

Assignment 14

Ivan Pocrnic, 24.V.2018.

Please note that the code provided is partial

Assignment 14: 1 and 2

Run RENUM:

```
echo renum.par | renumf90 | tee renum.log
```

File with genotyped animals IDs:

```
awk '{print $1}' snp3.2k_XrefID | sort +0 -1 > genotyped
```

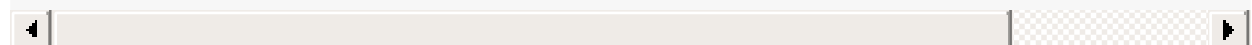
Please note that you can find list of genotyped animals in several ways!

Run BLUP:

```
mkdir blup; cd blup
cp ../renf90.par par.temp
grep -v OPTION par.temp > blup.par
sed -i 's:renf90.dat:../renf90.dat:g' blup.par
sed -i 's:renadd02.ped:../renadd02.ped:g' blup.par
echo "OPTION use_yams" >> blup.par
echo "OPTION sol se" >> blup.par
echo blup.par | blupf90 | tee blup.log
cp solutions blup_solutions
rm *.temp
```

Calculate accuracy:

```
awk '{if ($1==1 && $2==2) print $3,$4,(1-($5*$5)/0.40)**0.5}' blup_solutions | sort +0 -1 >
join -1 +1 -2 +1 ../genotyped blup_sol_acc > gb.tmp
awk '{ sum += $3; n++ } END { if (n > 0) print sum / n; }' gb.tmp
```



You don't have to use command line, you can use stat software.

Run ssGBLUP:

```
mkdir gblup; cd gblup
cp ../renf90.par par.temp
grep -v OPTION par.temp > gblup.par
sed -i 's:renf90.dat:../renf90.dat:g' gblup.par
sed -i 's:renadd02.ped:../renadd02.ped:g' gblup.par
echo "OPTION use_yams" >> gblup.par
echo "OPTION SNP_file ../snp3.2k" >> gblup.par
echo "OPTION chrinfo ../mrkmap.txt" >> gblup.par
echo "OPTION sol se" >> gblup.par
echo gblup.par | blupf90 | tee gblup.log
cp solutions gblup_solutions
rm *.temp
```

```
awk '{if ($1==1 && $2==2) print $3,$4,(1-($5*$5)/0.40)**0.5}' gblup_solutions | sort +0 -1
join -1 +1 -2 +1 ../genotyped gblup_sol_acc > gb.tmp
awk '{ sum += $3; n++ } END { if (n > 0) print sum / n; }' gb.tmp
```

Assignment 14: 3 and 4

```
awk 'BEGIN { for (i=1;i<45000;i++) print 1}' > W
```

Run ssGBLUP again if you like, including options as you are instructed.

```
echo "OPTION no_quality_control" >> gblup.par
echo "OPTION chrinfo ../mrkmap.txt" >> gblup.par
echo "OPTION saveGInverse" >> gblup.par
echo "OPTION weightedG W" >> gblup.par
```

Run postGS adding this options to existing parameter files:

```
echo 'OPTION chrinfo ../mrkmap.txt' >> post.par
echo 'OPTION readGInverse' >> post.par
echo 'OPTION weightedG W' >> post.par
```

Assignment 14: 5

Prediction for young/new animals:

```
echo new_animals | predF90 | tee predict.log
```

Assignment 14: 6

For this assignment you can use same parameter files as before! Only modify adding few details.

```
echo "OPTION Manhattan_plot" >> post.par  
echo "OPTION windows_variance 20" >> post.par
```

If you are running nonlinear A:

```
echo "OPTION which_weight nonlinearA" >> post.par
```

Assignment 14: 7

This assignments is repeating the 14.6. Just before running you will update the weights file "w" and run programs again.

For example:

```
cp solutions solutions_1  
cp snp_sol snp_sol_1  
cp chrSnp chrSnp_1  
cp W W_1  
awk '{print $7}' snp_sol > W  
echo gblup.par | blupF90 | tee gblup.log_2  
echo post.par | postGSf90 | tee post.log_2
```

Or if you want to write your own UNIX script for multiple (e.g. 3) iterations:

```
for i in {1..3}
```

do

```
echo gblup.par | blupf90 | tee gblup.log_${i}
cp solutions solutions_${i}
echo post.par | postGSf90 | tee post.log_${i}
cp snp_sol snp_sol_${i}
cp chrsnp chrsnp_${i}
cp w w_${i}
awk '{print $7}' snp_sol > W
```

done