Reference Manual for Set-Valued System Identification (SVSI) Analysis

Wenjian Bi and Guolian Kang

November 13, 2014

1 Overview

Set-Valued System Identification analysis attempts to test genetic associations with complex phenotypes such as ordered categorical phenotype [1] or binary phenotype [2] based on set-valued model and using a novel Set-Valued System Identification algorithm. Basically, the set-valued model models the classification/dichotomization process of a continuous complicated unobserved or unobervalble phenotype and models the relationship between categorical outcome and genetic variants. Set-Valued System Identification (SVSI) algorithm is a new algorithm based on expectation-maximization (EM) which has been proposed and appears in a wide range of applications such as sensor networks and telecommunications.

In this document, we describe how to perform SVSI analysis for ordered categorical phenotype and binary phenotype using example data sets provided with the package.

2 Installation

The latest version of SVSI package can be download from
`http://www.stjuderesearch.org/site/depts/biostats/software`.

```
> # Install packages
> # Load the package
> library(SVSI)
```

3 Testing associations between SNPs and an ordered categorical phenotype

3.1 Example dataset

SVSI package provides an example dataset (`data_cate_geno`) which has a genotype matrix of 697 individuals and 20 SNPs among which the last 10 SNPs are generated from the alternative hypothesis, two binary and one continuous covariates (`data_cate_cova`).

```
> # Install packages
> # Load the package
> library(SVSI)
> # Load the example data set
> data(data_cate_geno) #ordered categorical outcome data and genotype data
> data(data_cate_cova) #covariate data for categorical outcome data
> names(data_cate_geno)#colnames of data_cate_geno
```

```
 [1] "FID"    "IID"    "PAT"    "MAT"    "SEX"    "status" "SNP1"   "SNP2"
 [9] "SNP3"   "SNP4"   "SNP5"   "SNP6"   "SNP7"   "SNP8"   "SNP9"   "SNP10"
[17] "SNP11"  "SNP12"  "SNP13"  "SNP14"  "SNP15"  "SNP16"  "SNP17"  "SNP18"
[25] "SNP19"  "SNP20"
```

```
> names(data_cate_cova)#colnames of data_cate_cova
```

```
[1] "FID"   "IID"   "SEX"   "AGE"   "SMOKE"
```

```
> summary(data_cate_geno)#summary of data_cate_geno
```

```
     FID                 IID         PAT         MAT         SEX
 Length:697         Min.   :1   Min.   :0   Min.   :0   Min.   :1.000
 Class :character   1st Qu.:1   1st Qu.:0   1st Qu.:0   1st Qu.:1.000
 Mode  :character   Median :1   Median :0   Median :0   Median :2.000
                    Mean   :1   Mean   :0   Mean   :0   Mean   :1.531
                    3rd Qu.:1   3rd Qu.:0   3rd Qu.:0   3rd Qu.:2.000
                    Max.   :1   Max.   :0   Max.   :0   Max.   :2.000
     status            SNP1                SNP2              SNP3
 Min.   :0.0000   Min.   :0.000000   Min.   :0.00000   Min.   :0.00000
 1st Qu.:0.0000   1st Qu.:0.000000   1st Qu.:0.00000   1st Qu.:0.00000
 Median :0.0000   Median :0.000000   Median :0.00000   Median :0.00000
 Mean   :0.4878   Mean   :0.007174   Mean   :0.02869   Mean   :0.01004
 3rd Qu.:1.0000   3rd Qu.:0.000000   3rd Qu.:0.00000   3rd Qu.:0.00000
 Max.   :2.0000   Max.   :1.000000   Max.   :2.00000   Max.   :1.00000
      SNP4               SNP5               SNP6              SNP7
 Min.   :0.000000   Min.   :0.000000   Min.   :0.00000   Min.   :0.000000
 1st Qu.:0.000000   1st Qu.:0.000000   1st Qu.:0.00000   1st Qu.:0.000000
