Semi-Mixture Regression Model for Incomplete Data

Given a mixture regression model,

Where,

Note, *ck* is the prior probability that any *zi* belongs to the *k*th component. Let *Y* be random variable representing components, *Y* = 1, 2,…, *K*. We have:

At the *t*th iteration, the E-step aims to estimate missing values *xij* and *zi* for each partial distribution *Pk*(*zi* | *αk*, *σk*2) as follows:

Where,

At the *t*th iteration, the M-step aims to estimate the next parameter Θ(*t*+1) as follows:

Where,

In the E-step, for each *k*th component, it estimates missing values *xij* and *zi* according to one-iteration balance process.

In the M-step, for each *k*th component, we have:

Where,

Note, *μk*(*t*+1) is biased estimate of variance. The optimization problem of *αk*(*t*+1) estimation is solved by Lagrange duality.