Bayesian Updates

February 1, 2023

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 $Y_t \sim \mathrm{Ber}(p_t)$ 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 $\tilde{\lambda}_t = 4d_t + 2$ and $\lambda = 2d + 1 = \frac{1}{2}\tilde{\lambda}$.

* Note: actual θ runs from 0 to π. When working with $\tilde{\lambda}$, $\tilde{\theta}$ in the likelihood function has a factor $\frac{1}{2}$, which means $\tilde{\theta}$ runs from 0 to $\pi/2$. I.e., $\tilde{\theta} = \frac{1}{2}\theta$.

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.

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

This gives us:

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

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

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

gives

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

where in the penultimate line, we use the fact that

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

This gives us that

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

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

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

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

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

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.

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

2.3 Gaussian case

2.3.1 Posterior

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

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

Bayes rule, again, states that

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

We assume that

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

we know that

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

and by definition

$$\Pi(Y = y) = \int_{-\infty}^{\infty} \mathcal{L}(y, \tilde{\theta})\Pi(\tilde{\theta})\mathrm{d}\tilde{\theta}. \tag{2.4}$$

Let us now define the *bias* to be

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

which gives the likelihood as

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

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

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

$$b = \int_{-\infty}^{\infty} \Pi(\tilde{\theta})\Lambda(\tilde{\theta})\mathrm{d}\tilde{\theta} \tag{2.7}$$

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

Now, putting that in to equation (2.4) gives

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

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

$$b = \int_{-\infty}^{\infty} \Pi(\tilde{\theta})\Lambda(\tilde{\theta})\mathrm{d}\tilde{\theta} \tag{2.11}$$

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

which apparently ([1], equation 153 on page 25) is an 'identity' for $\sigma > 0$ and $\mu, \tilde{\lambda} \in \mathbb{R}$:

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

Putting this all together gives for the posterior:

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

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

2.3.2 Expected values

Now, we're interested in the following quantities:

- $\mathbb{E}_y(\mathrm{Var}_{\tilde{\theta}}(\tilde{\theta}|Y))$ - The expected posterior variance
- $\mathcal{V}$ - The variance reduction factor
- $\mathbb{E}(\tilde{\theta}|Y=0), \mathrm{Var}(\tilde{\theta}|Y=0)$ - Posterior mean and variance when measure $Y=0$
- $\mathbb{E}(\tilde{\theta}|Y=1), \mathrm{Var}(\tilde{\theta}|Y=1)$ - Posterior mean and variance when measure $Y=1$

Theorem 12 (together with equation (113)) of [1] states that the *expected posterior variance* is given by

$$\mathbb{E}_y(\mathrm{Var}_{\tilde{\theta}}(\tilde{\theta}|Y)) = \sigma^2(1-\sigma^2\mathcal{V}), \tag{2.16}$$

with

$$\mathcal{V} = \frac{1}{4}\left[\sum_{y\in\{0,1\}} \frac{I_1(y)^2}{I_0(y)} - \mu^2\right], \tag{2.17}$$

and with

$$I_k(y) = \int_{-\infty}^{\infty} \tilde{\theta}^k \mathcal{L}(y,\tilde{\theta})\Pi(\tilde{\theta})\mathrm{d}\tilde{\theta} \tag{2.18}$$

the k-th moment of the function $\mathcal{L}(y,\cdot)\Pi(\cdot)$.

Now by writing the expected bias b and chi function χ, equations (2.7) and (2.8) respectively, in terms of the moments, you can show (equations (132-135) from [1]) that for a two-outcome likelihood function, the variance reduction factor can be written as

$$\mathcal{V} = \begin{cases} \frac{\chi^2}{1-b^2}, & |b|<1 \\ 0, & |b|=1. \end{cases} \tag{2.19}$$

The Gaussian prior has a nice property: differentiating the expected bias w.r.t. the prior mean gives the chi function, i.e.

$$\chi(\mu,\sigma) = \frac{\partial}{\partial\mu} b(\mu,\sigma), \tag{2.20}$$

resulting in

$$\chi(\mu,\sigma) = -\tilde{\lambda}e^{-\frac{1}{2}\tilde{\lambda}^2\sigma^2}\sin(\tilde{\lambda}\mu) \tag{2.21}$$

and now the variance reduction factor can be written as

$$\mathcal{V} = \mathcal{V}(\mu,\sigma) = \frac{\partial_\mu b(\mu,\sigma)^2}{1-b(\mu,\sigma)^2}\mathbb{1}_{\Lambda\notin\{\pm 1\}}, \tag{2.22}$$

where $\mathbb{1}_{\Lambda\notin\{\pm 1\}}$ denotes the indicator function which is equal to 1 when $\Lambda \notin \{\pm 1\}$ and 0 otherwise.