 Median :0.000000   Median :0.000000   Median :0.00000   Median :0.000000
 Mean   :0.002869   Mean   :0.004304   Mean   :0.01435   Mean   :0.008608
 3rd Qu.:0.000000   3rd Qu.:0.000000   3rd Qu.:0.00000   3rd Qu.:0.000000
 Max.   :1.000000   Max.   :1.000000   Max.   :2.00000   Max.   :2.000000
      SNP8               SNP9              SNP10              SNP11
 Min.   :0.000000   Min.   :0.000000   Min.   :0.000000   Min.   :0.00000
 1st Qu.:0.000000   1st Qu.:0.000000   1st Qu.:0.000000   1st Qu.:0.00000
 Median :0.000000   Median :0.000000   Median :0.000000   Median :0.00000
 Mean   :0.002869   Mean   :0.007174   Mean   :0.007174   Mean   :0.01291
 3rd Qu.:0.000000   3rd Qu.:0.000000   3rd Qu.:0.000000   3rd Qu.:0.00000
 Max.   :1.000000   Max.   :2.000000   Max.   :1.000000   Max.   :1.00000
     SNP12            SNP13              SNP14             SNP15
 Min.   :0.0000   Min.   :0.000000   Min.   :0.00000   Min.   :0.00000
 1st Qu.:0.0000   1st Qu.:0.000000   1st Qu.:0.00000   1st Qu.:0.00000
 Median :0.0000   Median :0.000000   Median :0.00000   Median :0.00000
 Mean   :0.2912   Mean   :0.002869   Mean   :0.03443   Mean   :0.01004
 3rd Qu.:1.0000   3rd Qu.:0.000000   3rd Qu.:0.00000   3rd Qu.:0.00000
 Max.   :2.0000   Max.   :1.000000   Max.   :2.00000   Max.   :2.00000
     SNP16              SNP17          SNP18             SNP19
 Min.   :0.000000   Min.   :0.000   Min.   :0.00000   Min.   :0.000000
 1st Qu.:0.000000   1st Qu.:0.000   1st Qu.:0.00000   1st Qu.:0.000000
 Median :0.000000   Median :0.000   Median :0.00000   Median :0.000000
 Mean   :0.002869   Mean   :0.066   Mean   :0.01722   Mean   :0.005739
 3rd Qu.:0.000000   3rd Qu.:0.000   3rd Qu.:0.00000   3rd Qu.:0.000000
```

```
 Max.   :1.000000   Max.   :2.000   Max.   :1.00000   Max.   :1.000000  
     SNP20         
 Min.   :0.000000  
 1st Qu.:0.000000  
 Median :0.000000  
 Mean   :0.002869  
 3rd Qu.:0.000000  
 Max.   :1.000000  
```

```
> summary(data_cate_cova)#summary of data_cate_cova
```

```
     FID                 IID          SEX             AGE       
 Length:697         Min.   :1   Min.   :1.000   Min.   :16.00  
 Class :character   1st Qu.:1   1st Qu.:1.000   1st Qu.:25.00  
 Mode  :character   Median :1   Median :2.000   Median :40.00  
                    Mean   :1   Mean   :1.531   Mean   :41.79  
                    3rd Qu.:1   3rd Qu.:2.000   3rd Qu.:52.00  
                    Max.   :1   Max.   :2.000   Max.   :91.00  
     SMOKE       
 Min.   :0.0000  
 1st Qu.:0.0000  
 Median :0.0000  
 Mean   :0.2597  
 3rd Qu.:1.0000  
 Max.   :1.0000  
```

```
> print(data_cate_geno[1:3,]) #display the first three rows of data_cate_geno
```

```
  FID IID PAT MAT SEX status SNP1 SNP2 SNP3 SNP4 SNP5 SNP6 SNP7 SNP8 SNP9 SNP10
1 ID1   1   0   0   1      0    0    1    0    0    0    0    0    0    1     0
2 ID2   1   0   0   2      2    0    0    0    0    0    0    0    0    0     0
3 ID3   1   0   0   2      0    0    0    0    0    0    0    0    0    0     0
  SNP11 SNP12 SNP13 SNP14 SNP15 SNP16 SNP17 SNP18 SNP19 SNP20
1     0     0     0     0     0     0     0     0     0     0
2     0     0     0     0     0     0     0     0     0     0
3     0     1     0     0     0     0     0     0     0     0
```

```
> print(data_cate_cova[1:3,]) #display the first three rows of data_cate_cova
```

```
  FID IID SEX AGE SMOKE
1 ID1   1   1  82     0
2 ID2   1   2  86     0
3 ID3   1   2  84     1
```

