A Local Version of R to Improve MCMC Convergence Diagnostic
Théo Moins, Julyan Arbel, Stéphane Girard, Anne Dutfoy

To cite this version:
Théo Moins, Julyan Arbel, Stéphane Girard, Anne Dutfoy. A Local Version of R to Improve MCMC Convergence Diagnostic. BAYSM 2022 - Bayesian Young Statisticians Meeting, Jun 2022, Montréal, Canada. hal-03765339

HAL Id: hal-03765339
https://hal.science/hal-03765339
Submitted on 31 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Overview: has the chain converged?

- Diagnosing convergence of Markov chain Monte Carlo (MCMC) is crucial in Bayesian analysis.
- It is frequent to run multiple chains in parallel with different initial values.
- At convergence, all chains should have a similar distribution.

\[ R = \sqrt{\frac{W + B}{W}} \]

with

\[ \hat{B} = \frac{1}{M - 1} \sum_{i=1}^{M} (\hat{\theta}^{(i,m)} - \hat{\theta}^{(-)})^2, \]

\[ \hat{W} = \frac{1}{M} \sum_{i=1}^{M} s^{(i,m)}. \]

If \( R > 1.01 \), convergence issue detected.

Ways of fooling \( \hat{R} \) and improvements \[2]\)

Two cases where \( \hat{R} \) fails:

- Chains with infinite mean and different locations: \( \hat{R} \approx 1 \)
- Bulk-\( \hat{R} \): \( \hat{R} \) on \( z^{(m,n)} \), the normally transformed ranks of \( \theta^{(m,n)} \).
- Chains with same mean and different variances: \( \hat{R} \approx 1 \)
- Tail-\( \hat{R} \): \( \hat{R} \) on \( \zeta^{(m,n)} \), the deviations from the median.

Population version

Assume chain \( m \) has stationary distribution \( F_m \),
As \( \mathbb{E}[\{\theta^{(m,n)} \leq x\}] = F_m(x) \), and \( \text{Var} \{\theta^{(m,n)} \leq x\} = F_m(x) - (1 - F_m(x))^2 \),

\( W(x) = \frac{1}{M} \sum_{m=1}^{M} (F_m(x) - F_m(x))^2 \)

\( B(x) = \frac{1}{M} \sum_{m=1}^{M} F_m(x) - \left( \frac{1}{M} \sum_{m=1}^{M} F_m(x) \right)^2 \)

\( R(x) := \frac{W(x) + B(x)}{W(x)} \)

Fooling rank-\( \hat{R} \), robustness of \( \hat{R}_{\infty} \)

- \( F \rightarrow \mathcal{F} \) with a Markov chain central limit theorem:

\[ \sqrt{n/m} (\hat{F} - F(x)) \overset{\text{d}}{\rightarrow} \mathcal{N}(0, \sigma^2(x)) \]

Define \( \text{ESS}(x) := \frac{n}{m} \sum_{i=1}^{n} \sum_{j=1}^{m} I(\theta^{(i,j)} \leq x) \)

Then for any \( x \in \mathbb{R} \):

\[ \text{ESS}(x)/(\hat{R}^2(x) - 1) \xrightarrow{n \to \infty} \chi^2_{M - 1} \]

\[ R_{\text{lim},\alpha}(x) := \left[ 1 + \frac{\alpha}{\text{ESS}(x)} \right]^{-1} \]

\[ \Rightarrow \mathbb{P}(\hat{R}(x) \geq R_{\text{lim},\alpha}(x)) \approx \alpha. \]

With \( \alpha = 0.05 \), \( m = 4 \), \( \text{ESS}(x) = 400 \), \( R_{\text{lim},\alpha}(x) \approx 1.010 \)

Multivariate extension

- Compute the univariate \( \hat{R}_{\infty,i} \) separately on each coordinate \( i \).
- If \( \hat{R}_{\infty,i} < 1.01 \) for all \( i \), compute the multivariate \( \hat{R}_{\infty} \) to check convergence of the dependence structure.

References