Robust Multivariate Regression

D. Busarova[1]

[1] Moscow State University, Leninskiye Gory, 119899, Moscow, Russia

Keywords: Multivariate Regression, Oja median.

Consider the multivariate regression model

$$Y = XB + \varepsilon,$$

where Y is a $n \times q$ matrix of responses, X is a $n \times p$ matrix of predictors, and ε is a $n \times q$ matrix of random errors. We will write Y_i and X_i for the transposed ith rows of Y and X, respectively. If the model includes an intercept, then the first coordinate of X_i is 1. The ith observation is then the pair (Y_i, X_i). We assume, that both Y_i and X_i are random. The problem we solve is then that of affine equivariant estimating the unknown $p \times q$ matrix B of regression coefficients.

The estimates we propose are based on elemental sets and elemental regressions. Each elemental set is a subset $I = \{i_1, i_2, \ldots, i_p\}$ of size p of the original data. We will write $Y(I)$ for the $p \times q$ submatrix $(Y_{i_1}, Y_{i_2}, \ldots, Y_{i_p})^T$ of Y and write $X(I)$ for the $p \times p$ submatrix $(X_{i_1}, X_{i_2}, \ldots, X_{i_p})^T$ of X. If $\mathrm{rank}(X(I)) = p$, then elemental regression $B(I)$ is defined as $B(I) = X(I)^{-1}Y(I)$. Denote then $b(I) = \mathrm{vec}(B(I))$. If $\mathrm{rank}(X(I)) < p$, then elemental regression $b(I)$ is called degenerate.

Our approach is based on the Oja criterion function. We first calculate all vectorized nondegenerate elemental regressions. Then we construct the objective functions $U_n(\beta)$ and $D_n(\beta)$ for the Oja median and the weighted Oja median of data set $\{b(I)\}$, respectively. The choices of β to minimize $U_n(\beta)$ and $D_n(\beta)$ are then our estimates $\hat{\beta}_n^1$ and $\hat{\beta}_n^2$ of the vector $\mathrm{vec}(B)$.

We investigate the consistency and the asymptotic normality of both proposed estimates. Also we show that this multivariate regression estimates have the appropriate equivariance properties and bounded influence functions.