3.2 Genetic association test with adjusting for covariates using SVSI

Conduct genetic association testing on the input data (`data_cate_geno`) with adjusting for covariate (`data_cate_cova`):

```
> out=SVSI_OrderCate_Whole(data_cate_geno,data_cate_cova,prefix="",result.dir=NA)
> out[,1:6]
```

```
      [,1]    [,2]       [,3]        [,4]       [,5]       [,6]
      ""      "Variable" "Estimate"  "SE"       "Stat"     "Pval"
SNP1  "SNP1"  "geno"     " 0.735100" "0.50340"  "2.133"    "0.1442"
SNP2  "SNP2"  "geno"     "-0.765600" "0.350800" "4.7640"   "0.02907"
SNP3  "SNP3"  "geno"     "-2.657000" "3.179000" "0.6984"   "0.4033"
SNP4  "SNP4"  "geno"     "-2.641000" "5.948000" "0.1972"   "0.6570"
SNP5  "SNP5"  "geno"     "-2.791000" "4.814000" "0.3360"   "0.5621"
SNP6  "SNP6"  "geno"     " 0.378300" "0.304000" "1.549"    "0.2133"
SNP7  "SNP7"  "geno"     "-0.242700" "0.467000" "0.2700"   "0.6033"
SNP8  "SNP8"  "geno"     " 0.085240" "0.841500" "0.01026"  "0.9193"
SNP9  "SNP9"  "geno"     "-1.846000" "3.078000" "0.3599"   "0.5486"
SNP10 "SNP10" "geno"     "-2.667000" "3.75300"  "0.5050"   "0.4773"
SNP11 "SNP11" "geno"     " 0.971000" "0.378000" "6.5980"   "0.01021"
SNP12 "SNP12" "geno"     " 0.007681" "0.090120" "0.007265" "0.9321"
SNP13 "SNP13" "geno"     "-2.658000" "5.940000" "0.2003"   "0.6545"
SNP14 "SNP14" "geno"     "-0.305000" "0.260700" "1.3690"   "0.2421"
SNP15 "SNP15" "geno"     " 0.307800" "0.31300"  "0.9671"   "0.3254"
SNP16 "SNP16" "geno"     "-2.809000" "5.919000" "0.2252"   "0.6351"
SNP17 "SNP17" "geno"     "-0.110500" "0.17620"  "0.3936"   "0.5304"
SNP18 "SNP18" "geno"     " 0.913900" "0.328300" "7.749"    "0.005373"
SNP19 "SNP19" "geno"     " 0.080910" "0.596100" "0.01842"  "0.8920"
SNP20 "SNP20" "geno"     " 0.860600" "0.792800" "1.1780"   "0.2777"
```

