

For the parametric estimation without correlation, 3 probability distributions are considered, namely gaussian, lognormal and weibull. The network consists of 3 o-d pairs,  $x_1$ ,  $x_2$ ,  $x_3$  and 5 links, the flows on which is defined by the following mapping function:

$\text{fun\_map} = [\exp(x_1 + x_2 + 1), (x_1 - 1)^2, (x_2 - 1)^2, (x_3 - 1)^2, \exp(x_2 + x_3 + 1)]$

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

$\text{par\_loc} = 1$

$\text{par\_scale} = 0.2, 0.3, 0.4$

For Weibull distribution, following parameter values were used:

$\text{par\_loc} = 0.5$  (shape parameter)

$\text{par\_scale} = 1.3, 1.35, 1.5$  (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 correspond to  $x_1$ ,  $x_2$ ,  $x_3$  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 are sampled for moment matching. For observed data, total 250 samples are considered (equivalent to working days in 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.