Work Log for October

Logan Brown

October 13, 2014

Contents

1   Goals for the Month

1. Find out what causes some of the probabilties to go outrageously high

2. Use the Simulated Yeast Genome, compare to the Ecoli genome

3. If possible, NSE Patch

4. Debugging Thoughts

2   Progress/Notes

2.1   Find out what causes some of the probabilties to go outrageously high

2.1.1   Compare Differences in the Code

```
my.coef.r
my.drawPhiConditionalAllPred.r
my.drawPhiConditionalAll.r
my.estimatePhiOne.r
my.fitMultinomOne.r
my.logdmultinomCodOne.r
my.objectivePhiOne.Lfp.r
my.objectivePhiOne.nlogL.r
my.objectivePhiOne.nlogphiL.r
my.objectivePhiOne.phiLfp.r
my.pPropTypeNoObs.lognormal_bias.r
plotbin.r
plotmodel.r
simu.orf.r
```

Codes where NSE and ROC are the same:

1. my.coef.r

2. my.estimatePhiOne.r

2.2   Look at stddev(phi)

I have a run going that is tracking the scales of the MCMC for the phi values.

Since the scale of the MCMC should be about the Standard Devation of Phi, we'll see what happens.

PROBLEM: The ROC model has different Phi scales for each gene. The NSE model keeps all the scales the same between different genes. That could be contributing to the problem? ~~I think this is caused by the opatch we wrote that replaces Wei-Chen's scaling by .5 with subtracting out log(DBL_MAX).~~

ROC Model Phi Scales	NSE Model Phi Scales
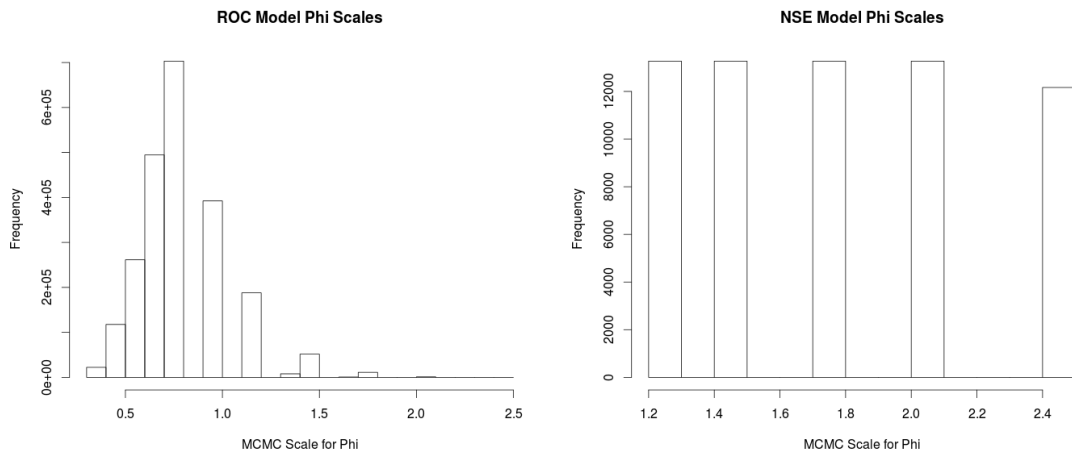
Every 6th time that the scale is updated, the code resets all the scales to 1. This definitely happens in Yeast (NSE and ROC), and I think it happens in E Coli. Needs further testing.

2.3 Simulated Yeast Genome

2.3.1 Use Sections of the Simulated Yeast Genome

I ran one 5th of the yeast genome(section1.fasta, section1.csv) and it FINISHED. but... not well.

See the results in data/10.01.yeastnse.*

- It took just over 30 hours to finish
  started at: 2014-10-01 15:57:54
  finished at: 2014-10-02 22:08:56

- The proposed phi values are still outrageously too high. The final proposed phi had mean value 2766786.09545948

- Still proposing some 0 (-Inf) values for lpProp. At most, around 45-49 Inf per run.

- Xobs (the original phi values) are generally much lower than the results of run

- The values are not very accurate.

2.3.2 Crashed Yeast Run

The vast vast majority of the phi values being proposed are fine. -Inf values are throwing it off.

2.4 E. Coli genome

A LARGE NSE RUN OF THE E. COLI GENOME RAN TO COMPLETION.

2.5 If possible, NSE Patch

3 Goals for next Month

1. Future Goal