Investigation of Monte Carlo methods for estimating Matrix Determinants

David Young
Supervisor: Mike Peardon

September 16, 2019

Trinity College Dublin
The University of Dublin

1 Abstract

My project was an investigation into different integral representaions for estimating determinants with Monte Carlo methods. I was concerned with how much noise is present in data produced by representations utitilsing different distributions. My focus was on sparse, positive-definite matrices. My hope was to find alternative representaions that produce a smaller error in the estimates than the Gaussian Distribution method which already exists.

2 Introduction

Monte Carlo methods are a broad class of of computational algorithms that uses repeated sampling to obtain numerical results. The methds I am interested rely on manipulation of integral representations so that the integrand is the product of some known distribution which vectors in $\mathbb{R}^n$ can be sampled from, and some function of vectors in $\mathbb{R}^n$. By generating a significantly large number of samples from the distribution, and averaging the results of the function evaluated at each sample, an estimate for the integral can be found.

3 Method for Finding Representations

I was interested in representations of the same form as the existing Gaussian method for an $n \times n$ matrix M;

$$\det M = \int_{\mathbb{R}^n} \frac{e^{-\frac{1}{2}x^t(M^2)^{-1}x}}{(\sqrt{2\pi})^n} \prod_{i=0}^{n}(dx_i)$$

This is derived by starting with the joint distribution of n Standrad Normal Distributions;

$$1 = \int_{\mathbb{R}^n} \frac{e^{-\frac{1}{2}x^t x}}{(\sqrt{2\pi})^n} \prod_{i=0}^{n}(dx_i)$$

We then make a substitution of $M^{-1}x$ for x, giving us the detrminant of the Jacobian M^{-1} in the integrand;

$$1 = \int_{\mathbb{R}^n} \frac{e^{-\frac{1}{2}(M^{-1}x)^t M^{-1}x}}{(\sqrt{2\pi})^n} \det M^{-1} \prod_{i=0}^{n}(dx_i)$$

Multiplying across by $\det M$ we get the desired integral representation, which can be seperated out to;

$$\det M = \int_{\mathbb{R}^n} \frac{e^{-\frac{1}{2}x^t x}}{(\sqrt{2\pi})^n} \prod_{i=0}^{n}(dx_i)\; e^{-\frac{1}{2}x^t((M^2)^{-1}-I)x}$$

Then letting $X \sim \mathcal{N}(0, I)$ and letting $f(x) = e^{-\frac{1}{2}x^t((M^2)^{-1}-I)x}$ we have;

$$\det M = E[f(X)]$$

And therefore we have for large N and $x_i \sim X$;

$$\det M \approx \frac{1}{N}\sum_{i=0}^{N} f(x_i)$$

In this way Monte Carlo sampling on the Standard Normal Distribution can be used to estiamte the determinant.

The issue of trying to compute the inverse of the matrix M can also be avoided by for each x sampled from the distribution, a y is computed such that $My = x$. This can be done with iterative solvers, such as conjugate gradient.

4 Integral Representaions using Alternative Distributions

Using the above method, hereafter refered to as the Jacobain method, on suitable distributions, I was able to derive several alternative representations.

4.1 Exponetial Reflected across Y-Axis

The first alternative we looked at was the exponential distribution with parameter 1, reflected across the Y-axis as so to be defined on the whole real line. The joint density function over $\mathbb{R}^n$ is then;

$$f(x) = \frac{1}{2^n} e^{-\sum\limits_{i=1}^{n} |x_i|}$$

Using the Jacobian method we then obatain an integral representation for $\det M$ of;

$$\det M = \int_{\mathbb{R}^n} \frac{1}{2^n} e^{-\sum\limits_{i=1}^{n} |x_i|} \prod_{i=0}^{n} (dx_i) e^{-\sum\limits_{i=1}^{n} (|I_i^t M^{-1} x| + |x_i|)}$$

4.2 Radial Exponential Distribution

We then altered this distribution again to one in which the points are distributed uniformly in any direction in $\mathbb{R}^n$ and their radial distance is distributed exponentialy. The density function for this distribution is then;

