Look Ma, No Sampling!

Colin Fox & Lennart Golks

ISBA BayesComp Singapore 2025

Tired of waiting for your MCMC to run? No problem, just skip the MCMC and evaluate expectations using function approximation and numerical integration.

Bayesian Formulation

θ Hyperparameters: Hyper-prior: $[\theta]$

x Latent field: Representation and prior $[x|\theta]$

y Observed data: Likelihood: $[y|x]$

The focus of inference is the posterior distribution:

$$[x,\theta|y] = \frac{[y|x]\,[x|\theta]\,[\theta]}{[y]}$$

We assume the normalizing constant $[y]$ is finite.

Posterior Inference

We wish to compute expectations:

$$\mathrm{E}_{x,\theta|y}[f(x)] = \int f(x)[x,\theta|y]\,\mathrm{d}x\,\mathrm{d}\theta$$

Notation

We learned this notation from Alan Gelfand: read $[a]$ as "the distribution over a", and $[a|b]$ as "the distribution over a given b". We will abuse this notation to also denote the density function.

Monte Carlo Integration

$$\mathrm{E}_{x,\theta|y}[f(x)] \approx \frac{1}{N}\sum_{i=1}^{N} f(x_i)$$

where $\{(x_i,\theta_i)\}$ is ergodic for $[x,\theta|y]$.

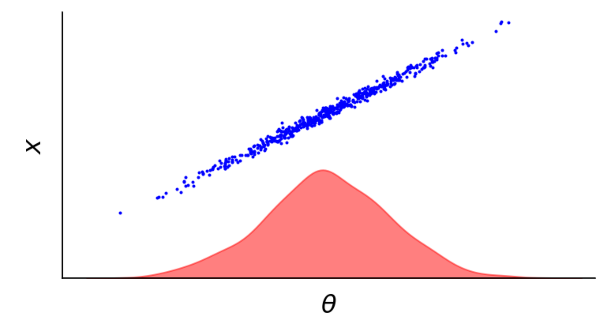
"Monte Carlo is an extremely bad method; it should be used only when all alternative methods are worse." (Alan Sokal)

Markov chain Monte Carlo

A representative MCMC scheme is the block Gibbs sampler

- Draw $\theta' \sim [\theta|x]$
- Draw $x' \sim [x|\theta', y]$

simulating a transition kernel that targets the posterior $[x,\theta|y]$.



Narrow scatter plot shows why this is slow

Better is to move in the marginal posterior over hyperparameters $[\theta|y]$
H. Rue and L. Held. *Gaussian Markov random fields : Theory and applications*. Chapman Hall, New York, 2005.

Factorize Posterior (MTC)

factorisation of the posterior density

$$[x,\theta|y] = [x|\theta,y]\,[\theta|y]$$

i.e. into the full conditional for x and the marginal posterior over θ.

Lemma Sampling $\theta' \sim [\theta|y]$ then $x' \sim [x|\theta', y]$ generates a sample from the posterior distribution, i.e.,

$$(x',\theta') \sim [x,\theta|y].$$

C. Fox and R. A. Norton. Fast sampling in a linear-Gaussian inverse problem. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):1191–1218, 2016.

Some kind of pictorial break

But not a box with text like this

Marginal Posterior over Hyperparameters

The marginal posterior distribution over hyperparameters is defined by the integral $[\theta|y] = \int_X [x,\theta|y]\,\mathrm{d}x$, as mentioned above, but this calculation is to be avoided because the integral is over the h ent field x. A cheap algebraic calculation is available when the full conditional for x

These terms simplify when $A^{-1}b = \mathcal{A}^{-1}\beta$

$$[x|\theta,y] = \frac{[y|x]\,[x|\theta]}{[y|\theta]}$$

has known form, implying that the normalising constant $[y|\theta]$ has known θ dependence, and hence one can evaluate the marginal posterior over θ

$$[\theta|y] \propto [y|\theta][\theta].$$

R. A. Norton, J. A. Christen, and C. Fox. Sampling hyperparameters in hierarchical models: improving on Gibbs for high-dimensional latent fields and large datasets. *Communications in Statistics - Simulation and Computation*, 47(9):2639–2655, 2018.

Tensor Train Representation of PDF

A surface rendering of a density in TT format
TT formula

S. Dolgov, K. Anaya-Izquierdo, C. Fox, and R. Scheichl. Approximation and sampling of multivariate probability distributions in the tensor train decomposition. *Statistics and Computing*, 30(3):603–625, 2020.

Inverse Problem of Recovering Ozone Profile

I'm thinking picture of satellite, DAG, RTE

Posterior Expectation by TT and Affine Representations

Picture of posterior inference, or timings, or some output summaries – perhaps a box below with some snappy summary or conclusions

R. A. Norton and C. Fox, Efficiency and computability of MCMC with Langevin, Hamiltonian, and other matrix-splitting proposals, arxiv:1501.03150.