Bayesian Updates

December 20, 2022

Chapter 1

Preliminaries

1.1 Circular Distributions

I'm anticipating that I might need to put more words into this later on, so am leaving space for them here.

1.1.1 Von Mises distribution

The Von-Mises distribution is given by:

$$f(x, \mu, \kappa) = \frac{1}{2\pi I_0 \kappa} \exp(\kappa \cos(x - \mu)), \quad -\pi \leq x \leq \pi,$$

where $I_0(\cdot)$ is the 0th modified Bessel function, where the nth modified Bessel function is given by

$$I_n(\kappa) = \frac{1}{\pi} \int_0^\pi \cos(n\theta) \exp(\kappa \cos\theta) \mathrm{d}\theta.$$

The circular mean of the Von-Mises distribution is given by:

$$\mathbb{E}\left[\exp i\theta\right] = \frac{I_1(\kappa)}{I_0(\kappa)} e^{i\mu}.$$

In general, this can be seen via

$$\begin{aligned}
\mathbb{E}[e^{in\theta}] &= \frac{1}{2\pi I_0\left(\kappa\right)} \int_{-\pi}^{\pi} \exp\left(in\theta\right) \exp\left(\kappa \cos(\theta - \mu)\right) \mathrm{d}\theta \\
&= \frac{1}{2\pi I_0\left(\kappa\right)} \int_{-\pi-\mu}^{\pi-\mu} \exp(in(\psi + \mu)) \exp\left(\kappa \cos\psi\right) \mathrm{d}\psi \\
&= \frac{e^{in\mu}}{2\pi I_0\left(\kappa\right)} \int_{-\pi}^{\pi} \exp\left(in\theta\right) \exp\left(\kappa \cos\theta\right) \mathrm{d}\theta \\
&= \frac{e^{in\mu}}{2\pi I_0(\kappa)} \int_{-\pi}^{\pi} \left(\cos(n\theta) + i\sin(n\theta)\right) \exp\left(\kappa \cos\theta\right) \mathrm{d}\theta \\
&= \frac{e^{in\mu}}{2\pi I_0\left(\kappa\right)} \int_{-\pi}^{\pi} \cos\left(n\theta\right) \exp\left(\kappa \cos\theta\right) \mathrm{d}\theta \\
&= \frac{I_{|n|}(\kappa)}{I_0(\kappa)} e^{in\mu}.
\end{aligned}$$

Note that we remove the sin integral by using the fact that the integral of an odd function over a symmetric, periodic interval is 0.

Chapter 2

Problem Statement

2.1 Setup

Goal: Given a single measurement of a Bernoulli random variable and a Von-Mises prior distribution, calculate the posterior distribution and approximate to a Von-Mises distribution.

- t - time step
- d_t - Grover depth of quantum circuit at time t
- Y_t - random variable representing a single shot measurement y_t of the quantum circuit at time t
- $\Pi(\theta|Y_1 = y_1, \ldots, Y_t = y_t) = \Pi(\theta|\mathbf{Y}_t)$ - 'true' posterior at time t (though values for $t' < t$ have been used to approximate the earlier distributions)
- $\hat{\Pi}(\theta|Y_1 = y_1, \ldots, Y_t = y_t) = \hat{\Pi}(\theta|\mathbf{Y}_t)$ - approximate posterior at time t.

According to Bayes rule:

$$\Pi(\theta|Y_t = y_t, \mathbf{Y}_{t-1}) = \frac{\Pi(Y_t = y_t|\theta)\Pi(\theta|\mathbf{Y}_{t-1})}{\Pi(Y_t = y_t)},$$

so we need to compute each of the quantities on the RHS.

At time t, we make a measurement y_t of $Y_t \sim \text{Ber}(p_t)$ at a Grover depth of d_t where

$$p_t = \frac{1}{2}(1 - \cos((4d_t + 2)\hat{\mu}_{t-1}).$$

Thus,

$$\Pi(Y_t = y_t|\theta) = \frac{1}{2}(1 + (-1)^{y_t} \cos((4d_t + 2)\hat{\mu}_{t-1})).$$

For convenience, let $\lambda_t = 4d_t + 2$.

