Parameterization Issues and Diagnostics in MCMC

Gill Chapter 10 & 12

November 10, 2008

Convergence to Posterior Distribution

Theory tells us that if we run the Gibbs sampler long enough the samples we obtain will be samples from the joint posterior distribution (target or stationary distribution). This does not depend on the starting point (forgets the past).

Important questions

- How long until we get there?
- Once we are there, how long do we need to run it?

Mixing of the chain plays a critical role in how fast we can obtain good results.

WinBUGS Model - Non-Centered

```
for (n in 1:N){
 muj[n] <- alpha[station[n]]+beta[station[n]]*X[n]
 Y[n] ~ dnorm(muj[n], phi)
}

for (j in 1:J) {
  alpha[j] ~ dnorm(alpha.mu, alpha.phi)
  beta[j] ~ dnorm(beta.mu,  beta.phi)
}
phi ~ dgamma(.001, .001)
alpha.mu ~ dnorm(0.0, 1.0E-6)
alpha.sigma ~ dunif(0, 100)
alpha.phi <-1/(alpha.sigma*alpha.sigma)
beta.mu ~ dnorm(0.0, 1.0E-6)
beta.phi <- pow(beta.sigma, -2)
beta.sigma ~ dunif(0, 100)
```

CODA

The CODA package provides many popular diagnostics for assessing convergence of MCMC output from WinBUGS (and other programs)

```
> out2 =  read.coda.interactive()
Enter CODA index file name
(or a blank line to exit)
1: codaIndex.txt
Enter CODA output file names, separated by return key
(leave a blank line when you have finished)
1: coda1.txt
2: coda2.txt
3:

> codamenu()  # run CODA interactively
```

Useful Diagnostics/Functions

- Geweke: `geweke.diag()`
- Gelman-Rubin: `gelman.diag()`
- Heidelberg & Welch: `heidel.diag()`
- Raftery-Lewis: `raftery.diag()`
- Effective Sample Size: `effectiveSize()`
- Autocorrelation `autocorr.plot()`
- Cross Variable Correlations:`crosscorr.plot()`

Geweke

Geweke (1992) proposed a convergence diagnostic for Markov chains based on a test for equality of the means of the first and last part of a Markov chain (by default the first 10% and the last 50%).

If the samples are drawn from the stationary distribution of the chain, the two means are equal and Geweke's statistic has an asymptotically standard normal distribution.

Gelman-Rubin

Gelman and Rubin (1992) propose a general approach to monitoring convergence of MCMC output in which $m > 1$ parallel chains are run.

- Use different starting values that are overdispersed relative to the posterior distribution.
- Convergence is diagnosed when the chains have "forgotten" their initial values, and the output from all chains is indistinguishable.
- The diagnostic is applied to a single variable from the chain. It is based a comparison of within-chain and between-chain variances (similar to a classical analysis of variance)
- Assumes that the target is normal (transformations may help)
- Values of $\hat{R}$ near 1 suggest convergence

Heidelberg-Welch

The convergence test uses the Cramer-von-Mises statistic to test the null hypothesis that the sampled values come from a stationary distribution.

- The test is successively applied, firstly to the whole chain, then after discarding the first 10%, 20%, of the chain until either the null hypothesis is accepted, or 50% of the chain has been discarded.
- The latter outcome constitutes "failure" of the stationarity test and indicates that a longer MCMC run is needed.
- If the stationarity test is passed, the number of iterations to keep and the number to discard (burn-in) are reported.

Raftery-Lewis

Calculates the number of iterations required to estimate the quantile q to within an accuracy of $\pm r$ with probability p.

- Separate calculations are performed for each variable within each chain. If the number of iterations in data is too small, an error message is printed indicating the minimum length of pilot run.
- The minimum length is the required sample size for a chain with no correlation between consecutive samples. An estimate I (the 'dependence factor') of the extent to which autocorrelation inflates the required sample size is also provided.
- Values of I larger than 5 indicate strong autocorrelation which may be due to a poor choice of starting value, high posterior correlations or stickiness of the MCMC algorithm.
- The number of burn-in iterations to be discarded at the beginning of the chain is also calculated.

Others

- Effective Sample Size: Provides an estimate of the sample size (number of MCMC draws) adjusted for autocorrelation.
- Autocorrelations: lag correlations within a variable
- CrossCorrelations: correlation between variables

Summary

- Diagnostics cannot guarantee that chain has converged
- Can indicate that it has not converged

Solutions?

- Run longer and thin output
- Reparametrize model
- “Block” correlated variables together
- Add auxiliary variables (Slice-sampler for example)
- Use “Rao-Blackwellization” in estimation