Homotopy Sampling and Data Assimilation

Juan Restrepo[1], Jorge Ramirez[2], and Robert Miller[1]

[1]Oregon State University
[2]Universidad Nacional de Colombia

November 24, 2022

Abstract

Importance sampling is modified via {\it homotopy continuation} to improve the efficiency and success of the sampler. The homotopy will use a known distribution as a starting empirical importance sampling distribution and generate a continuous schedule which culminates with the target distribution. The focus is the estimation of the normalization constant of the target distribution. The homotopy method is extended to a Bayesian setting, for stationary and time dependent posterior distributions. The numerical implementation uses a combination of sample averages, with sampling parameter N, and homotopy stages M, where M is typically a small number. The algorithm replaces homotopy stages for sampling steps, potentially resulting in a better or more efficient importance sampler. Numerical experiments suggest this is the case. The results also suggest that the method may improve the efficiency of the sampler by concentrating the samples in regions of greater impact.

Homotopy Sampling and Data Assimilation

Juan M. Restrepo[1,2,3], **Jorge M. Ramírez**[4], **Robert Miller**[3].

[1] Departments of Mathematics & [2] Statistics, [3] College of Earth Oceans and Atmospheric Sciences, Oregon State University, Corvallis OR USA
[4] Universidad Nacional de Colombia, Sede Medellín, Colombia,
restrepo@math.oregonstate.edu, http://www.math.oregonstate.edu/∼restrepo

Abstract

Importance sampling is modified via homotopy continuation to improve the efficiency and success of the sampler. The homotopy will use a known distribution as a starting empirical importance sampling distribution and generate a continuous schedule which culminates with the target distribution. The focus is the estimation of the normalization constant of the target distribution. The homotopy method is extended to a Bayesian setting, for stationary and time dependent posterior distributions.
The numerical implementation uses a combination of sample averages, with sampling parameter N, and homotopy stages M, where M is typically a small number. The algorithm replaces homotopy stages for sampling steps, potentially resulting in a better or more efficient importance sampler. Numerical experiments suggest this is the case. The results also suggest that the method may improve the efficiency of the sampler by concentrating the samples in regions of greater impact.

1. HOMOTOPY IMPORTANCE SAMPLING

Find $Z_1 = \int q(x)\,dx$, where $q(x)$, an improper probability density function (pdf) via a homotopy procedure, starting with a known pdf $p(x)$:
Let Z_s, $s \in [0,1]$, a continuous function

$$Z_s = \int q^s(x) p^{(1-s)}(x) dx,$$

and $p(x)/Z_0$ is a known pdf. With

$$\phi_s(x) = \frac{q^s(x)p^{(1-s)}(x)}{Z_s}.$$

then, $p = \phi_0$ to $q = \phi_1$ and

$$\ln\left(\phi_s(x)\right) = s\ln\left(q(x)\right) + (1-s)\ln\left(p(x)\right) - \ln Z_s, \qquad 0 \le s \le 1,$$

Assuming continuity of Z_s, we find that

$$\frac{dZ_s}{ds} = \int \log\left(\frac{q}{p}\right) q^s p^{1-s} dx := \left\langle \log\left(\frac{q}{p}\right)\right\rangle_s Z_s.$$

Hence

$$\frac{dZ_s}{Z_s} = \left\langle \log\left(\frac{q}{p}\right)\right\rangle_s ds.$$

We note that

$$\frac{Z_{s+\epsilon}}{Z_s} = \frac{1}{Z_s}\int \left(\frac{q(x)}{p(x)}\right)^{\epsilon} p^{(1-s)}(x) q^s(x) dx := \left\langle \left(\frac{q(x)}{p(x)}\right)^{\epsilon}\right\rangle_s.$$

This expression is generally true, however, we will be assuming the $\epsilon \ll 1$ when used in the numerical homotopy procedure.