3.3 Genetic association test without adjusting for covariates using SVSI

Conduct genetic association testing on the input data (`data_cate_geno`) without adjusting for any covariate:

```
> out=SVSI_OrderCate_Whole(data_cate_geno,data_cate_cova=NA)
> out
```

```
      [,1]    [,2]       [,3]       [,4]      [,5]       [,6]
      ""      "Variable" "Estimate" "SE"      "Stat"     "Pval"
SNP1  "SNP1"  "geno"     "0.7858"   "0.5015"  "2.455"    "0.1172"
SNP2  "SNP2"  "geno"     "-0.7582"  "0.3467"  "4.781"    "0.02877"
SNP3  "SNP3"  "geno"     "-2.782"   "3.423"   "0.6606"   "0.4164"
SNP4  "SNP4"  "geno"     "-2.775"   "6.404"   "0.1877"   "0.6648"
SNP5  "SNP5"  "geno"     "-2.776"   "5.229"   "0.2818"   "0.5955"
SNP6  "SNP6"  "geno"     "0.3518"   "0.3031"  "1.348"    "0.2457"
SNP7  "SNP7"  "geno"     "-0.2108"  "0.458"   "0.2119"   "0.6453"
SNP8  "SNP8"  "geno"     "0.06611"  "0.8391"  "0.006207" "0.9372"
SNP9  "SNP9"  "geno"     "-1.896"   "3.188"   "0.3535"   "0.5521"
SNP10 "SNP10" "geno"     "-2.779"   "4.05"    "0.4708"   "0.4926"
SNP11 "SNP11" "geno"     "0.9571"   "0.3761"  "6.476"    "0.01093"
SNP12 "SNP12" "geno"     "0.007383" "0.08981" "0.006758" "0.9345"
SNP13 "SNP13" "geno"     "-2.775"   "6.404"   "0.1877"   "0.6648"
SNP14 "SNP14" "geno"     "-0.3143"  "0.2597"  "1.465"    "0.2261"
SNP15 "SNP15" "geno"     "0.3065"   "0.3123"  "0.9638"   "0.3262"
SNP16 "SNP16" "geno"     "-2.775"   "6.404"   "0.1877"   "0.6648"
SNP17 "SNP17" "geno"     "-0.1025"  "0.1754"  "0.3411"   "0.5592"
SNP18 "SNP18" "geno"     "0.9155"   "0.3261"  "7.884"    "0.004986"
```

```
SNP19 "SNP19" "geno"     "0.06632"  "0.5942"  "0.01245"  "0.9111"
SNP20 "SNP20" "geno"     "0.7809"   "0.791"   "0.9746"   "0.3235"
```

