MCMC Diagnostics

Review

In the practical you used *Metropolis-Hastings* with a *Gaussian* proposal distribution to infer *one* parameter, R_0

In this session we will:

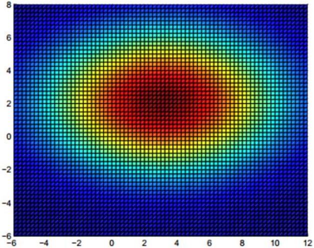
- extend to multivariate inference
- learn about MCMC diagnostics
- think about accuracy and efficiency

Interlude: Multivariate Gaussian distribution

To infer more multiple parameters we can use multivariate Gaussian

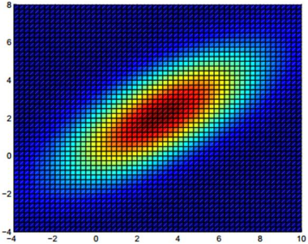
mean $\mu = \begin{bmatrix} 3 & 2 \end{bmatrix}$

covariance $\Sigma = \begin{bmatrix} 25 & 0 \\ 0 & 9 \end{bmatrix}$



mean $\mu = \begin{bmatrix} 3 & 2 \end{bmatrix}$

covariance $\Sigma = \begin{bmatrix} 10 & 5 \\ 5 & 5 \end{bmatrix}$



For accurate and efficient MCMC we tune the variance and covariance of the proposal distribution.

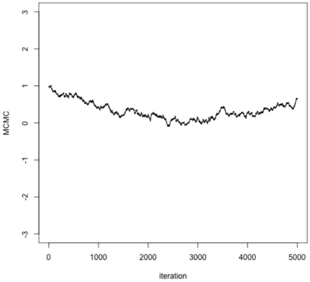
Why I like hairy caterpillars

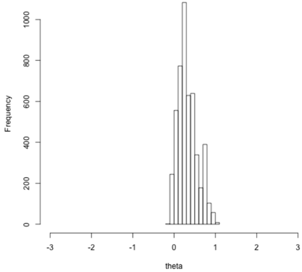
Key characterisitics

- Straight
- Plump head, plump rear!
- Multiple colours

Choosing a proposal distribution

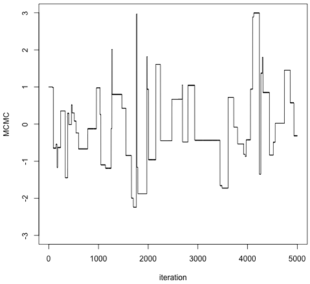
If variance is too small, the chain will be slow to reach the target distribution.

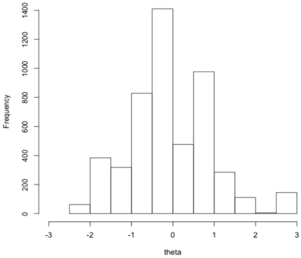




Choosing a proposal distribution

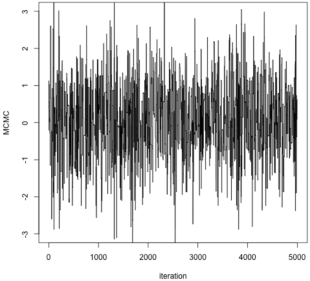
If variance is too high, many proposed values will be rejected and the chain will *stick* in one place for many steps.

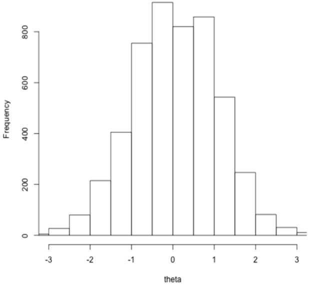




Choosing a proposal distribution

If variance is just right, the chain will efficiently explore the full shape of the target distribution.





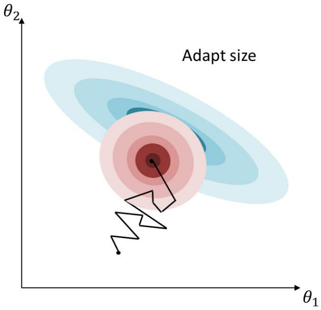
Try several different proposal distributions (pilot runs), aiming for an acceptance rate between 24% and 40%.

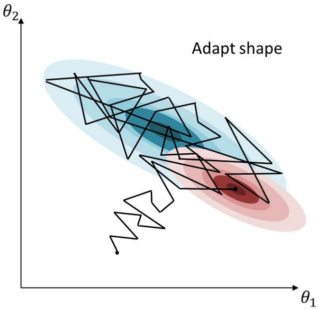
Adaptive MCMC

- Adaptive MCMC alters proposal distribution while chain is running.
- Start with large symmetric variance, scan around to find a mode.
- Then alter shape of proposal distribution to match covariance matrix of accepted values.
- Eventually proposal density should match the shape of target density.

Adaptive MCMC

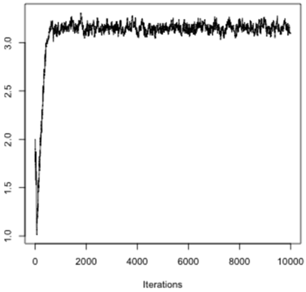
Two-stage adaptation





Burn-in

- We can start our MCMC chain anywhere.
- It can take a while to reach and explore the target density $f(\theta)$.
- Throw away early samples: burn-in phase.
- How much to discard?



Lepidoptera caterpillar

MCMC sample size

- In MCMC, each sample depends on the one before - auto-correlation
- Reduce degree of auto-correlation by thinning, only retain every n^{th} sample.
- Information content of MCMC samples is given by the effective sample size (ESS).
- We use the R package *coda*.

Accuracy and efficiency

How does each element influence accuracy and efficiency?

- Burn-in
- MCMC iterations (after burn-in)
- Thinning
- Number of chains (with different initial conditions)
- Proposal distribution
- Transforming parameters