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**

Impure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 0

Number of false positives: 20

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: 72

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: 153

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: 44

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: 523

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

Average true positive position (penalized) ranked by p-value: 278

Impure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 0

Number of false positives: 163

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: 182

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: 25

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: 136

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: 12

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: 12

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

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: 97

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

Average true positive position (penalized) ranked by p-value: 51.5

Pure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 1

Number of false positives: 506

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

Average true positive position (penalized) ranked by p-value: 254.5

Pure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha15

Number of true positives: 1

Number of false positives: 8

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: 1

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

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

Pure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 2

Number of false positives: 738

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

Average true positive position (penalized) ranked by p-value: 636.75

Pure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 1

Number of false positives: 1004

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

Average true positive position (penalized) ranked by p-value: 881

Impure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 0

Number of false positives: 710

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: 889

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: 244

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: 520

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: 204

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: 227

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: 290

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: 288

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: 266

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: 245

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

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

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: 1333

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

Average true positive position (penalized) ranked by p-value: 696

Impure\_Recessive\_TwoPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 1320

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

Average true positive position (penalized) ranked by p-value: 682.5

Pure\_Recessive\_OnePair\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 1

Number of false positives: 1656

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: 1566

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: 537

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

Average true positive position (penalized) ranked by p-value: 319.5

Pure\_Recessive\_TwoPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 542

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

Average true positive position (penalized) ranked by p-value: 276.5

**CONTROL DATASETS**

Pure\_Control\_TwoPairs\_BaselineAlpha10

Number of true positives: 0

Number of false positives: 0

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

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

Pure\_Control\_OnePair\_BaselineAlpha10

Number of true positives: 0

Number of false positives: 0

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

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

Pure\_Control\_EightPairs\_BaselineAlpha10

Number of true positives: 0

Number of false positives: 0

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

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

Impure\_Control\_TwoPairs\_BaselineAlpha10

Number of true positives: 0

Number of false positives: 99

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

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

Impure\_Control\_OnePair\_BaselineAlpha10

Number of true positives: 0

Number of false positives: 30

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

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

Impure\_Control\_EightPairs\_BaselineAlpha10

Number of true positives: 0

Number of false positives: 1

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

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

**XOR DATASETS:**

Pure\_XOR\_EightPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 0

Number of false positives: 245

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: 969

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: 1130

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: 731

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

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

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

Control: 0, 0, 0, 0, 0, 0

XOR: 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): 0,1, 0, 0, 278, 0, 0, 0, 1, 1, 0, 0, 51.5, 254.5, 6.5, 0, 636.75, 881, 0, 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, 696, 682. 5, 192, 149, 319.5, 276.5

Control: 0, 0, 0, 0, 0, 0

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: 1/1, ½, 1/1, 1/1, ½, ½, ½, 2/8, 1/8, 1/1, 1/1, 2/2, 2/2, 1/1, 1/1, 2/2, 2/2, ½, ½, 1/1, 1/1, ½, 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: 0, 278, 1, 51.5, 881, 0

Interaction Alpha 3.0: 1, 0, 1, 254.5, 636.75, 0

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: 1, 1, 0, 0, 51.5, 254.5, 6.5, 0, 636.75, 881

Impure: 0, 1, 0, 0, 278, 0, 0, 0, 0, 0

Pure epistasis is easier to detect as expected.

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