Bayesian Updates

October 24, 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:

$$\begin{aligned} \mathbb{E}[e^{i\theta}] &= \frac{1}{2\pi I_0(\kappa)} \int_{-\pi}^{\pi} e^{i\theta} \exp(\kappa \cos(\theta - \mu)) \mathrm{d}\theta \\ &= \frac{I_1(\kappa)}{I_0(\kappa)} e^{i\mu} \end{aligned}$$

In particular, since $\frac{I_1(\kappa)}{I_0(\kappa)}$ is an increasing function, the circular mean uniquely defines a Von-Mises distribution.

For higher circular moments,

$$\mathbb{E}[e^{in\theta}] = \frac{I_{|n|}(\kappa)}{I_0(\kappa)} e^{in\mu}$$

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
- $\Pi_(\theta|Y_1 = y_1, \dots, Y_t = y_t) = \Pi(\theta|\mathbf{Y}_t)$ - 'true' posterior at time t (though values for $t' < t$ have been used to approximate the earlier distributions)
- $\hat{\Pi}(\theta|Y_1 = y_1, \dots, Y_t = y_t) = \hat{\Pi}(\theta|\mathbf{Y}_t)$ - 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 $Y_t \sim \mathrm{Ber}(p_t)$ 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

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

Now, via Bayes rule, we're computing

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

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.

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

This gives us:

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

Putting this all together, and letting

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

gives

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

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

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

This is problematic. We're not expecting μ to move anywhere on average, but we're hoping that κ is going to increase ($1/\kappa$ is analogous to σ^2 for a normal distribution.