Matrix Epistasis Report (with EpiGEN)

Ways of evaluating epistasis detection algorithm performance:

1. Number of interacting pairs (true positives) identified.
2. Number of false positives.
3. Position of true positive(s) (averaged if multiple) when ranked by p-value. If any true positives are missing their position is considered to be the “last” in the list.
   1. The above definition will be referred to as penalized true positive position.
   2. Alternatively, average true positive position can be computed only using the detected true positives. This will be referred to as unpenalized true positive position.
4. Runtime (~1second for 1000 SNPs, > 15 minutes for 100000 SNPs)
   1. This is an R package. Was executing multiple hours before I interrupted it for 100000 SNPs, there may be some way to speed it up.

Compare these performance parameters across:

1. Multiplicative vs Dominant vs Recessive vs XOR Interaction.
2. Interaction Alpha: 1.25, 1.5, 2, 3, 8, 16.
3. Pure vs Impure.

Each EpiGEN dataset is named as follows: PURE/IMPURE (indicates presence of marginal effects) \_ MULTIPLICATIVE/DOMINANT/RECESSIVE/XOR (indicates type of interaction) \_ ONEPAIR / TWOPAIRS / EIGHTPAIRS (indicates number of interacting disease snp pairs) \_ BASELINEALPHA10 (indicates quantitative phenotype value in absence of marginal and interaction effects) \_ InteractionAlpha 1.25/1.5/2/3/8/16 (indicates strength of interaction effect)

**RESULTS (pages 1-11)**

**MULTIPLICATIVE DATASETS**

Pure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 2

Number of false positives: 779

F1 Score: 0.0051

Average true positive position (unpenalized) ranked by p-value: 324

Average true positive position (penalized) ranked by p-value: 667.5

Pure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 1

Number of false positives: 1061

F1 Score: 0.0019

Average true positive position (unpenalized) ranked by p-value: 6

Average true positive position (penalized) ranked by p-value: 930.875

Impure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 0

Number of false positives: 778

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 0

Number of false positives: 954

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 0

Number of false positives: 25

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 1

Number of false positives: 84

F1 Score: 0.023

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 1

Impure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha15

Number of true positives: 0

Number of false positives: 166

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha125

Number of true positives: 0

Number of false positives: 49

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 1

Number of false positives: 548

F1 Score: 0.0036

Average true positive position (unpenalized) ranked by p-value: 31

Average true positive position (penalized) ranked by p-value: 290.5

Impure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 0

Number of false positives: 177

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha15

Number of true positives: 0

Number of false positives: 206

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha125

Number of true positives: 0

Number of false positives: 29

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Pure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 1

Number of false positives: 147

F1 Score: 0.013

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 1

Pure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 1

Number of false positives: 18

F1 Score: 0.1

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 1

Pure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha15

Number of true positives: 0

Number of false positives: 26

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Pure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha125

Number of true positives: 0

Number of false positives: 0

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Pure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 1

Number of false positives: 113

F1 Score: 0.017

Average true positive position (unpenalized) ranked by p-value: 4

Average true positive position (penalized) ranked by p-value: 59.5

Pure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 1

Number of false positives: 542

F1 Score: 0.0037

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 272.5

Pure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha15

Number of true positives: 1

Number of false positives: 8

F1 Score: 0.18

Average true positive position (unpenalized) ranked by p-value: 3

Average true positive position (penalized) ranked by p-value: 6.5

Pure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha125

Number of true positives: 0

Number of false positives: 2

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

**DOMINANT DATASETS**

Impure\_Dominant\_OnePair\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 1

Number of false positives: 265

F1 Score: 0.0075

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 1

Impure\_Dominant\_OnePair\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 547

F1 Score: 0.0036

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 1

Impure\_Dominant\_TwoPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 2

Number of false positives: 223

F1 Score: 0.018

Average true positive position (unpenalized) ranked by p-value: 1.5

Average true positive position (penalized) ranked by p-value: 1.5

Impure\_Dominant\_TwoPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 2

Number of false positives: 245

F1 Score: 0.016

Average true positive position (unpenalized) ranked by p-value: 1.5

Average true positive position (penalized) ranked by p-value: 1.5

Pure\_Dominant\_OnePair\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 1

Number of false positives: 313

F1 Score: 0.0063

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 1

Pure\_Dominant\_OnePair\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 320

F1 Score: 0.0062

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 1

Pure\_Dominant\_TwoPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 2

Number of false positives: 291

F1 Score: 0.013

Average true positive position (unpenalized) ranked by p-value: 1.5

Average true positive position (penalized) ranked by p-value: 1.5

Pure\_Dominant\_TwoPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 2

Number of false positives: 380

F1 Score: 0.010

Average true positive position (unpenalized) ranked by p-value: 1.5

Average true positive position (penalized) ranked by p-value: 1.5

**RECESSIVE DATASETS**

Impure\_Recessive\_OnePair\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 0

