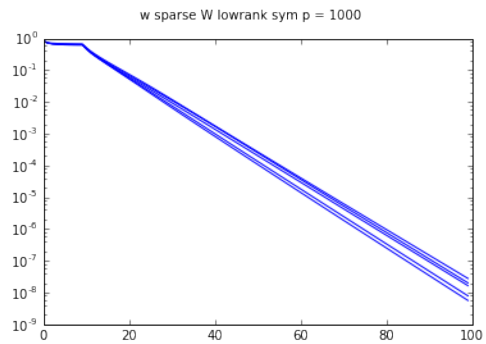
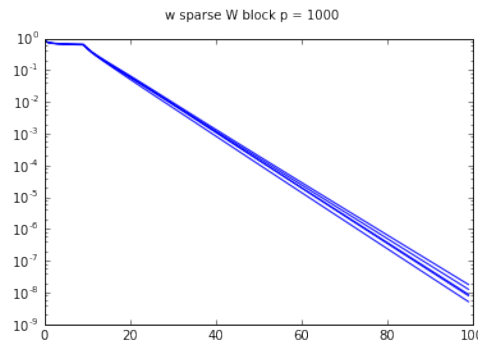
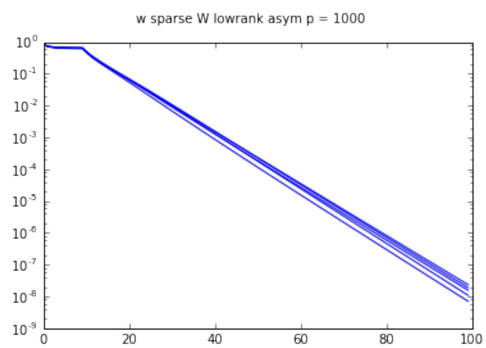


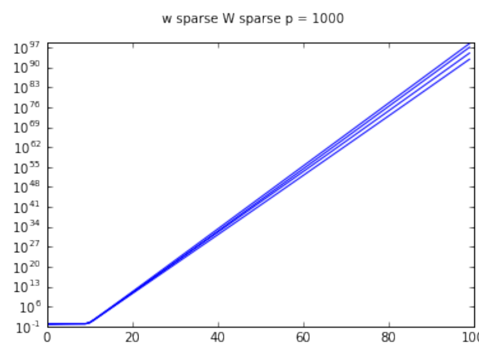
(a) w sparse W low rank and symmetric



(b) w sparse W block



(c) w sparse W low rank and asymmetric



(d) w sparse W sparse

Figure 1: caption

1 generalized Factorization Machine

We observed that when gFM is dealing with W that has a sparse structure, the error rate on the training data explodes, probably because its updating through the sequence of estimation does not necessarily require the error change in a descent manner.

Similarly, choosing the rank k of the training data as close to the original rank k in is important. As it determines the initialization of U in the algorithm described, estimating k by a large number would almost certainly cause the error rate explodes. The updating rule used in this algorithm is highly sensitive to apriori knowledge about the structure though it can't be easily estimated without certain domain knowledges.

The learning rate, as used to describe the algorithm of gFM, differs from the learning rate we know from algorithms we familiar with from the first order or second order algorithms. Instead, they use learning rate here to denote the degree a single estimation can learn. We tried setting up different learning rate and found that within a certain small window of learning rate, the error rate can decrease linearly with a considerable rate, while with a slightly larger learning rate, linearly increased error rate can be observed.

Generally, generalized factorization machine and its associated algorithm bridges the theoretical gap between the matrix sensing problem and the factorization machine. The problem it deals with is naturally non-convex which means the conventional gradient descent framework can hardly achieve a global convergence. It is proved that the algorithm they proposed can achieve a linear convergence rate and it is also shown in our experiment. However, we found it to be tricky to fine tune the parameters such as learning rate and the rank k before applying this algorithm, which makes it somewhat impractical in a real world setting.