2. Numerical Approximation of the Continuous Dynamic

Let $s_m := m\epsilon$, $m = 1, ..., M$, and $\epsilon = 1/M$. We can write Z_1/Z_0 as the expanded product of fractions:

$$\frac{Z_1}{Z_0} = \frac{Z_\epsilon}{Z_0}\cdot\frac{Z_{2\epsilon}}{Z_\epsilon}\cdots\frac{Z_1}{Z_{(M-1)\epsilon}} = \prod_{m=1}^{M}\frac{Z_{m\epsilon}}{Z_{(m-1)\epsilon}} = \prod_{m=1}^{M}\left\langle\left(\frac{q(x)}{p(x)}\right)^{\epsilon}\right\rangle_{(m-1)\epsilon}$$

$$\ln\left(\frac{Z_1}{Z_0}\right) \approx \sum_{m=1}^{M}\ln\left(\frac{1}{N}\sum_{n=1}^{N}\left(\frac{q[X(n)_{(m-1)}])}{p([X(n)_{(m-1)}]}\right)^{\epsilon}\right),$$

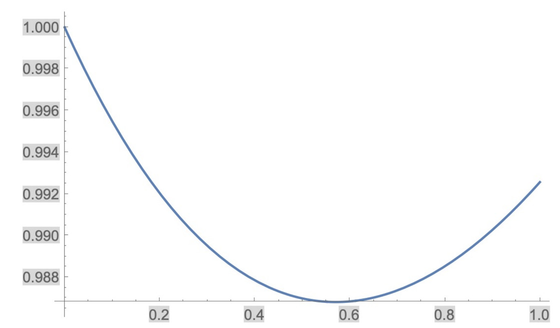
where the n samples

$$[X(n)_{(m-1)}] \sim \frac{1}{Z_{(m-1)\epsilon}} q^{(m-1)\epsilon}(x) p^{1-(m-1)\epsilon}(x).$$

the $(m-1)^{th}$ distribution (known).
Gaussian Example:
Find $Z_1 = \int_{-\infty}^{\infty} q(x)\,dx$, where $q(x) = \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_q^2}\right]$, via homotopy. Starting pdf: $p(x) = \frac{1}{\sqrt{4\pi\sigma_q^2}}\exp[-\frac{(x-\mu_q)^2}{2\sigma_q^2}]$,

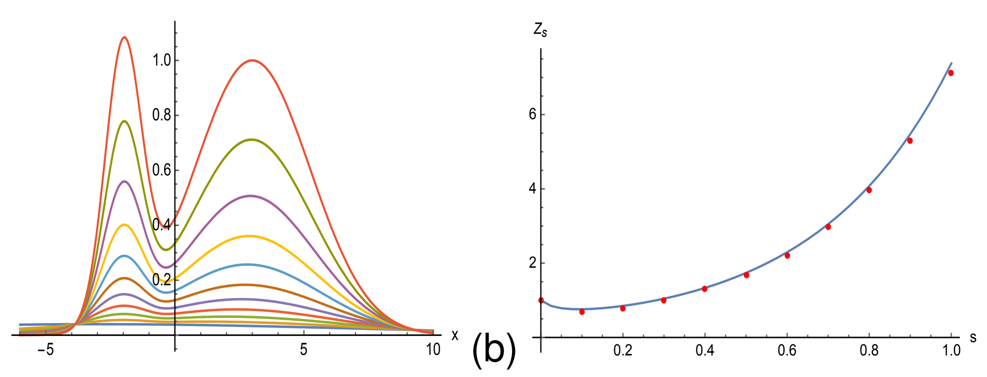
$$Z_s = \frac{2^s\pi^{s/2}\sigma_p^{s-1}\sigma_q\sigma_p}{w(s)}\exp\left(-\frac{1}{4}\frac{(\mu_q-\mu_p)^2(s-1)s}{w(s)}\right).$$

where $w = t\sigma_p^2 + (s-1)\sigma_q^2$, and $0 \le s \le 1$. Analytically, $Z_1 = \frac{1}{\sqrt{4\pi\sigma_q^2}}$. For the case $\mu_q = \mu_p = 0$, the Figure shows $Z(s)$:



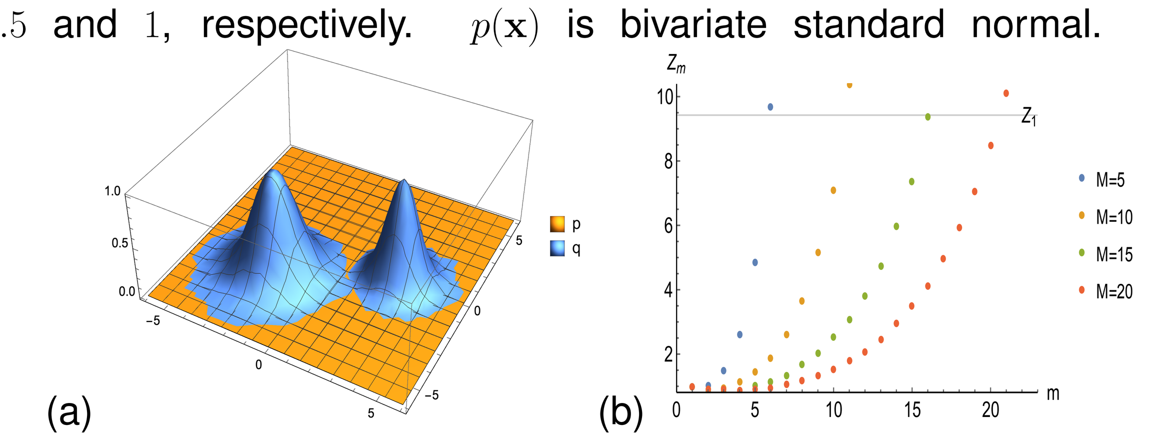
Plot of $Z(s)$, with $\mu_q = \mu_p = 0$ and $\sigma_q = 0.1$ and $\sigma_p = 0.2$. Comparison of analytical and numerical approximation to $Z(s)$.

Bimodal Example:



(a) Evolution of the pdf during the homotopy, from $p(x) = \exp[-(x-3)^2/100]/\sqrt{200\pi}$, to target $q(x) = (\exp[0.1(x-3)^2] + \exp[-(x+2)^2])/Z_1$; (b) Exact and estimated $Z(s)$. $N = 100$, $M = 10$.

Finding Z_1 for a Bivariate Gaussian: Target $q(x)$ is bivariate distribution with maximas at $(-2,2)$ and $(3,0)$, and widths 0.5 and 1, respectively. $p(\mathbf{x})$ is bivariate standard normal.



(a) $p(x)$ and $q(\mathbf{x})$; (b) plot of $Z(s)$, as a function of the number of homotopy steps M. $N = 30$.

3. BAYESIAN HOMOTOPY: Find $Z_1 = \int p(x|y)\,dx$

Homotopy without Updates Choose $p(x)$ *either* $\pi(x)$ prior or $\pi(y|x)$. Choice is dictated by requiring support of $p(x)$ greater than $q(x)$ and knowledge of Z_0.

$$Z_s = \int [q(x)]^s p(x) dx.$$

If $q(x) = \pi(y|x)$ then .$p(x) = \pi(x)/Z_0$. If $q(x) = \pi(x)$ then $p(x) = \pi(y|x)/Z_0$.

Homotopy with Updates If neither the prior or likelihood are known, use an importance distribution $p = I(x)$, for which $Z_0 = \int I(x)\,dx$ is known. The

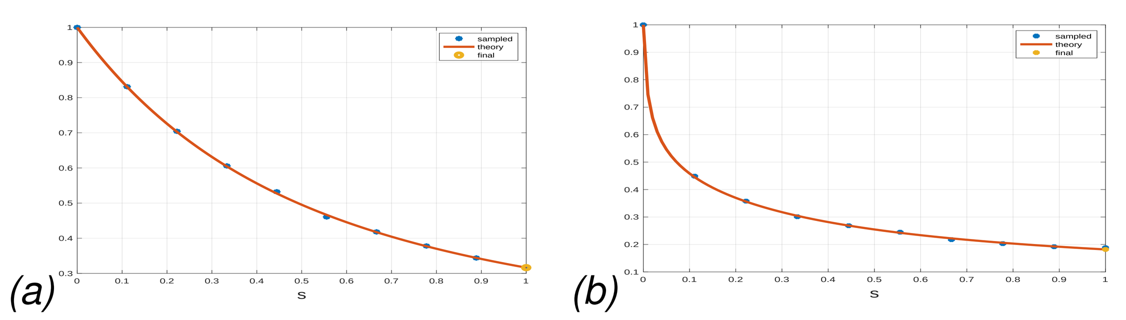
$$Z_s = \int\left[\frac{\pi(x)\pi(y|x)}{I(x)}\right]^s I(x)dx.$$

The samples are now drawn from $I(x)$.