To simplify some of the computations, we're going to assert that the posterior follows a Von-Mises distribution after every update, so we calculate the new values $\hat{\mu}_t, \hat{\kappa}_t$ and generate our approximate posterior

$$\hat{\Pi}(\theta|\mathbf{Y}_t) \sim VM(\hat{\mu}_t, \hat{\kappa}_t).$$

2.2 Single shot updates

For simplicity, we're going to consider the first step of the update, which makes things a lot nicer. In this case, we want to know what the circular mean of the posterior distribution is after updating.

- $\Pi(\theta) \sim \mathrm{VM}(\mu, \kappa)$ - prior
- $\Pi(Y|\theta) \sim \mathrm{Ber}(\frac{1}{2}(1 - \cos(\lambda\theta)))$

This gives us:

$$\begin{aligned}
\Pi(Y = y) &= \int_{-\pi}^{\pi} \Pi(Y = y|\theta)\Pi\left(\theta\right) \mathrm{d}\theta \\
&= \frac{1}{2\pi I_0(\kappa)} \int_{-\pi}^{\pi} \frac{1}{2}(1 + (-1)^y \cos(\lambda\theta)) \exp(\kappa \cos(\theta - \mu))\mathrm{d}\theta \\
&= \frac{1}{2\pi I_0(\kappa)} \left(\int_{-\pi}^{\pi} \frac{1}{2} \exp(\kappa \cos(\theta - \mu))\mathrm{d}\theta \right. \\
&\left. + (-1)^y \int_{-\pi}^{\pi} \frac{1}{2} \cos(\lambda\theta) \exp(\kappa \cos(\theta - \mu))\mathrm{d}\theta \right) \\
&= \frac{1}{4\pi I_0\left(\kappa\right)} \left(2\pi I_0\left(\kappa\right) + (-1)^y \int_{-\pi}^{\pi} \frac{e^{i\lambda\theta} + e^{-i\lambda\theta}}{2} \exp\left(\kappa \cos\left(\theta - \mu\right)\right) \mathrm{d}\theta \right) \\
&= \frac{1}{2} \left(1 + (-1)^y \cos(\lambda\mu) \frac{I_\lambda(\kappa)}{I_0(\kappa)} \right)
\end{aligned}$$

where in the penultimate line, we use the expression for the nth circular moment. Putting this all together, and letting

$$C(y, \lambda, \mu, \kappa) = \frac{\frac{1}{2}\frac{1}{2\pi I_0(\kappa)}}{\frac{1}{2}(1 + (-1)^y \cos(\lambda\mu)\frac{I_\lambda(\kappa)}{I_0(\kappa)})} = \frac{1}{2\pi(I_0(\kappa) + (-1)^y \cos(\lambda\mu) I_\lambda(\kappa))}$$

gives

$$\begin{aligned}
\mathbb{E}[e^{i\theta}|Y=y] &= C(y,\lambda,\mu,\kappa)\int_{-\pi}^{\pi} e^{i\theta}(1+(-1)^y\cos(\lambda\theta))\exp(\kappa\cos(\theta-\mu))\mathrm{d}\theta \\
&= C(y,\lambda,\mu,\kappa)\left(\int_{-\pi}^{\pi} e^{i\theta}\exp(\kappa\cos(\theta-\mu))\mathrm{d}\theta\right. \\
&\quad \left. +(-1)^y\int_{-\pi}^{\pi} e^{i\theta}\cos(\lambda\theta)\exp(\kappa\cos(\theta-\mu))\mathrm{d}\theta\right) \\
&= C(y,\lambda,\mu,\kappa)\left(2\pi I_1(\kappa)e^{i\mu}\right. \\
&\quad \left. +(-1)^y\int_{-\pi}^{\pi} e^{i\theta}\left(\frac{e^{i\lambda\theta}+e^{-i\lambda\theta}}{2}\right)\exp(\kappa\cos(\theta-\mu))\mathrm{d}\theta\right) \\
&= 2\pi C(y,\lambda,\mu,\kappa)\left(I_1(\kappa)e^{i\mu}+\frac{(-1)^y}{2}\left(I_{\lambda+1}(\kappa)e^{i(\lambda+1)\mu}+I_{\lambda-1}(\kappa)e^{-i(\lambda-1)\mu}\right)\right)
\end{aligned}$$

