

Manual for Running the Scripts

Determination of Nonlinear Genetic Architecture using Compressed Sensing

Chiu Man Ho and Stephen D. H. Hsu

Block-Diagonal (BD) Model:

First run `BD_Generate_A_Randn_5.m`, `BD_Generate_A_Randn_50.m` and `BD_Generate_A_Randn_100.m` to generate the synthetic genotypes and noises for $s = 5$, 50 and 100 respectively.

These synthetic genotypes and noises will be loaded into `BD_5.m`, `BD_50.m` and `BD_100.m` respectively to produce the results of the paper.

Then, one can run `BD_plot_figures.m` to make the figures for the BD model.

Promiscuous (PS) Model:

First run `PS_Generate_A_Randn_32.m`, `PS_Generate_A_Randn_3020.m` and `PS_Generate_A_Randn_6040.m` to generate the synthetic genotypes and noises for $(s = 3, s' = 2)$, $(s = 30, s' = 20)$ and $(s = 60, s' = 40)$ respectively.

These synthetic genotypes and noises will be loaded into `PS_32.m`, `PS_3020.m` and `PS_6040.m` respectively to produce the results of the paper.

Then, one can run `PS_plot_figures.m` to make the figures for the BD model.

Real Genome:

We only perform this study in the BD model with $s = 5$ and the PS model with $(s = 3, s' = 2)$.

For the comparison of real data with synthetic data in the BD model, one can run `Data_BD_5.m` and then `Data_BD_plot_figures.m` for the results.

For the comparison of real data with synthetic data in the PS model, one can run `Data_PS_32.m` and then `Data_PS_plot_figures.m` for the results.

Of course, these running require the presence of real genome data.

Continuous Case:

We only perform this study in the PS model with ($s = 3$, $s' = 2$).

One can first run Continue_Generate_A_Randn_32.m to generate the continuous data. Then, run Continue_32.m and Continue_PS_plot_figures.m for the results.

Lasso and p-value

Lasso_sv.m is executed for running LASSO, and lse.m is used for computing the p-values of the genetic markers that have non-zero support. These two scripts should be present in all the studies.

Explanatory Comments

Since the scripts for generating the main results are similar, we only add explanatory comments to one of them. If one understands the explanatory comments in **PS_6040.m**, he/she will understand what's going in each step and hence other similar scripts.

Figure 2 and Figure 3

These are generated by using PS_32_y_vs_A.m and PS_32_Pvalue.m respectively.

Table 1

This can be generated by running Table.m