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{2\pi^{\frac{n+1}{2}}} \frac{e^{-|x|}}{|x|^{n-1}}$$

Using the Jacobian method we then obtain an integral representation for $\det M$ of;

$$\det M = \int_{\mathbb{R}^n} \frac{\Gamma(\frac{n+1}{2})}{2\pi^{\frac{n+1}{2}}} \frac{e^{-|x|}}{|x|^{n-1}} \prod_{i=0}^{n} (dx_i) \frac{|x|^{n-1} e^{-|M^{-1}x| + |x|}}{|M^{-1}x|^{n-1}}$$

4.3 Radial Weibull Distribution

The final distribution we tried was the same as above excpet with the radial distance having a Weibull Dsitribution with paramaters $(\frac{1}{2}, 1)$. The density function for this distribution is then;

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{4\pi^{\frac{n+1}{2}}} \frac{e^{-|x|^{\frac{1}{2}}}}{|x|^{n-\frac{1}{2}}}$$

Using the Jacobian method we then obtain an integral representation for $\det M$ of;

$$\det M = \int_{\mathbb{R}^n} \frac{\Gamma(\frac{n+1}{2})}{4\pi^{\frac{n+1}{2}}} \frac{e^{-|x|^{\frac{1}{2}}}}{|x|^{n-\frac{1}{2}}} \prod_{i=0}^{n}(dx_i) \frac{|x|^{n-\frac{1}{2}} e^{-|M^{-1}x|^{\frac{1}{2}}+|x|^{\frac{1}{2}}}}{|M^{-1}x|^{n-\frac{1}{2}}}$$

5 Testing of Representations

The matrices we wanted to test on were positive-definite matrices of the form;

$$M = \begin{bmatrix} 1 & -\alpha & 0 & \dots & 0 & 0 & -\alpha \\ -\alpha & 1 & -\alpha & 0 & \dots & 0 & 0 \\ 0 & -\alpha & 1 & -\alpha & \ddots & \dots & 0 \\ \vdots & 0 & -\alpha & 1 & \ddots & & \vdots \\ 0 & \vdots & \ddots & \ddots & \ddots & -\alpha & 0 \\ 0 & 0 & \vdots & & -\alpha & 1 & -\alpha \\ -\alpha & 0 & 0 & \dots & 0 & -\alpha & 1 \end{bmatrix}$$

For each of the above representations I built Python programmes that estimated the determinant of M for values of incrementaly increasing values of alpha that kept M positive definite. Each programme generated one million samples from the desired distributions to compute the Monte Carlo estimate for each value of alpha, and also computed the standard error for each value of alpha.

The nature of the representations meant that it was very easy to adapt them to compute the inverse of the determinant, so I gathered data on these estimates too,

I tested the programmmes for matrices of sizes from 2x2 up to 100x100, so as to keep the analytic computation of the determinant quick. For each size I graphed the errors and estimates produced by each representation against alpha.

6 Results

One of the immediate results apparent from the data gathered was that the standard error of the Gaussian Method blows up beyond a certain alpha value. This can be seen below in the 2x2 case.

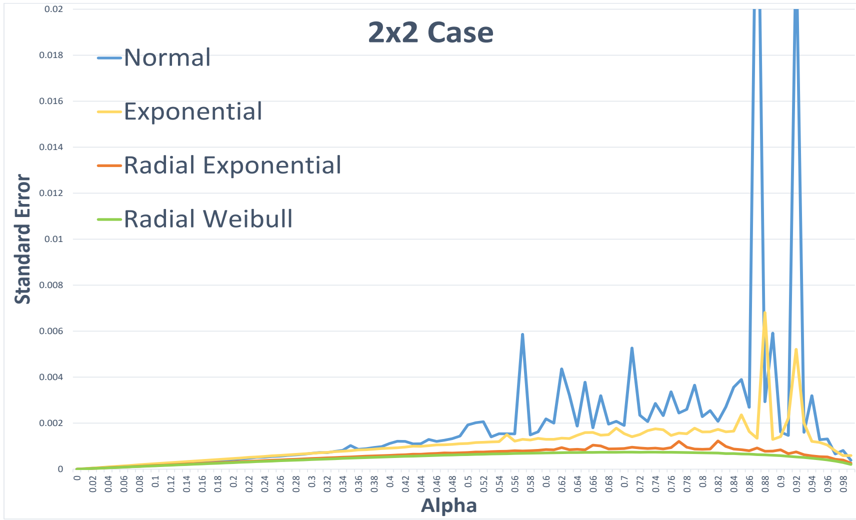

Figure 1: Graph of standard error for 2x2 case

After noticing this we were able to explain it analytically by using the variance of the Gausssian representation method.

