Now you see it? Now you don’t?
The role of graphics in identifying MCMC convergence

ASC ISCB, Melbourne
Monday 27th August, 2018

Nicholas Tierney, Di Cook, Miles McBain
Why care about convergence?
Converged?
Converged?
We *know* when it looks converged.

Right?
Yes, and no.
Research Question

Can visualisations be used to detect convergence?

Are some visualisations better than others for detecting convergence?

Is confidence related to detecting convergence?

Is experience with Bayesian statistics related to detecting convergence?
Experimental Method

Model

Bad Specification
60,000 Burn in
10,000 Samples

Good Specification
60,000 Burn in
10,000 Samples
Experimental Method

Slice 10,000 samples into 10 sets
Experimental Method

Select nine bad and one good
Experimental Method

Generate graphics

Select the one that is most converged
"@#$ I have to log on" The poster innovations at #bayesonbeach2017 continue @nj_tierney
Demo
Experience (years)

Please provide an estimate of the number of years experience you have with Bayesian Statistics

- 0 - 1 years
- 1 - 2 years
- 2 - 3 years
- 3 - 4 years
- 4 - 5 years
- 5 plus years

< Previous  Next >
Experience (Rank)

Please evaluate your experience with Bayesian Statistics, selecting one of the following:

- No experience at all
- Some experience
- Moderate experience
- Very experienced
- Expert
Identify which graphic is the most converged

Choose one:

1 2 3 4 5 6 7 8 9 10

< Previous  Next >
From 0 to 100, with 0 being 'no confidence at all' and 100 being 'absolutely confident', rank your confidence with this decision.
Identify which graphic is the most converged

Choose one:

1 2 3 4 5 6 7 8 9 10
From 0 to 100, with 0 being 'no confidence at all' and 100 being 'absolutely confident', rank your confidence with this decision.
Identify which graphic is the most converged

Choose one:

1 2 3 4 5 6 7 8 9 10
From 0 to 100, with 0 being 'no confidence at all' and 100 being 'absolutely confident', rank your confidence with this decision.
model{
  for( i in 1 : N ) {
    y[i] ~ dnorm(mu[i], tau)
    mu[i] <- lambda[T[i]]
    T[i] ~ dcat(P[])
  }
  P[1:maxT] ~ ddirch(alpha[])
  lambda[1] ~ dnorm(0.0, 1.0E-6)
  for (j in 2:maxT){
    lambda[j] <- lambda[j-1] + theta[j-1]
    theta[j-1] ~ dnorm(0.0, 1.0E-6)I(0.0, )
  }
  tau ~ dgamma(0.001, 0.001)
  sigma <- 1 / sqrt(tau)
}
Bad Model

22 Mixtures

Good model

2 Mixtures
# Results

## Descriptive statistics

**N = 16**

<table>
<thead>
<tr>
<th>Expertise</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some Experience</td>
<td>3</td>
</tr>
<tr>
<td>Moderate Experience</td>
<td>6</td>
</tr>
<tr>
<td>Very Experienced</td>
<td>4</td>
</tr>
<tr>
<td>Expert</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Expertise</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 1 years</td>
<td>2</td>
</tr>
<tr>
<td>1 - 2 years</td>
<td>1</td>
</tr>
<tr>
<td>2 - 3 years</td>
<td>2</td>
</tr>
<tr>
<td>3 - 4 years</td>
<td>3</td>
</tr>
<tr>
<td>5 plus years</td>
<td>8</td>
</tr>
</tbody>
</table>
Results

Can visualisations be used to detect convergence?
Results

Are some visualisations better than others for detecting convergence?
Results

Is confidence related to detecting convergence?
Results

Is confidence related to detecting convergence?
Summary

Can visualisations be used to detect convergence?

Are some visualisations better than others for detecting convergence?

Is confidence related to detecting convergence?

Is experience with Bayesian statistics related to detecting convergence?
Discussion

Are we measuring the right thing?

Multiple Chains would be more useful

Do we tell people the model before we show them the diagnostics?
Future Work

Using a variety of diagnostics to make a decision

- Geweke diagnostic
- Gelman-Rubin
- Effective Sample Size

Generating non convergent samples from a variety of models

Use and test multiple different chains
Future Work

Implement Training

- Explain convergence
- Users understand good/bad

Get feedback and create a better experiment
Acknowledgements

Dr. David Frazier for his helpful suggestions on the methodology

Dr. Sam Clifford for discussions on the model implementation, and writing the JAGS code to create the “Good” and “Bad” models.
References


Learn more

njtierney/dualchain

nj_tierney

njtierney.com

nicholas.tierney@gmail.com