathbf{p} with respect to the i^{th} population, say $D(\mathbf{x}_{new}, \mathbf{X}_i)$.
- 5: **if** $i = k$ **then Stop**
- 6: **else**
- 7: Set $i \leftarrow i + 1$, **goto top**
- 8:
- 9: Find c that maximizes the PQD of \mathbf{x}_{new} w.r.t. all possible classes:

$$D(\mathbf{x}_{new}, \mathbf{X}_c) = \max\{D(\mathbf{x}_{new}, \mathbf{X}_i) : i = 1, 2, \dots, k\}$$

- 10: Assign class c to new data \mathbf{x}_{new} .
-

Note One can define a weighted version of PQD by replacing $X_{\mathbf{p}}$ by their weighted version $\tilde{X}_{\mathbf{p}}$. A weighted classification scheme follows similarly.

New notion of data depth

We define a data depth based on projection quantiles (which, of course, can be extended to the weighted versions), on the lines of Zuo's Projection Depth [ref?].

Definition Given a random vector $X \in \mathbb{R}^p$ that follows a multivariate distribution F , and a point $x \in \mathbb{R}^p$, define an outlyingness function

$$O(x, F) = \sup_{\|u\|=1} \left| F_u(x_u) - \frac{1}{2} \right|$$

where F_u is the distribution function of X_u , and $x_u = \langle x, u \rangle / \|u\|$. Then we define the **Projection Quantile Depth** at x with respect to F as

$$PQD(x, F) = \frac{1}{1 + O(x, F)}$$

We now follow the order in Zuo's paper to obtain the properties of the PQD.

(P1) Affine invariance Say G is the cdf of $AX + b$. Then we have $O(Ax + b, G) = O(x, F)$ because for all $u \in \mathbb{R}^p$ we have

$$F(x) = G(Ax + b) \Rightarrow F_u(x_u) = G_{Au+b}((Ax + b)_{Au+b})$$

where subscripting denotes projection.

(P2) Quasi-concavity for $0 < \lambda < 1, z = \lambda x + (1 - \lambda)y$,

$$\begin{aligned} F_u(z_u) &= F_u(\lambda x_u + (1 - \lambda)y_u) \\ \Rightarrow \left| F_u(z_u) - \frac{1}{2} \right| &\leq \begin{cases} \left| \max\{F_u(x_u), F_u(y_u)\} - \frac{1}{2} \right| & \text{if } F_u(z_u) < 1/2 \\ \left| \min\{F_u(x_u), F_u(y_u)\} - \frac{1}{2} \right| & \text{otherwise} \end{cases} \\ &\leq \max \left\{ \left| F_u(x_u) - \frac{1}{2} \right|, \left| F_u(y_u) - \frac{1}{2} \right| \right\} \end{aligned}$$

and we finally get $PQD(z, F) \geq \max\{PQD(x, F), PQD(y, F)\}$.

(P3) Vanishing at infinity For $\|x\| \rightarrow \infty$ we have $F_u(x_u) \rightarrow 1$ hence $PQD(x, F) \rightarrow 2/3$. So we can define $PQD^*(x, F) = PQD(x, F) - 2/3$ to obtain P3.

(P4) Maximized at center of symmetric F Halfspace symmetry of F w.r.t. center θ implies (halfspace) symmetry of F_u around θ_u for any $u \in \mathbb{R}^p$, so that $F_u(\theta_u) = 1/2$ and subsequently PQD achieves its max value 1 at θ .

Example 1 For $X \sim \mathcal{N}_2((0, 0)', I_2)$ we have $X_u \sim N(0, 1)$ for any $u \in \mathbb{R}^p : \|u\| = 1$. Due to the symmetry of the distribution function maximizing $|F_u(x_u) - 1/2|$ is equivalent to maximizing $F_u(x_u) = \Phi(x_u) = \Phi(\langle x, u \rangle)$.

Converting u to polar coordinates we can maximize over the angle θ that u makes with the X-axis, and can easily find that the above function maximizes at $\theta = \tan^{-1}(x_2/x_1)$, which is the direction for x itself. Thus for given x the sup is obtained at $u = x$, and we have

$$PQD(x, F) = \frac{1}{1 + \Phi(\|x\|) - 1/2} = \frac{2}{1 + 2\Phi(\|x\|)} \quad (1)$$

Fig. 1 shows a comparison of PQD with Zuo's PD for standard bivariate normal.

Note In place of standard bivariate normal, one can use any circularly symmetric F , i.e. a distribution that has the same marginal distribution F_0 along all one-dimensional projections to come up with an exact formula for $PQD(x, F)$. In that general scenario, $\Phi(\cdot)$ will be replaced by $F_0(\cdot)$ in (1).