4 Testing associations between SNPs and a binary phenotype

4.1 Example dataset

SVSI package provides an example dataset (`data_bina_geno`) which has a genotype matrix of 697 individuals and 20 SNPs among which the last 2 SNPs are generated from the alternative hypothesis, two binary and one continuous covariates (`data_bina_cova`).

```
> # Install packages
> # Load the package
> library(SVSI)
> # Load the example data set
> data(data_bina_geno) #ordered binagorical outcome data and genotype data
> data(data_bina_cova) #covariate data for binagorical outcome data
> names(data_bina_geno)#colname of data_bina_geno
```

```
 [1] "FID"    "IID"    "PAT"    "MAT"    "SEX"    "status" "SNP1"   "SNP2"  
 [9] "SNP3"   "SNP4"   "SNP5"   "SNP6"   "SNP7"   "SNP8"   "SNP9"   "SNP10" 
[17] "SNP11"  "SNP12"  "SNP13"  "SNP14"  "SNP15"  "SNP16"  "SNP17"  "SNP18" 
[25] "SNP19"  "SNP20" 
```

```
> names(data_bina_cova)#colname of data_bina_cova
```

```
[1] "FID"   "IID"   "SEX"   "AGE"   "SMOKE"
```

```
> summary(data_bina_geno)#summary of data_bina_geno
```

```
     FID                 IID         PAT         MAT         SEX       
 Length:697         Min.   :1   Min.   :0   Min.   :0   Min.   :1.000  
 Class :character   1st Qu.:1   1st Qu.:0   1st Qu.:0   1st Qu.:1.000  
 Mode  :character   Median :1   Median :0   Median :0   Median :2.000  
                    Mean   :1   Mean   :0   Mean   :0   Mean   :1.531  
                    3rd Qu.:1   3rd Qu.:0   3rd Qu.:0   3rd Qu.:2.000  
                    Max.   :1   Max.   :0   Max.   :0   Max.   :2.000  
     status            SNP1               SNP2              SNP3        
 Min.   :0.0000   Min.   :0.000000   Min.   :0.00000   Min.   :0.00000  
 1st Qu.:0.0000   1st Qu.:0.000000   1st Qu.:0.00000   1st Qu.:0.00000  
 Median :0.0000   Median :0.000000   Median :0.00000   Median :0.00000  
 Mean   :0.2999   Mean   :0.007174   Mean   :0.02869   Mean   :0.01004  
 3rd Qu.:1.0000   3rd Qu.:0.000000   3rd Qu.:0.00000   3rd Qu.:0.00000  
 Max.   :1.0000   Max.   :1.000000   Max.   :2.00000   Max.   :1.00000  
      SNP4               SNP5               SNP6              SNP7         
 Min.   :0.000000   Min.   :0.000000   Min.   :0.00000   Min.   :0.000000  
 1st Qu.:0.000000   1st Qu.:0.000000   1st Qu.:0.00000   1st Qu.:0.000000  
 Median :0.000000   Median :0.000000   Median :0.00000   Median :0.000000  
 Mean   :0.002869   Mean   :0.004304   Mean   :0.01435   Mean   :0.008608  
 3rd Qu.:0.000000   3rd Qu.:0.000000   3rd Qu.:0.00000   3rd Qu.:0.000000  
 Max.   :1.000000   Max.   :1.000000   Max.   :2.00000   Max.   :2.000000  
```

```
      SNP8               SNP9             SNP10              SNP11         
 Min.   :0.000000   Min.   :0.00000   Min.   :0.000000   Min.   :0.000000  
 1st Qu.:0.000000   1st Qu.:0.00000   1st Qu.:0.000000   1st Qu.:0.000000  
 Median :0.000000   Median :0.00000   Median :0.000000   Median :0.000000  
 Mean   :0.002869   Mean   :0.04304   Mean   :0.002869   Mean   :0.002869  
 3rd Qu.:0.000000   3rd Qu.:0.00000   3rd Qu.:0.000000   3rd Qu.:0.000000  
 Max.   :1.000000   Max.   :2.00000   Max.   :1.000000   Max.   :1.000000  
     SNP12             SNP13              SNP14             SNP15         
 Min.   :0.00000   Min.   :0.000000   Min.   :0.00000   Min.   :0.000000  
 1st Qu.:0.00000   1st Qu.:0.000000   1st Qu.:0.00000   1st Qu.:0.000000  
 Median :0.00000   Median :0.000000   Median :0.00000   Median :0.000000  
 Mean   :0.01578   Mean   :0.004304   Mean   :0.05882   Mean   :0.002869  
 3rd Qu.:0.00000   3rd Qu.:0.000000   3rd Qu.:0.00000   3rd Qu.:0.000000  
 Max.   :1.00000   Max.   :1.000000   Max.   :2.00000   Max.   :1.000000  
     SNP16            SNP17              SNP18              SNP19       
 Min.   :0.0000   Min.   :0.00000   Min.   :0.000000   Min.   :0.0000  
 1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.000000   1st Qu.:0.0000  
 Median :0.0000   Median :0.00000   Median :0.000000   Median :0.0000  
 Mean   :0.3831   Mean   :0.02152   Mean   :0.005739   Mean   :0.1334  
 3rd Qu.:1.0000   3rd Qu.:0.00000   3rd Qu.:0.000000   3rd Qu.:0.0000  
 Max.   :2.0000   Max.   :1.00000   Max.   :1.000000   Max.   :1.0000  
     SNP20        
 Min.   :0.00000  
 1st Qu.:0.00000  
 Median :0.00000  
 Mean   :0.05595  
 3rd Qu.:0.00000  
 Max.   :1.00000  
```

```
> summary(data_bina_cova)#summary of data_bina_cova
```

```
     FID                 IID         SEX              AGE       
 Length:697         Min.   :1   Min.   :0.0000   Min.   :16.00  
 Class :character   1st Qu.:1   1st Qu.:0.0000   1st Qu.:25.00  
 Mode  :character   Median :1   Median :1.0000   Median :40.00  
                    Mean   :1   Mean   :0.5308   Mean   :41.79  
                    3rd Qu.:1   3rd Qu.:1.0000   3rd Qu.:52.00  
                    Max.   :1   Max.   :1.0000   Max.   :91.00  
     SMOKE       
 Min.   :0.0000  
 1st Qu.:0.0000  
 Median :0.0000  
 Mean   :0.2425  
 3rd Qu.:0.0000  
 Max.   :1.0000  
```

```
> print(data_bina_geno[1:3,]) #display the first three rows of data_bina_geno
```

```
  FID IID PAT MAT SEX status SNP1 SNP2 SNP3 SNP4 SNP5 SNP6 SNP7 SNP8 SNP9 SNP10
1 ID1   1   0   0   1      1    0    1    0    0    0    0    0    0    0     0
2 ID2   1   0   0   2      1    0    0    0    0    0    0    0    0    0     0
3 ID3   1   0   0   2      0    0    0    0    0    0    0    0    0    0     0
```

```
  SNP11 SNP12 SNP13 SNP14 SNP15 SNP16 SNP17 SNP18 SNP19 SNP20
1     0     0     0     1     0     1     0     0     0     0
2     0     0     0     0     0     1     0     0     0     0
3     0     0     0     0     0     0     0     0     0     0
```

```
> print(data_bina_cova[1:3,]) #display the first three rows of data_bina_cova
```

```
  FID IID SEX AGE SMOKE
1 ID1   1   0  82     0
2 ID2   1   1  86     0
3 ID3   1   1  84     0
```