where in the penultimate line, we use the fact that

$$\int_{-\pi}^{\pi} e^{in\theta}\exp(\kappa\cos(\theta-\mu))\mathrm{d}\theta = 2\pi I_0(\kappa)\mathbb{E}[e^{in\theta}] = I_{|n|}(\kappa)e^{in\mu}.$$

This gives us that

$$\mathbb{E}[e^{i\theta}|Y=y] = \frac{I_1(\kappa)e^{i\mu}+\frac{(-1)^y}{2}\left(I_{\lambda+1}(\kappa)e^{i(\lambda+1)\mu}+I_{\lambda-1}(\kappa)e^{-i(\lambda-1)\mu}\right)}{I_0(\kappa)+(-1)^y\cos(\lambda\mu)I_\lambda(\kappa)}.$$

If we then take expectations over Y (i.e. multiply by $\Pi(Y=y)$) and sum) this gives us

$$\mathbb{E}[e^{i\theta}] = \frac{I_1(\kappa)}{I_0(\kappa)}e^{i\mu}.$$

So, we can infer that we do not expect the angular parameter μ to move. To infer something about κ, we need to consider $\mathbb{E}[R]$. As before, let's consider $\mathbb{E}[R|Y=y]$. From the above, we can deduce that

$$\begin{aligned}
\mathbb{E}[R|Y=y]^2 &= \frac{\left(I_1+\frac{(-1)^y}{2}\left(e^{i\lambda\mu}I_{\lambda+1}+e^{-i\lambda\mu}I_{\lambda-1}\right)\right)\left(I_1+\frac{(-1)^y}{2}\left(e^{-i\lambda\mu}I_{\lambda+1}+e^{i\lambda\mu}I_{\lambda-1}\right)\right)}{\left(I_0+(-1)^y\cos(\lambda\mu)I_\lambda\right)^2} \\
&= \frac{I_1^2+\frac{1}{4}\left(I_{\lambda+1}^2+I_{\lambda-1}^2\right)+\frac{1}{2}\cos\left(2\lambda\mu\right)I_{\lambda+1}I_{\lambda-1}+(-1)^yI_1\left(I_{\lambda+1}+I_{\lambda-1}\right)\cos\lambda\mu}{\left(I_0+(-1)^y\cos(\lambda\mu)I_\lambda\right)^2} \\
&= \frac{N_y^2}{\left(I_0+(-1)^y\cos(\lambda\mu)I_\lambda\right)^2},
\end{aligned}$$

where for brevity, we have suppressed the argument κ for each of the Bessel functions I_ν.

Calculating $\mathbb{E}[R]$ then, is achieved by multiplying by $\Pi(Y = y)$, square-rooting, and summing. This results in the sum of the square roots of the numerators multiplied by a constant factor of $\frac{1}{2I_0(\kappa)}$, i.e.

$$\begin{aligned}\mathbb{E}[R] &= \frac{N_0 + N_1}{2I_0(\kappa)} \\ &= \frac{1}{2I_0(\kappa)}\sqrt{I_1^2 + \frac{1}{4}\left(I_{\lambda+1}^2 + I_{\lambda-1}^2\right) + \frac{1}{2}\cos(2\lambda\mu)I_{\lambda+1}I_{\lambda-1} + I_1\left(I_{\lambda+1} + I_{\lambda-1}\right)\cos(\lambda\mu)} \\ &+ \frac{1}{2I_0(\kappa)}\sqrt{I_1^2 + \frac{1}{4}\left(I_{\lambda+1}^2 + I_{\lambda-1}^2\right) + \frac{1}{2}\cos(2\lambda\mu)I_{\lambda+1}I_{\lambda-1} - I_1\left(I_{\lambda+1} + I_{\lambda-1}\right)\cos(\lambda\mu)}\end{aligned}$$

