

```
In [1]: include("/home/nicole/Jupyter/SSBRnoJ/src/SSBR.jl")
        using SSBR
```

```
In [2]: function getPos(ped,IDs)
        posAi = Array{Int64,size(IDs,1)}
        for (i,id) = enumerate(IDs[:,1])
            posAi[i] = ped.idMap[id].seqID
        end
        return posAi
    end
```

```
Out[2]: getPos (generic function with 1 method)
```

```
In [3]: ; cd Data/0.5a0/G/7

        /home/nicole/Jupyter/JG3/Data/0.5a0/G/7
```

```
In [4]: ;ls
        G5.noGenotype.ID
        GenNF.txt
        PedAll.txt
        Phe.txt
        PheAll.txt
        Regression.G5.G.J.txt
        Regression.G5.G.JC.txt
        all.ID
        alphaEstimates
        genotype.ID
        meanOfSNPGA11
        meanOfSNPGG0
        meanOfSNPGG1
        meanOfSNPGG2
        meanOfSNPGG3
        meanOfSNPGG4
        meanOfSNPGG5
        noGenotype.ID
        sim.bv
        sim.phenotype
```

```
In [5]: ;awk '{print $1}' PedAll.txt | sort -b > all.ID
```

```
In [6]: ;awk '{print $1}' GenNF.txt | sort -b > genotype.ID
```

```
In [7]: ;join -v1 all.ID genotype.ID > noGenotype.ID
```

```
In [8]: ;awk '{print $1,$2}' Phe.txt > sim.phenotype
```

```
In [9]: ;awk '{print $1,$3}' PheAll.txt > sim.bv
```

```
In [10]: ; awk 'NR >=1 && NR <=8000 {print $1}' PedAll.txt | sort -b > G0.ID
```

```
In [11]: ; awk 'NR >=8001 && NR <=16000 {print $1}' PedAll.txt | sort -b > G1.ID
```

```
In [12]: ; awk 'NR >=16001 && NR <=24000 {print $1}' PedAll.txt | sort -b > G2.ID
```

```
In [13]: ; awk 'NR >=24001 && NR <=32000 {print $1}' PedAll.txt | sort -b > G3.ID
```

```
In [14]: ; awk 'NR >=32001 && NR <=40000 {print $1}' PedAll.txt | sort -b > G4.ID
```

```
In [15]: ; awk 'NR >=40001 && NR <=48000 {print $1}' PedAll.txt | sort -b > G5.ID
```

```
In [16]: ;join G0.ID genotype.ID > G0.Genotype.ID
```

```
In [17]: ;join G1.ID genotype.ID > G1.Genotype.ID
```

```
In [18]: ;join G2.ID genotype.ID > G2.Genotype.ID
```

```
In [19]: ;join G3.ID genotype.ID > G3.Genotype.ID
```

```
In [20]: ;join G4.ID genotype.ID > G4.Genotype.ID
```

```
In [21]: ;join G5.ID genotype.ID > G5.Genotype.ID
```

```
In [22]: ;join -v1 G0.ID genotype.ID > G0.noGenotype.ID
```

```
In [23]: ;join -v1 G1.ID genotype.ID > G1.noGenotype.ID
```

```
In [24]: ;join -v1 G2.ID genotype.ID > G2.noGenotype.ID
```

```
In [25]: ;join -v1 G3.ID genotype.ID > G3.noGenotype.ID
```

```
In [26]: ;join -v1 G4.ID genotype.ID > G4.noGenotype.ID
```

```
In [27]: ;join -v1 G5.ID genotype.ID > G5.noGenotype.ID
```

```
In [28]: ;wc G0.Genotype.ID;wc G1.Genotype.ID;wc G2.Genotype.ID;wc G3.Genotype.ID;wc G4.Genotype.ID;wc G5.Genotype.ID
200 200 1200 G0.Genotype.ID
200 200 1200 G1.Genotype.ID
200 200 1200 G2.Genotype.ID
200 200 1200 G3.Genotype.ID
200 200 1200 G4.Genotype.ID
8000 8000 48000 G5.Genotype.ID
```

```
In [29]: ;wc G0.noGenotype.ID;wc G1.noGenotype.ID;wc G2.noGenotype.ID;wc G3.noGenotype
7800 7800 46800 G0.noGenotype.ID
7800 7800 46800 G1.noGenotype.ID
7800 7800 46800 G2.noGenotype.ID
7800 7800 46800 G3.noGenotype.ID
7800 7800 46800 G4.noGenotype.ID
0 0 0 G5.noGenotype.ID
```

```
In [30]: ped,A_Mats,numSSBayes = calc_Ai("PedAll.txt","genotype.ID",calculateInbreeding
nothing
df      = read_genotypes("GenNF.txt",numSSBayes)
M_Mats = make_MMats(df,A_Mats,ped,center=true);           # with
y_Vecs = make_yVecs("sim.phenotype",ped,numSSBayes);
Z_Mats = make_ZMats(ped,y_Vecs,numSSBayes)
X_Mats, W_Mats = make_XWMats(Z_Mats,M_Mats,numSSBayes)    # no
nothing
```

```
In [31]: vRes      = 0.738
vG       = 0.738
nIter    = 50000
@time aHat1,alphaHat,betaHat,epsHat =
ssGibbs(M_Mats,y_Vecs,Z_Mats,X_Mats,W_Mats,A_Mats, numSSBayes,vRes,vG,nIter,
nothing

This is iteration 5000
This is iteration 10000
This is iteration 15000
This is iteration 20000
This is iteration 25000
This is iteration 30000
This is iteration 35000
This is iteration 40000
This is iteration 45000
This is iteration 50000
2525.739572 seconds (23.04 G allocations: 723.480 GB, 7.19% gc time)
```

```
In [32]: betaHat
```

```
Out[32]: 1-element Array{Float64,1}:
2.47785
```

```
In [33]: using DataFrames
```

```
In [34]: df = readtable("sim.bv", eltypes=[UTF8String, Float64], separator = ' ',head
a = Array{Float64,numSSBayes.num_ped}
for (i,ID) in enumerate(df[:,1])