Bayesian Examples *Find* $Z_1 = \int \pi(x)\pi(y|x)\,dx$, *where* $\pi(x)\exp[(x-y)^2/2Q^2]$ *and* $\pi(x) = \frac{x}{R^2}\exp[-x^2/2R^2]$, *a Rayleigh distribution. For this case*

$$Z_s = \frac{\pi}{2}\frac{QRys}{(Q^2+R^2s)^{3/2}}\exp\left[\frac{sy^2}{2(Q^2+R^2s)}\right]\left(2-\mathcal{P}\left[-\frac{1}{2},\frac{(R^2s^2Y^2)}{2W}\right]\right),$$

$W = (Q^4 + Q^2R^2s)$, *where $\mathcal{P}$ is the regularized Gamma function.*



Analytical and numerical comparisons: (a) $Z(s)$ corresponding to to a Rayleigh distribution; (b) $Z(s)$ for case with $\pi(x)$ Gaussian and $\pi(y|x)$ a χ^2 distribution.

4. Markovian Homotopy Data Assimilation

Find $Z_1(k = 0 : T)$, of time-discrete posterior distribution

$$\pi(x_{0:T}|y_{1:T}) \propto \pi(y_{1:T}|x_{1:T})\pi(x_{0:T}).$$

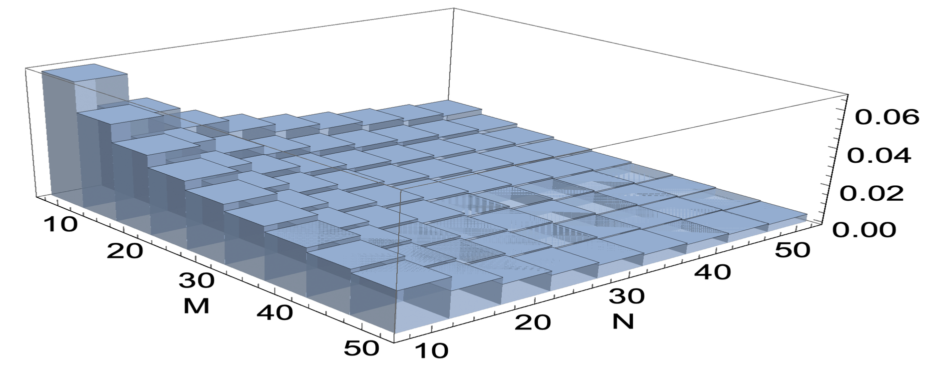
Assuming $Z_1(k-1)$ is known, to find $Z_1(k)$, let $Z_0(k) = Z_1(k-1)$, known. In this case, for $s \in [0,1]$,

$$Z_s(k) = \int\left[\frac{\pi(x_k|x_{k-1})\pi(y_{1:T}|x_k))}{I(x)}\right]^s I(x)dx.$$

where $I(x) = \pi(x_{k-1}|y_{1:k-1})/Z_1(k-1)$

5. Computational Complexity

Homotopy Sampling Replaces reduces the number of N samples required for the sample averages for M the number of homotopy steps.



$\ln[|Z(M,N) - Z_1|]$, *$Z(M,N)$ is estimate of Z_1, homotopy steps: M and samplies used: N.*

Typically $M = \mathcal{O}(10)$, whereas N is large.

6. APPLICATIONS OF THE METHOD

- Sampling
- Canonical Partition Ensemble Calculations
- Data-informed Sample Generation
- Model-informed Sample Generation
- Stochastic and Statistical Emulators

7. SUMMARY

- *We develop a computational method to estimate $Z_1 = \int q(x)\,dx$, via homotopy continuation, generating Z_s, $s \in [0,1]$, starting with $Z_0 = \int p(x)\,dx$, known.*
- *When implemented numerically the method estimates Z_1 using M steps of homotopy and N sample averages. The total computational complexity is $\mathcal{O}(MN)$ and requires no additional storage.*
- *The discretized version delivers Z_1 with a cost comparable to MC, however, it is more efficient when the sample distribution can take advantage of importance sampling.*

FUNDING: *We received financial support from NSF DMS grant 0304890.*