For the parametric estimation without correlation, 3 probability distributions are considered, namely gaussian, lognormal and weibull. The network consists of 3 o-d pairs, x1, x2, x3 and 5 links, the flows on which is defined by the following mapping function:

fun\_map = [exp(x1 + x2 + 1), (x1-1)^2, (x2-1)^2, (x3-1)^2, exp(x2+x3+1)]

For normal and lognormal distributions, following parameter values are assumed:

par\_loc = 1

par\_scale = 0.2, 0.3, 0.4

For Weibull distribution, following parameter values were used:

par\_loc = 0.5 (shape parameter)

par\_scale = 1.3, 1.35, 1.5 (scale parameter)

Thus, for each distribution, total 3 simulations were run for 50 samples each, keeping ‘gradobj’ to ‘off’. Finally, for each distribution, in total 6 plots were generated by fitting normal curves to the estimated parameter values. 6 plots are correspond to x1, x2, x3 location parameters and scale parameters, and in each plot, 3 lines correspond to 3 different scale parameters.

For each distribution, the samples are generated using uniform distribution, whereas observed values are generated using the respective distribution. Number of observations generated is sample are 1e5, but for each iteration, randomly 1e4 observations the sampled for moment matching. For observed data, total 250 samples are considered (equivalent to working days is a year, I think) with 200 observations each. For all the estimation, moment bound is considered to be 4.

As the fitted curves are normal, they are supposed to be bell curved with peak at the mean, i.e., corresponding to whatever value that curve is plotted for. The not-so-good shape of the curves could be dedicated to fewer samples used for estimation, hopefully increasing the number of iterations to 200 samples, will make the curves look better and hence, more informative.

The problem with the above method was the singularity of the gmmweight matrix, which was solved by only taking its diagonal terms for estimation. All the above plots were recreated for gaussian distribution case with moment bound of 2 and 4.

Next, we moved onto testing multimodal distributions.

GM Distributions

Initially we started with using only 2 dimensions of X and 6 dimensions of y. The mapping was as follows:

Fun\_map = [(x(:,1)-1).^2, (x(:,2)-1).^2,...

(x(:,1)+1).^2.\*(x(:,2)+1).^2, (x(:,1).\*x(:,2)),...

x(:,1).^1.5+x(:,2).^1.5,...

x(:,1)./(1e-3+x(:,2))]

The mean and standard deviation parameters for modes of both the dimensions were as follows:

Dimension 1:

Mode 1: mean = 0.08, std = 0.01; Mode 2: mean = 1.2, std = 0.01

Dimension 2:

Mode 1: mean = 0.08, std = 0.02; Mode 2: mean = 1.2, std = 0.02

Next, the sample specifications are similar to above, number of observation samples = 200, number of instances in each sample = 250, num\_smp = 1e5, num\_smp\_iter = 1e4. Moment bound used was 4. In addition to this, another constraint was introduced that for each dimension, the 2nd mode appears after the first mode.