    j = ped.idMap[ID].seqID
    a[j] = df[i,2]
end
```

```
In [35]: IDs = readtable("all.ID", eltypes =[UTF8String], separator = ' ',header=false)
posAi = getPos(ped,IDs)
cor1 = cor(a[posAi],aHat1[posAi])[1,1]
reg1 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - all.ID : correlation = %6.3f\n", cor1 ) # with
@printf("SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = %6.3f\n",
JCA11 = cor1
```

```
SSBRJC from Gibbs - all.ID : correlation = 0.920
SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.977
```

```
Out[35]: 0.9195631158733599
```

```
In [36]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[36]: 0.1535232081517967
```

```
In [37]: IDs = readtable("genotype.ID", eltypes =[UTF8String], separator = ' ',header=
posAi = getPos(ped,IDs)
cor2 = cor(a[posAi],aHat1[posAi])[1,1]
reg2 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - genotype.ID : correlation = %6.3f\n", cor2 ) #
@printf("SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = %6.3f\n",
JCA11 = cor2
```

```
SSBRJC from Gibbs - genotype.ID : correlation = 0.978
SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 1.070
```

```
Out[37]: 0.9778942299406397
```

```
In [38]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[38]: 1.443143207817809
```

```
In [39]: IDs = readtable("noGenotype.ID", eltypes =[UTF8String], separator = ' ',header=
posAi = getPos(ped,IDs)
cor3 = cor(a[posAi],aHat1[posAi])[1,1]
reg3 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - noGenotype.ID : correlation = %6.3f\n", cor3 ) ;
@printf("SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = %6.3f\n",
JCA11 = cor3
```

```
SSBRJC from Gibbs - noGenotype.ID : correlation = 0.881
SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.963
```

```
Out[39]: 0.881005315201903
```

```
In [40]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[40]: -0.14408140715574463
```

```
In [41]: IDs = readtable("G0.ID", eltypes=[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor4 = cor(a[posAi],aHat1[posAi])[1,1]
reg4 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G0.ID : correlation = %6.3f\n", cor4 ) # with e
@printf("SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = %6.3f\n", re
JCall = cor4
```

```
SSBRJC from Gibbs - G0.ID : correlation = 0.720
SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = 0.883
```

```
Out[41]: 0.7202608307901012
```

```
In [42]: GEBV = aHat1[posAi]
G1GEBV=mean(GEBV)
```

```
Out[42]: -1.318630015658899
```

```
In [43]: IDs = readtable("G1.ID", eltypes=[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor4 = cor(a[posAi],aHat1[posAi])[1,1]
reg4 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G1.ID : correlation = %6.3f\n", cor4 ) # with e
@printf("SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV = %6.3f\n", re
JCall = cor4
```

```
SSBRJC from Gibbs - G1.ID : correlation = 0.772
SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV = 0.973
```

```
Out[43]: 0.7721915193469753
```

```
In [44]: GEBV = aHat1[posAi]
G1GEBV=mean(GEBV)
```

```
Out[44]: -0.6591388607826953
```

```
In [45]: IDs = readtable("G2.ID", eltypes=[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor5 = cor(a[posAi],aHat1[posAi])[1,1]
reg5 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G2.ID : correlation = %6.3f\n", cor5 ) # with e
@printf("SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = %6.3f\n", re
JCall = cor5
```

```
SSBRJC from Gibbs - G2.ID : correlation = 0.772
SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = 1.006
```

```
Out[45]: 0.7717442595368251
```

```
In [46]: GEBV = aHat1[posAi]
G2GEBV=mean(GEBV)
```

```
Out[46]: -0.07516717307834041
```

```
In [47]: IDs = readtable("G3.ID", eltypes=[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor6 = cor(a[posAi],aHat1[posAi])[1,1]
reg6 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G3.ID : correlation = %6.3f\n", cor6 ) # with e
@printf("SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = %6.3f\n", r
JCA11 = cor6
```

```
SSBRJC from Gibbs - G3.ID : correlation = 0.774
SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = 1.010
```

```
Out[47]: 0.7744359883377436
```

```
In [48]: GEBV = aHat1[posAi]
G3GEBV=mean(GEBV)
```

```
Out[48]: 0.4872341540840679
```

```