4.2 Genetic association test with adjusting for covariates using SVSI

Conduct genetic association testing on the input data (`data_bina_geno`) with adjusting for covariate (`data_bina_cova`):

```
> out=SVSI_Bina_Whole(data_bina_geno,data_bina_cova)
> out[,1:6]
```

```
      [,1]    [,2]       [,3]       [,4]      [,5]      [,6]
      ""      "Variable" "Estimate" "SE"      "Stat"    "Pval"
SNP1  "SNP1"  "geno"     "-1.077"   "0.6357"  "2.87"    "0.09027"
SNP2  "SNP2"  "geno"     "0.4505"   "0.3391"  "1.765"   "0.184"
SNP3  "SNP3"  "geno"     "0.1673"   "0.5342"  "0.09815" "0.7541"
SNP4  "SNP4"  "geno"     "-0.187"   "0.9583"  "0.03809" "0.8453"
SNP5  "SNP5"  "geno"     "-0.2635"  "0.8461"  "0.097"   "0.7555"
SNP6  "SNP6"  "geno"     "0.06807"  "0.3769"  "0.03262" "0.8567"
SNP7  "SNP7"  "geno"     "0.3829"   "0.5616"  "0.4648"  "0.4954"
SNP8  "SNP8"  "geno"     "3.469"    "8.931"   "0.1509"  "0.6977"
SNP9  "SNP9"  "geno"     "0.3005"   "0.2652"  "1.284"   "0.2571"
SNP10 "SNP10" "geno"     "2.812"    "9.129"   "0.09487" "0.7581"
SNP11 "SNP11" "geno"     "2.998"    "9.002"   "0.1109"  "0.7391"
SNP12 "SNP12" "geno"     "0.2289"   "0.4629"  "0.2445"  "0.621"
SNP13 "SNP13" "geno"     "2.123"    "7.446"   "0.0813"  "0.7755"
SNP14 "SNP14" "geno"     "0.04715"  "0.2184"  "0.04662" "0.8291"
SNP15 "SNP15" "geno"     "-0.5882"  "0.916"   "0.4124"  "0.5208"
SNP16 "SNP16" "geno"     "0.06238"  "0.09372" "0.443"   "0.5057"
SNP17 "SNP17" "geno"     "0.3464"   "0.4094"  "0.7158"  "0.3975"
SNP18 "SNP18" "geno"     "-0.3562"  "0.6647"  "0.2872"  "0.592"
SNP19 "SNP19" "geno"     "-0.6177"  "0.1555"  "15.78"   "7.125e-05"
SNP20 "SNP20" "geno"     "-0.9028"  "0.2253"  "16.05"   "6.158e-05"
```

