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 tensor train function representation and numerical integration.

Bayesian Formulation

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

$\downarrow$

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

$\downarrow$

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

Posterior Inference

The focus of inference is the posterior distribution:

try blblbl

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

We assume the normalizing constant $[y]$ is finite.
We wish to compute expectations:

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

This 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.

DO THIS

DON'T DO THIS

Quadrature and the Law of Total Expectation

The posterior expectation of any function $h(x)$ may be written

$$\mathrm{E}_{x,\theta|y}[h(x)] = \mathrm{E}_{\theta|y}\left[\mathrm{E}_{x|\theta,y}[h(x)]\right]$$

We compute the outer expectation by quadrature (see next panel). When the inner expectation is cheap to calculate, evaluating $\mathrm{E}_{x,\theta|y}[h]$ requires **no MCMC**.

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)

Marginal Posterior over Hyperparameters $[\theta|y]$

$$[\theta|y] = \int_X [x,\theta|y]\,\mathrm{d}x$$

Avoid calculating this integral over high-dimensional latent field x.

A cheap algebraic calculation is available when the full conditional for x

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

has known form. Then the normalising constant $[y|\theta]$ has known θ dependence, and

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

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.

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 –

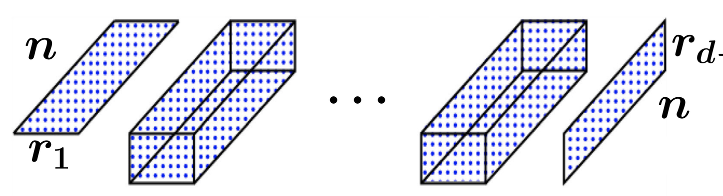
Better is to move in the marginal posterior over hyperparameters $[\theta|y]$

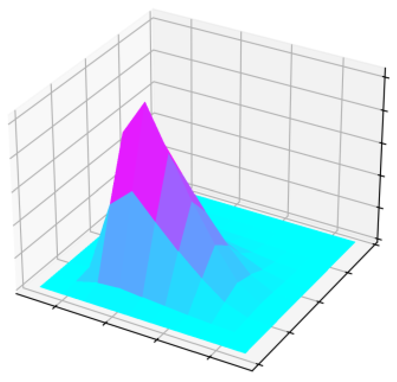
H. Rue and L. Held. *Gaussian Markov random fields : Theory and applications*. Chapman Hall, New York, 2005.

Tensor Train Representation of $[\theta|y]$

A surface rendering of a density in TT format

TT formula





$$f_{X_k}(x_k) = \frac{1}{z}\left(\gamma' \prod_{i=1}^{k-1} \lambda_i(X_i) \prod_{i=k+1}^{d} \lambda_i(X_i) + \sum_{l_{k-1}=1}^{r_{k-1}} \sum_{l_k=1}^{r_k} \left(\sum_{i=1}^{n} \phi_k^{(i)}(x_k) D_k[l_{k-1}, i, l_k]\right)^2\right) \lambda_k(x_k)$$

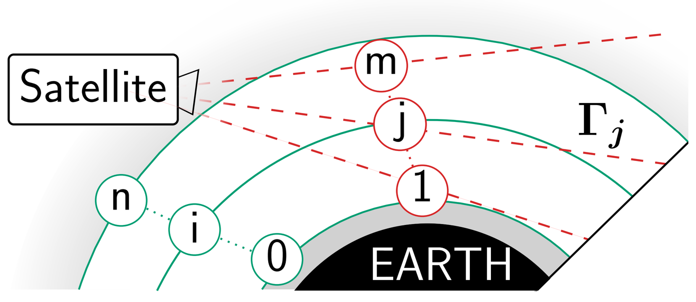
S. Dolgov, K. Anaya-Izquierdo, C. Fox, and R. Scheichl. Approximation and sampling of multivariate probability distributions in the tensor train decomposition. *Statistics and Computing*, 30(3):603–625, 2020.

Independent Posterior Sampling (MTC)

If the inner expectation is not available exactly, draw *independent* posterior samples by:
Use TT representation to sample $\theta' \overset{\text{iid}}{\sim} [\theta|y]$ then $x' \sim [x|\theta', y]$

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.

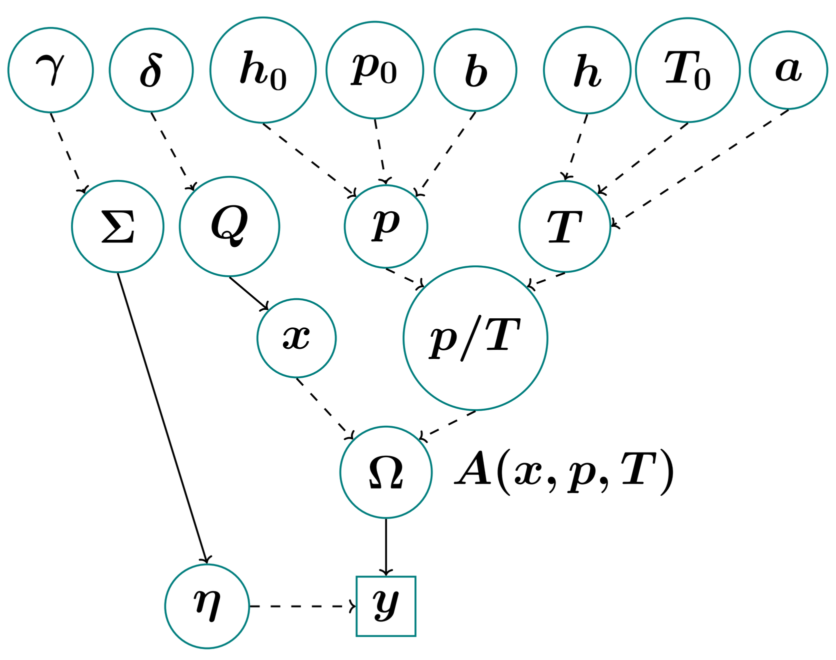
Inverse Problem of Recovering Ozone Profile



$$y_j = \int_{\Gamma_j} B(\nu, T) k(\nu, T) \frac{p(T)}{k_B T(r)} x(r) \tau(r) \mathrm{d}r + \eta_j$$

$$\tau(r) = \exp\left\{-\int_{r_{\text{obs}}}^{r} k(\nu, T) \frac{p(T)}{k_B T(r')} x(r') \mathrm{d}r'\right\}$$

$$y = MAx + \eta$$



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