$$\mathrm{Var}[f(X)] = E[(f(X))^2] - E[f(X)]^2 = \int_{\mathbb{R}^n} \frac{e^{-\frac{1}{2}x^t(2(M^2)^{-1}-I)x}}{(\sqrt{2\pi})^n} \prod_{i=0}^{n}(dx_i) - (\det M)^2$$

And then by comparison to the Gaussian representation we have

$$\mathrm{Var}[f(X)] = \frac{1}{\sqrt{\det\left(2(M^2)^{-1} - I\right)}} - \left(\frac{1}{\det M}\right)^2$$

Therefore when the determinant of $2((M^2)^{-1} - 1)$ is less than or equal to zero the variance does not exist.

The data also shows clearly that the Radial Weibull and Radial Exponential Distributions produced significantly less error than the Normal Distribution in almost all cases.

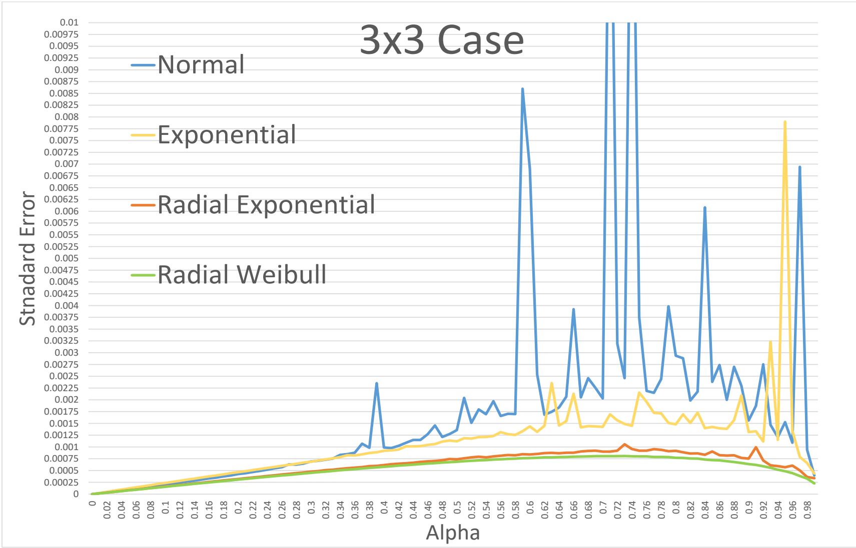