Number of false positives: 0

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Recessive\_OnePair\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 0

Number of false positives: 3

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Recessive\_TwoPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 1

Number of false positives: 1473

F1 Score: 0.0013

Average true positive position (unpenalized) ranked by p-value: 57

Average true positive position (penalized) ranked by p-value: 766

Impure\_Recessive\_TwoPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 1424

F1 Score: 0.0014

Average true positive position (unpenalized) ranked by p-value: 43

Average true positive position (penalized) ranked by p-value: 734.5

Pure\_Recessive\_OnePair\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 1

Number of false positives: 1782

F1 Score: 0.0011

Average true positive position (unpenalized) ranked by p-value: 192

Average true positive position (penalized) ranked by p-value: 192

Pure\_Recessive\_OnePair\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 1698

F1 Score: 0.0012

Average true positive position (unpenalized) ranked by p-value: 149

Average true positive position (penalized) ranked by p-value: 149

Pure\_Recessive\_TwoPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 1

Number of false positives: 565

F1 Score: 0.0035

Average true positive position (unpenalized) ranked by p-value: 100

Average true positive position (penalized) ranked by p-value: 333.5

Pure\_Recessive\_TwoPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 607

F1 Score: 0.0033

Average true positive position (unpenalized) ranked by p-value: 9

Average true positive position (penalized) ranked by p-value: 309

**XOR DATASETS:**

Pure\_XOR\_EightPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 0

Number of false positives: 277

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Pure\_XOR\_EightPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 0

Number of false positives: 1030

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_XOR\_EightPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 0

Number of false positives: 1215

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_XOR\_EightPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 0

Number of false positives: 775

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

**DISCUSSION (Page 12-13)**

Unpenalized true positive positions by interaction type.

Multiplicative: 324,6,0,0,0,1,0,0,31,0,0,0,1,1,0,0,4,1,3,0

Dominant: 1,1,1.5,1.5,1,1,1.5,1.5

Recessive: 0,0,57,43,192,149,100,9

XOR: 0,0,0,0

F1 Score by grouped by interaction type.

Multiplicative: 0.0051, 0.0019, 0.023, 0.0036, 0.013, 0.1, 0.017, 0.0037, 0.18, 0,0,0,0,0,0,0,0,0,0,0 => 0.017

Dominant: 0.0075, 0.0036, 0.018, 0.016, 0.0063, 0.0062, 0.013, 0.01 => 0.01

Recessive: 0,0,0.0013,0.0014, 0.0011, 0.0012, 0.0035, 0.0033 => 0.0015

XOR: 0,0,0,0 => 0

From this point onward assume all true positive positions are penalized.

Multiplicative Interaction Average True Positive Positions (note that 0 means no true positives were detected): 667.5,930.875,0,0,0,1,0,0,290.5,0,0,0,1,1,0,0,59.5,272.5,6.5,0

Dominant Interaction Average True Positive Positions: 1,1,1.5,1.5,1,1,1.5,1.5

Recessive Interaction Average True Positive Positions: 0,0,766,734.5,192,149,333.5,309

XOR Interaction Average True Positive Positions: 0,0,0,0

Notably, Matrix Epistasis performance is identical to that of Plink’s epistasis, since both seem to be based in logistic regression. Since, tools like Cassi and CAPE are also based on logistic regression model they should return identical results.

Pairs Detected (# of SNP pairs correctly detected / total # of interacting disease SNP pairs) – only count if at least one pair was detected: 2/8, 1/8, 1/1, ½, 1/1, 1/1, ½, ½, ½, 1/1, 1/1, 2/2, 2/2, 1/1, 1/1, 2/2, 2/2, ½, ½, 1/1, 1/1, 1/2, ½

Now let us examine when Matrix Epistasis could detect multiplicative interaction.

Multiplicative Interaction Average True Positive Position Clustered by Interaction Alpha:

Interaction Alpha 1.25: 0, 0, 0, 0

Interaction Alpha 1.5: 0, 0, 0, 6.5

Interaction Alpha 2.0: 930.875, 0, 0, 290.5, 1, 59.5

Interaction Alpha 3.0: 667.5, 0, 1, 0, 1, 272.5

Clear trend here that interactions with larger effect were more easily detectable. These results are also identical to that of Plink’s Epistasis.

Multiplicative Interaction Average True Positive Position Clustered by Pure Status:

Pure: 667.5, 930.875, 1, 1, 0, 0, 59.5, 272.5, 6.5, 0

Impure: 0, 0, 0, 1, 0, 0, 290.5, 0, 0, 0

Pure epistasis is easier to detect as expected.

Notably, all multiplicative pure interactions with alpha >= 2 were detected.