**PRSlmm Documentation**

*PRSlmm* is a follow up script for runPRS. It performes a linear mixed effects model to assess the variance explained of a PRS on related individuals. It receives the following four positional arguments:

* prsfile PRS file: FID IID S1 S2 ....
* phenofile Phenotypes file: FID IID Age Sex PC1 ... PCX ... Pheno1 Pheno2 Pheno3 ... PhenoK
* covdescript Describe the variables on phenofile: ColumnName [covar|qcovar]
* jobname Job Name

Additionally the following optional flags are accepted:

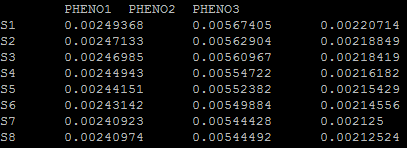
* -h, Print this help message and exit
* -r9, Use release9 grm (default is release8) (there has been problems with release 9 GRM)
* -std , This flag causes the phenotype to be centered around zero with a standard deviation of 1 (secondary output)
* -grm Use a specific (not default) GRM (useful for diagplus in R9 or other datasets such as HCP or UKB) must be full path!
* -grmext default =’gz’ you must set it if using a specific grm that is not in gz format
* -noimput, Do not include imputation covariates automatically (use if they are already in your phenofile or you are not working with genepi GWAS releases)
* -noPC Do not include PC covariates automatically (use if they are already in your phenofile or you would like more than 5 or you are not working with genepi GWAS releases)
* -noStdID Do not standardize IDs to length of 7 by adding zeros before (it us usually done automatically)
* -sexage Include the sex \* age interaction as a qcovariate (the phenofile must contain sex and age columns capslock insensitive)
* -agesqrd Include age squared as a covariate (the phenofile must contain age column capslock insensitive)
* -agesqrdsex Include age squared times sex as a covariate
* -namrk All values after this flag (use this flag at the end) will be interpreted as NA in phenotypes and covariates (e.g. -namrk -99 -9 . would interpret -99, -9 and . as NA in the phenotype)

This script performs the following:

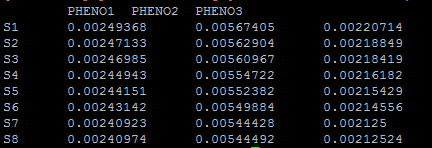
1. Opens the PRSfile and the phenofile and matches them by IID (note, it should be IID not iid nor id. This is the standard output of runPRS)
2. Ensures IDs are 7 digits (adds 0 before)
3. Opens covdescript to interpret the phenofile:
   1. Detects covars
   2. Detects qcvoars
      1. Generates interaction qcovars such as age2 and age\*sex
   3. Everything else is a phenotype
4. For every phenotype and PRS cutoff a GCTA script is made on the *lmm\_calc* directory
5. This scripts are submitted and wait for completion
6. Merges all the results into different output files:
   1. PRS\_Fixed\_eff – Fixed effect (coefficient)
   2. PRS\_SE – Standard error of fixed effect
   3. PRS\_R2 – R2
   4. PRS\_Pval- Pvalue of fixed effect
7. Saves the results

**Output Example**

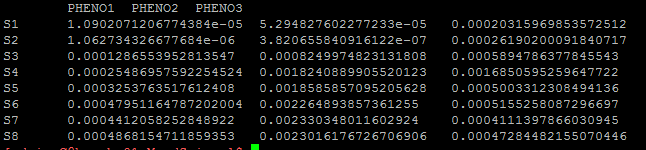
* Fixed Effects:



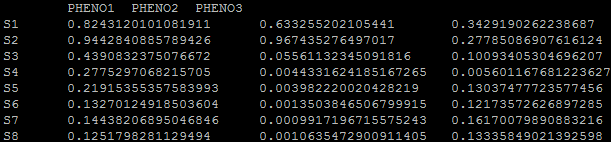
* Standard Errors



* R squared:



* Pvalues:



This format is quite convenient for plotting and assessing variance explained. In fact all the files have the same starting name (jobName) followed by the data eg jobName+Feff would be the fixed effects file. For a sample script og how to plot this results (either as bars or heatmaps) using python please refer to the script plotPRS provided (it is designed to receive this output without transformations).