Figure 2: Graph of standard error for 3x3 case

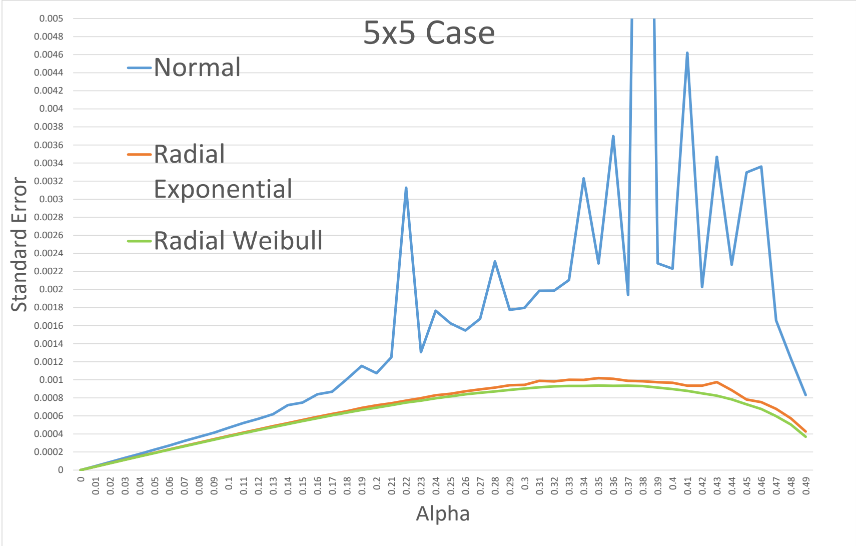

Figure 3: Graph of standard error for 5x5 case

Also for the estiamtor of the inverse of the determinant the Radial Weibull and Radial Exponential Distributions converged to the analytically computed result when the Normal did not.

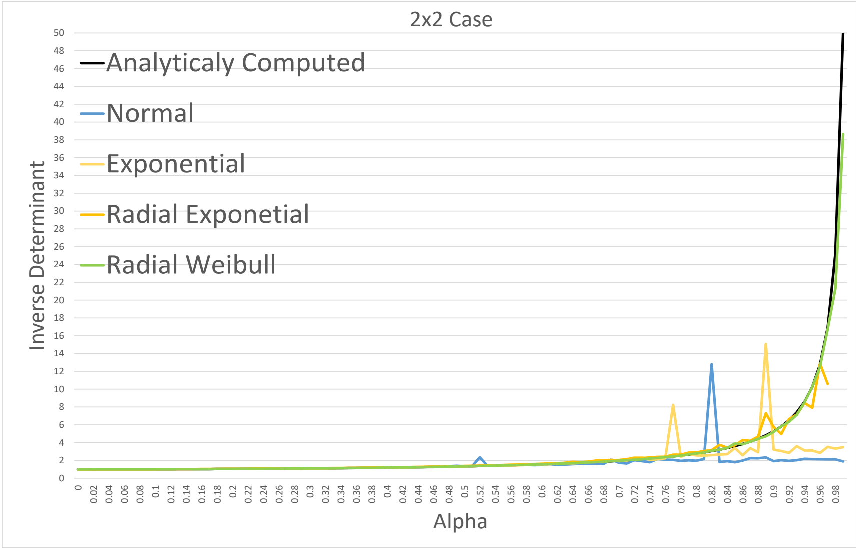

Figure 4: Graph of inverse determinant for 2x2 case

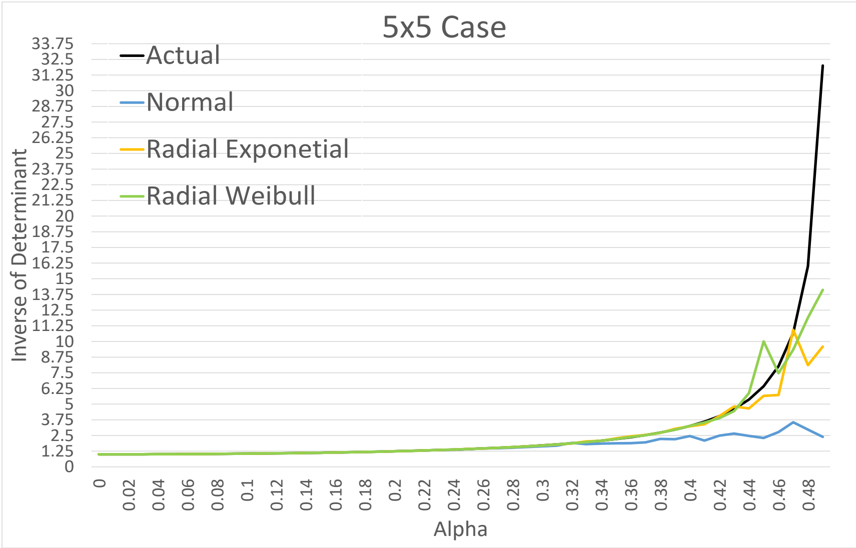

Figure 5: Graph of inverse determinant for 5x5 case

7 Acknowledgements

I would like to thank Professor Mike Peardon for his supervision and advice throughout the project.