In [49]: IDs = readtable("G4.ID", eltypes=[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor7 = cor(a[posAi],aHat1[posAi])[1,1]
reg7 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G4.ID : correlation = %6.3f\n", cor7 ) # with e
@printf("SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = %6.3f\n", r
JCA11 = cor7
```

```
SSBRJC from Gibbs - G4.ID : correlation = 0.769
SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = 1.007
```

```
Out[49]: 0.769256469583214
```

```
In [50]: GEBV = aHat1[posAi]
G4GEBV=mean(GEBV)
```

```
Out[50]: 0.9916698893451885
```

```
In [51]: IDs = readtable("G5.ID", eltypes=[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor8 = cor(a[posAi],aHat1[posAi])[1,1]
reg8 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G5.ID : correlation = %6.3f\n", cor8 ) # with e
@printf("SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = %6.3f\n", r
JCA11 = cor8
```

```
SSBRJC from Gibbs - G5.ID : correlation = 0.975
SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = 1.066
```

```
Out[51]: 0.9747204818800306
```

```
In [52]: GEBV = aHat1[posAi]
G5GEBV=mean(GEBV)
```

```
Out[52]: 1.4951712550014589
```

```
In [53]: IDs = readtable("G0.Genotype.ID", eltypes =[UTF8String], separator = ' ', head=10)
posAi = getPos(ped,IDs)
cor9 = cor(a[posAi],aHat1[posAi])[1,1]
reg9 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G0.Genotype.ID : correlation = %6.3f\n", cor9 )
@printf("SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg9 )
JCall = cor9
```

```
SSBRJC from Gibbs - G0.Genotype.ID : correlation = 0.974
```

```
SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 1.063
```

```
Out[53]: 0.9738721824580475
```

```
In [54]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[54]: 0.041354963943827726
```

```
In [55]: IDs = readtable("G1.Genotype.ID", eltypes =[UTF8String], separator = ' ', head=10)
posAi = getPos(ped,IDs)
cor9 = cor(a[posAi],aHat1[posAi])[1,1]
reg9 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G1.Genotype.ID : correlation = %6.3f\n", cor9 )
@printf("SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg9 )
JCall = cor9
```

```
SSBRJC from Gibbs - G1.Genotype.ID : correlation = 0.967
```

```
SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 1.063
```

```
Out[55]: 0.967372023575404
```

```
In [56]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[56]: 0.5611659588224627
```

```
In [57]: IDs = readtable("G2.Genotype.ID", eltypes =[UTF8String], separator = ' ', head=10)
posAi = getPos(ped,IDs)
cor10 = cor(a[posAi],aHat1[posAi])[1,1]
reg10 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G2.Genotype.ID : correlation = %6.3f\n", cor10 )
@printf("SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg10 )
JCall = cor10
```

```
SSBRJC from Gibbs - G2.Genotype.ID : correlation = 0.967
```

```
SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = 1.037
```

```
Out[57]: 0.9667542438782681
```

```
In [58]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[58]: 1.0895964943313317
```

```
In [59]: IDs = readtable("G3.Genotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
posAi = getPos(ped,IDs)
cor11 = cor(a[posAi],aHat1[posAi])[1,1]
reg11 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G3.Genotype.ID : correlation = %6.3f\n", cor11)
@printf("SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg11)
JCA11 = cor11
```

```
SSBRJC from Gibbs - G3.Genotype.ID : correlation = 0.962
SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = 1.039
```

```
Out[59]: 0.9619710516459823
```

```
In [60]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[60]: 1.4865708922172312
```

```
In [61]: IDs = readtable("G4.Genotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
posAi = getPos(ped,IDs)
cor12 = cor(a[posAi],aHat1[posAi])[1,1]
reg12 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G4.Genotype.ID : correlation = %6.3f\n", cor12)
@printf("SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg12)
JCA12 = cor12
```

```
SSBRJC from Gibbs - G4.Genotype.ID : correlation = 0.968
SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 1.053
```

```
Out[61]: 0.9680421044476997
```

```
In [62]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[62]: 1.9559058424281812
