More on MCMC sampling

Tutorial

26.06.2019

Mixing time of MCMC

"Burn-in" phase

It takes time for the MCMC process to converge to the stationary distribution. This transitional phase is called the **"burn-in" phase** and its duration is called the **mixing time**.

Mixing time of MCMC

"Burn-in" phase

It takes time for the MCMC process to converge to the stationary distribution. This transitional phase is called the **"burn-in" phase** and its duration is called the **mixing time**.

How to determine the mixing time?

Bad news: hard to determine in general!

- One of the issue when using MCMC inference;
- High if well separated modes exists.

More on this subject: Markov Chains and Mixing Times, Levin et al.
`https://pages.uoregon.edu/dlevin/MARKOV/mcmt2e.pdf`

Autocorrelation of samples

MCMC samples are not i.i.d.!

The sequence of samples generated from an MCMC process will not be i.i.d.: in fact, neighboring samples follow the conditional distribution specified by the transition kernel.

Autocorrelation of samples

MCMC samples are not i.i.d.!

The sequence of samples generated from an MCMC process will not be i.i.d.: in fact, neighboring samples follow the conditional distribution specified by the transition kernel.

Reducing the autocorrelation with thinning

A practical approach to reduce autocorrelation between generated samples is to perform **thinning**, i.e. use only every k-th sample.

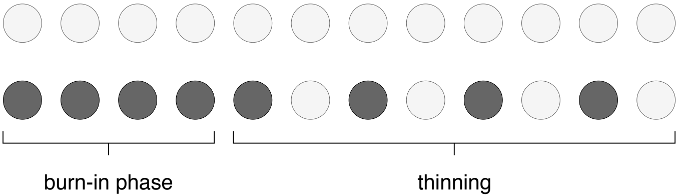
Autocorrelation of samples

MCMC samples are not i.i.d.!

The sequence of samples generated from an MCMC process will not be i.i.d.: in fact, neighboring samples follow the conditional distribution specified by the transition kernel.

Reducing the autocorrelation with thinning

A practical approach to reduce autocorrelation between generated samples is to perform **thinning**, i.e. use only every k-th sample.



Choice of q for Metropolis-Hastings

Admissible proposal distribution

The proposal distribution q is **admissible** (or **valid**) if its support include that of the target distribution p^*:

$$\mathrm{supp}(p^*) \subseteq \cup_x \mathrm{supp}(q(\cdot|x)) \tag{1}$$

Choice of q for Metropolis-Hastings

Admissible proposal distribution

The proposal distribution q is **admissible** (or **valid**) if its support include that of the target distribution p^*:

$$\mathrm{supp}(p^*) \subseteq \cup_x \mathrm{supp}(q(\cdot|x)) \tag{1}$$

Choice of proposal distribution

Of course, it should be valid!

common choice: Gaussian distribution

Effect of variance:

- Too high: high rejection rate $\Rightarrow$ rarely changes state, "sticky";
- Too low: short-move random walk $\Rightarrow$ stuck at isolated mode.

Blocking Gibbs sampling

Why in blocks?

Gibbs sampling uses coordinate-wise update $\Rightarrow$ highly correlated samples

Blocking Gibbs sampling

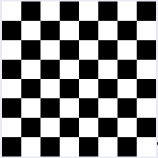
Why in blocks?

Gibbs sampling uses coordinate-wise update $\Rightarrow$ highly correlated samples

Checkerboard pattern for grid structure

For CV applications, we often encounter grid-shaped graphical models. A classic blocking design is to use the checkerboard pattern.

- nodes in a block are conditionally independent given nodes in the other block!

[a]

[a] From Wikipedia: https://en.wikipedia.org/wiki/Checkerboard accessed on June 25th, 2019