4.3 Genetic association test without adjusting for covariates using SVSI

Conduct genetic association testing on the input data (`data_bina_geno`) without adjusting for covariate:

```
> out=SVSI_Bina_Whole(data_bina_geno,data_bina_cova=NA)
> out
```

```
      [,1]    [,2]       [,3]         [,4]      [,5]        [,6]
      ""      "Variable" "Estimate"   "SE"      "Stat"      "Pval"
```

```
SNP1  "SNP1"  "geno"     "-0.7844"    "0.5693"  "1.898"     "0.1682"
SNP2  "SNP2"  "geno"     "0.1221"     "0.2719"  "0.2016"    "0.6534"
SNP3  "SNP3"  "geno"     "-0.3486"    "0.4791"  "0.5292"    "0.4669"
SNP4  "SNP4"  "geno"     "-0.5265"    "0.8876"  "0.3518"    "0.5531"
SNP5  "SNP5"  "geno"     "-0.0945"    "0.7502"  "0.01587"   "0.8998"
SNP6  "SNP6"  "geno"     "-0.0002912" "0.3541"  "6.762e-07" "0.9993"
SNP7  "SNP7"  "geno"     "0.3542"     "0.5474"  "0.4186"    "0.5176"
SNP8  "SNP8"  "geno"     "2.785"      "9.131"   "0.09305"   "0.7603"
SNP9  "SNP9"  "geno"     "-0.0003642" "0.2237"  "2.65e-06"  "0.9987"
SNP10 "SNP10" "geno"     "2.785"      "9.131"   "0.09305"   "0.7603"
SNP11 "SNP11" "geno"     "2.785"      "9.131"   "0.09305"   "0.7603"
SNP12 "SNP12" "geno"     "-0.179"     "0.3896"  "0.2111"    "0.6459"
SNP13 "SNP13" "geno"     "2.787"      "7.456"   "0.1397"    "0.7086"
SNP14 "SNP14" "geno"     "-0.04427"   "0.1956"  "0.05123"   "0.8209"
SNP15 "SNP15" "geno"     "-0.5265"    "0.8876"  "0.3518"    "0.5531"
SNP16 "SNP16" "geno"     "-0.02228"   "0.08365" "0.07096"   "0.7899"
SNP17 "SNP17" "geno"     "0.1002"     "0.3511"  "0.08145"   "0.7753"
SNP18 "SNP18" "geno"     "-0.5281"    "0.6287"  "0.7058"    "0.4008"
SNP19 "SNP19" "geno"     "-0.4073"    "0.1415"  "8.284"     "0.003999"
SNP20 "SNP20" "geno"     "-0.7318"    "0.2082"  "12.35"     "0.0004401"
```

5 Contact Information

If you have any technique questions about using the SVSI package, please email us at: guolian.kang@stjude.org or biwenjian@amss.ac.cn or zhifa.liu@stjude.org.

References

[1] Wenjian Bi, Guolian Kang, Yanlong Zhao, Yuehua Cui, Song Yan, Yun Li, Cheng Cheng, Stanley B. Pounds, Michael J. Borowitz, Mary V. Relling, Jun J. Yang, Ching-Hon Pui, Stephen P. Hunger, Christine M Hartford, Wing Leung, Ji-Feng Zhang. (2014). SVSI: Fast and Powerful Set-Valued System Identification Approach to Identifying Rare Variants in Sequencing Studies for Ordered Categorical Traits. *Annals of Human Genetics* **Accepted with a major revision**.

[2] Guolian Kang, Wenjian Bi, Yanlong Zhao, Ji-Feng Zhang, Jun J Yang, Heng Xu, Mignon L Loh, Stephen P Hunger, Mary V Relling, Stanley Pounds, Cheng Cheng. (2014). A New System Identification Approach to Identify Genetic Variants in Sequencing Studies for a Binary Phenotype. *Human Heredity* **78**:104–116.