```

```
In [63]: IDs = readtable("G5.Genotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
posAi = getPos(ped,IDs)
cor13 = cor(a[posAi],aHat1[posAi])[1,1]
reg13 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G5.Genotype.ID : correlation = %6.3f\n", cor13)
@printf("SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg13)
JCA13 = cor13
```

```
SSBRJC from Gibbs - G5.Genotype.ID : correlation = 0.975
SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = 1.066
```

```
Out[63]: 0.9747204818800306
```

```
In [64]: writedlm("Correlation.G5.G.C.txt",cor13)
```

```
In [65]: writedlm("Regression.G5.G.C.txt",reg13)
```



```
In [66]: GEBVG5Gall = aHat1[posAi]
         GEBVG5G=mean(GEBVG5Gall)
```

```
Out[66]: 1.4951712550014589
```

```
In [67]: IDs = readtable("G0.noGenotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
         posAi = getPos(ped,IDs)
         cor14 = cor(a[posAi],aHat1[posAi])[1,1]
         reg14 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G0.noGenotype.ID : correlation = %6.3f\n", cor14)
         @printf("SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = %6.3f\n", reg14)
         JCall = cor14
```

```
SSBRJC from Gibbs - G0.noGenotype.ID : correlation = 0.700
SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = 0.886
```

```
Out[67]: 0.7004831706735479
```

```
In [68]: GEBV = aHat1[posAi]
         mean(GEBV)
```

```
Out[68]: -1.3535014253923023
```

```
In [69]: IDs = readtable("G1.noGenotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
         posAi = getPos(ped,IDs)
         cor14 = cor(a[posAi],aHat1[posAi])[1,1]
         reg14 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G1.noGenotype.ID : correlation = %6.3f\n", cor14)
         @printf("SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV = %6.3f\n", reg14)
         JCall = cor14
```

```
SSBRJC from Gibbs - G1.noGenotype.ID : correlation = 0.755
SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV = 0.977
```

```
Out[69]: 0.755183980784623
```

```
In [70]: GEBV = aHat1[posAi]
         mean(GEBV)
```

```
Out[70]: -0.6904287279520583
```

```
In [71]: IDs = readtable("G2.noGenotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
         posAi = getPos(ped,IDs)
         cor15 = cor(a[posAi],aHat1[posAi])[1,1]
         reg15 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G2.noGenotype.ID : correlation = %6.3f\n", cor15)
         @printf("SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = %6.3f\n", reg15)
         JCall = cor15
```

```
SSBRJC from Gibbs - G2.noGenotype.ID : correlation = 0.753
SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = 1.007
```

```
Out[71]: 0.7530022845122945
```

```
In [72]: GEBV = aHat1[posAi]
         mean(GEBV)
```

```
Out[72]: -0.10503290814012683
```

```
In [73]: IDs = readtable("G3.noGenotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
         posAi = getPos(ped,IDs)
         cor16 = cor(a[posAi],aHat1[posAi])[1,1]
         reg16 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G3.noGenotype.ID : correlation = %6.3f\n", cor16)
         @printf("SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = %6.3f\n", reg16)
         JCall = cor16
```

```
SSBRJC from Gibbs - G3.noGenotype.ID : correlation = 0.760
SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = 1.009
```

```
Out[73]: 0.7595196964537299
```

```
In [74]: GEBV = aHat1[posAi]
         mean(GEBV)
```

```
Out[74]: 0.4616101351575765
```

```
In [75]: IDs = readtable("G4.noGenotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
         posAi = getPos(ped,IDs)
         cor17 = cor(a[posAi],aHat1[posAi])[1,1]
         reg17 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G4.noGenotype.ID : correlation = %6.3f\n", cor17)
         @printf("SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = %6.3f\n", reg17)
         JCall = cor17
```

```
SSBRJC from Gibbs - G4.noGenotype.ID : correlation = 0.752
SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 1.003
```

```
Out[75]: 0.7517949320721838
```

```
In [76]: GEBV = aHat1[posAi]
         mean(GEBV)
```

```
Out[76]: 0.9669458905481887
```

```
In [77]: numSSBayes
```

```
Out[77]: SSBR.NumSSBayes(54885,45885,9000,40000,39000,1000,200)
```