2.3 Single shot updates - Gaussian

Now we're going to assume a different prior and posterior:

- $\Pi(\theta) \sim \mathrm{N}(\mu, \sigma^2)$ - prior
- $\Pi(Y = y|\theta) \sim \mathrm{Ber}(\lambda) = \frac{1}{2}(1 + (-1)^y \cos(\lambda\theta))$

Bayes rule, again, states that

$$\Pi'(\theta|Y = y) = \frac{\Pi(\theta)\mathcal{L}(y, \theta)}{\Pi(Y = y)}. \tag{2.1}$$

We assume that

$$\Pi(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\theta-\mu)^2}{2\sigma^2}}, \tag{2.2}$$

we know that

$$\mathcal{L}(y, \theta) = \frac{1}{2}(1 + (-1)^y \cos(\lambda\theta)), \tag{2.3}$$

and by definition

$$\Pi(Y = y) = \int \mathcal{L}(y, \theta)\Pi(\theta)\mathrm{d}\theta. \tag{2.4}$$

Let us now define the *bias* to be

$$\Lambda(\theta) = 2\mathcal{L}(0, \theta) - 1, \tag{2.5}$$

which gives the likelihood as

$$\mathcal{L}(y, \theta) = \frac{1}{2}(1 + (-1)^y \Lambda(\theta)). \tag{2.6}$$

Recognise that in this case $\Lambda(\theta) = \cos(\lambda\theta)$.

Let us define *expected bias* b and the *chi function* χ as

$$b = \int \Pi(\theta)\Lambda(\theta)\mathrm{d}\theta \tag{2.7}$$

$$\chi = \frac{1}{\sigma^2}\int (\theta - \mu)\Pi(\theta)\Lambda(\theta)\mathrm{d}\theta \tag{2.8}$$

Now, putting that in to equation (2.4) gives

$$\Pi(Y=y) = \frac{1}{2}\Big[\int \Pi(\theta)\mathrm{d}\theta + (-1)^y \int \Pi(\theta)\Lambda(\theta)\mathrm{d}\theta\Big]. \tag{2.9}$$

The first part equals 1 by normalisation. The second part is the expected bias by definition, i.e.,

$$\Pi(Y=y) = \frac{1}{2}\Big[1 + (-1)^y b\Big]. \tag{2.10}$$

The expected bias b is given by

$$b = \int \Pi(\theta)\Lambda(\theta)\mathrm{d}\theta \tag{2.11}$$

$$= \int \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\theta-\mu)^2}{2\sigma^2}} \cos(\lambda\theta)\mathrm{d}\theta, \tag{2.12}$$

which apparently ([**?**], equation 153 on page 25) is an 'identity' for $\sigma > 0$ and $\mu, \lambda \in \mathbb{R}$:

$$b = e^{-\frac{1}{2}\lambda^2\sigma^2} \cos(\lambda\mu). \tag{2.13}$$

Putting this all together gives for the posterior:

$$\Pi'(\theta|Y=y) = \frac{\Pi(\theta)\mathcal{L}(y,\theta)}{\Pi(Y=y)} \tag{2.14}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \frac{e^{-\frac{1}{2}(\frac{\theta-\mu}{\sigma})^2}(1 + (-1)^y \cos(\lambda\theta))}{1 + (-1)^y e^{-\frac{1}{2}\lambda^2\sigma^2}\cos(\lambda\mu)}. \tag{2.15}$$

Appendix A

Integrals

A.1 Normal Distribution

First, let us note that

$$\frac{\mathrm{d}}{\mathrm{d}x}(e^{-x^2}) = -2xe^{-x^2}.$$

The integrals we are interested in computing, are either of the form

$$I_{2n}(k) = \int_{-\infty}^{\infty} \theta^{2n} \cos(k\theta) e^{-\theta^2} \mathrm{d}\theta \text{ or } I_{2n+1} = \int_{-\infty}^{\infty} \theta^{2n+1} \sin(k\theta) e^{-\theta^2} \mathrm{d}\theta$$