PMCMC Reading Group - Particle marginal Metropolis-Hastings sampler

Dennis Prangle

17th February 2009

Idealised Algorithm

- Target is $p(\theta, x_{1:T}|y_{1:T})$
- Metropolis-Hastings algorithm updating θ and $x_{1:T}$ jointly

Idealised algorithm:

1. Sample proposal θ^* from $q(\theta^*|\theta)$ (pre-specified arbitrary proposal distribution)
2. Sample $x^*_{1:T}$ from $p_{\theta^*}(x^*_{1:T}|y_{1:T})$ (i.e. posterior conditional on θ^*)
3. Accept with standard MH acceptance ratio:

$$\min\left[1, \frac{p(\lambda^*|y_{1:T})q(\lambda|\lambda^*)}{p(\lambda|y_{1:T})q(\lambda^*|\lambda)}\right]$$

(where $\lambda = (\theta, x_{1:T})$ i.e. all the unknown parameters)

- More appealing than either MCMC or SMC alone
- Idea: replace step 2 with SMC draw - sample from an approximation to the conditional posterior

Particle Marginal Metropolis-Hastings sampler

- (Initialisation step omitted)

1. Sample proposal θ^* from $q(\theta^*|\theta)$
2. Run an SMC algorithm targetting $p_{\theta^*}(x_{1:T}|y_{1:T})$.
3. Sample $x^*_{1:T}$ from $\hat{p}_{\theta^*}(.|y_{1:T})$ (i.e. from SMC output)
4. Accept with probability:

$$\min\left[1, \frac{\hat{p}_{\theta^*}(y_{1:T})p(\theta^*)}{\hat{p}_{\theta}(y_{1:T})p(\theta)} \frac{q(\lambda|\lambda^*)}{q(\lambda^*|\lambda)}\right]$$

Recall:

- $p(\theta)$ is prior
- $\hat{p}_{\theta^*}(y_{1:T})$ is the SMC marginal likelihood estimate from step 2

Properties

Under mild conditions:

- Update leaves target distribution invariant
- Algorithm converges to target distribution
- Acceptance probability converges to that of idealised algorithm as number of particles $\rightarrow \infty$

Proof

1. Write down proposal and target density of one iteration
   - On an extended space of all variables involved in SMC
   - The marginal target for $(\theta, x_{1:T})$ is as required
2. Show that the acceptance probability required to produce a MH update is as given
3. Prove convergence - done in a more general setting in Andrieu and Roberts pseudo-marginal paper (theorem 1)

Extensions / problems

- Choice of N (number of particles)
  - Performance (acceptance rate) suffers if $N << T$
  - Possible adaptive choice of N
- Mix PMMH update with other MH updates (to reduce computation cost)?
- All SMC particles can be used to estimate $Ef(\theta, x_{1:T})$ (sec 4.6)