

Last update: 7/01/2014

MMAP Covariance Structures

MMAP has options to compute pedigree-based, genomic and environmental covariance matrices as well as user-defined covariance structures. These covariance structures represent random effects in the mixed model. These matrices can be used in variance component estimation and other genetic analyses such as GWA. Interactions of the random effects can be modeled using the pointwise (Hadamard product) of individual matrices. See documentation on genomic matrices for options of working with genetic data.

MMAP requires the first 5 fields of the pedigree file to be pedigree id, ego, father id, mother id and sex, although the token used for each is not used. All other fields, if any, are ignored.

MMAP requires ancestral order of the pedigree for the gametic and kinship matrix calculations and will exit if the not properly ordered. The file *pedigree.err* contains the list of subjects that are out of order and their location in the pedigree file and can be used for manual correction of the pedigree file. Automated correction may be added in the future.

Pedigree

--compute_binary_relationship_matrix_by_groups

Required

--ped <pedigree>

Pedigree csv input filename

--binary_output_filename <file>

Output filename for the covariance matrix.

--group_size <num>

Controls the subpedigree size to control the memory. This generally is set to the number of subjects in the file unless the pedigree size is above several hundred thousand.

--single_pedigree

Treats all pedigrees as a single pedigree. The covariance file will have zero values for subjects in different pedigrees, but the option is **HIGHLY RECOMMENDED** if the relationship matrix is to be combined with any other covariance matrices that have between-pedigree covariance values such as genomic matrices. MMAP uses the grouping variable in the pedigree file, so for analysis with the

--xlinked_relationship_matrix

X linked relationship matrix

Example Commands

Autosome

```
mmap --compute_binary_relationship_matrix_by_groups --ped Amish.ped.csv --  
binary_output_filename Amish.autosome.relationship.bin --group_size 25000 --  
single_pedigree
```

X chromosome

```
mmap --compute_binary_relationship_matrix_by_groups --ped Amish.ped.csv --
binary_output_filename Amish.X.relationship.bin --group_size 25000 --single_pedigree
--xlinked_relationship_matrix
```

Pedigree Common Environment and Cytoplasmic inheritance

MMAP can construct 0-1 matrices to model common environment and cytoplasmic inheritance.

--y_chromosome_covariance

Traces Y lineage.

--mitochondrial_covariance

Traces maternal lineage

--full_sib_covariance

Groups full sibs

--spouse_covariance

Groups spouses

--nuclear_family_covariance

Groups nuclear families within a pedigree

Each option REQUIRES

--ped <pedigree>

--binary_output_filename <file>

Name of output file

--pedigree_output_filename <file>

Contains the pedigree plus an extra column with numerical value for each group

Permanent Environment or Group Covariance Matrices

MMAP can construct a random effect based on a grouping variable such as sex, study, plate, generation, etc, from a categorical or numerical variable in a csv file. The matrix is 0-1 with 1 in the (J,K) entry if subject J and subject K are in the same group, 0 otherwise. Missing data is treated as an independent group, that is, with a 1 for that individual on the diagonal.

--compute_group_covariance_file

--phentotype_id <ID>

Required. The header token that identifies the subject ID

--group_variable <string>

Required. The header token that identifies the grouping variable

--csv_input_filename <file>

Required. The csv input file with subjects and grouping variable

--binary_output_filename <file>

Required. Binary covariance filename.

Custom Covariance Matrices

MMAP can import user generated matrices in two formats to facilitate integration of output from 3rd party software that for example compute IBD matrices or omics-type matrices such as normalized gene expression. These options assume txt not csv files since txt is more common output format from other programs.

For files formatted IDA IDB <val> use the following option. **This option assumes upper triangular order as the default.** That is if there are N subjects in the file, then IDA is the first subject for lines 1 to N, the second subject for lines N+1 to N-1, etc, with IDB varying for each

line. Thus the file is expected to have $N*(N+1)/2$ lines plus 1 for the header, which can be skipped. The file is assumed to have a header.

--pairwise_value_txt2mmap

--txt_input_filename <file>

Required.

--binary_output_filename <file>

Required.

--no_header

Optional. Add if the file has no header line.

--idB

Optional. Add if the IDB column is constant while IDA varies.

For rectangular NxN files with the first line containing the subject ids use the following option

--square_matrix_txt2mmap

--txt_input_filename <file>

Required.

--binary_output_filename <file>

Required.

Extracting Values from Matrix

There are two options to extract the pairwise values from a binary covariance matrix and an option to extract the full matrix. Subject set options are being added to be able extract specified subjects.

--variance_component_matrix_mmap2pairs

Output each pair once. The order depends on the order of the subjects in the file. The output is the upper diagonal of the matrix

--variance_component_matrix_mmap2pairs_all

Output each pair of different subjects twice with opposite orders. The output is the full matrix

--variance_component_matrix_mmap2matrix

Rectangular output with the subject ids as the first row.

Both options require:

--binary_input_filename <file> --csv_output_filename <file>

Interaction

MMAP can model interactions between random effects using the pointwise product of two covariance matrices. MMAP assumes the matrices have the same subjects in each matrix.

--hadamard_product_two_covariance_matrices

--variance_component_filename <file1> <file2>

The two covariance matrices to be multiplied

--binary_output_filename <file>

The product matrix filename