Combining the above, gives, together with equation (2.13) for the Gaussian prior:

$$\mathcal{V} = \frac{\tilde{\lambda}^2 e^{-\tilde{\lambda}^2\sigma^2}\sin^2(\tilde{\lambda}\mu)}{1-e^{-\tilde{\lambda}^2\sigma^2}\cos^2(\tilde{\lambda}\mu)}\mathbb{1}_{\Lambda\notin\{\pm 1\}}, \tag{2.23}$$

and thereby the expected posterior variance is

$$\mathbb{E}_y(\mathrm{Var}_{\tilde{\theta}}(\tilde{\theta}|Y)) = \sigma^2(1-\sigma^2\frac{\tilde{\lambda}^2 e^{-\tilde{\lambda}^2\sigma^2}\sin^2(\tilde{\lambda}\mu)}{1-e^{-\tilde{\lambda}^2\sigma^2}\cos^2(\tilde{\lambda}\mu)}\mathbb{1}_{\Lambda\notin\{\pm 1\}}). \tag{2.24}$$

The next quantities of interest are $\mathbb{E}(\tilde{\theta}|Y=y)$ and $\mathrm{Var}(\tilde{\theta}|Y=y)$ for $y \in \{0,1\}$. By definition:

$$\mathbb{E}(\tilde{\theta}|Y=y) = \int_{-\infty}^{\infty} \tilde{\theta}\Pi'(\tilde{\theta}|Y=y)\mathrm{d}\tilde{\theta}, \tag{2.25}$$

and

$$\mathrm{Var}(\tilde{\theta}|Y=y) = \mathbb{E}(\tilde{\theta}^2|Y=y) - (\mathbb{E}(\tilde{\theta}|Y=y))^2, \tag{2.26}$$

with

$$\mathbb{E}(\tilde{\theta}^2|Y=y) = \int_{-\infty}^{\infty} \tilde{\theta}^2 \Pi'(\tilde{\theta}|Y=y)\mathrm{d}\tilde{\theta}. \tag{2.27}$$

Starting with $\mathbb{E}(\tilde{\theta}|Y=y)$, we write

$$\begin{aligned}
\mathbb{E}(\tilde{\theta}|Y=y) &= \frac{1}{\Pi(Y=y)} \int_{-\infty}^{\infty} \tilde{\theta}\Pi(\tilde{\theta})\mathcal{L}(y,\tilde{\theta})\mathrm{d}\tilde{\theta} && (2.28)\\
&= \frac{1}{\Pi(Y=y)} \int_{-\infty}^{\infty} \tilde{\theta}\Pi(\tilde{\theta})\Big(\frac{1}{2}(1+(-1)^y\cos(\tilde{\lambda}\tilde{\theta}))\Big)\mathrm{d}\tilde{\theta} && (2.29)\\
&= \frac{1/2}{\Pi(Y=y)} \left(\int_{-\infty}^{\infty} \tilde{\theta}\Pi(\tilde{\theta})\mathrm{d}\tilde{\theta} + (-1)^y \int_{-\infty}^{\infty} \tilde{\theta}\Pi(\tilde{\theta})\Lambda(\tilde{\theta})\mathrm{d}\tilde{\theta}\right) && (2.30)\\
&= \frac{1/2}{\Pi(Y=y)} \left(\mu + (-1)^y \int_{-\infty}^{\infty} \tilde{\theta}\Pi(\tilde{\theta})\Lambda(\tilde{\theta})\mathrm{d}\tilde{\theta}\right). && (2.31)
\end{aligned}$$

Now by equation (2.8),

$$\chi = \frac{1}{\sigma^2}\left(\int_{-\infty}^{\infty} \tilde{\theta}\Pi(\tilde{\theta})\Lambda(\tilde{\theta})\mathrm{d}\tilde{\theta} - \mu \int_{-\infty}^{\infty} \Pi(\tilde{\theta})\Lambda(\tilde{\theta})\mathrm{d}\tilde{\theta}\right), \tag{2.32}$$

and by using the definition for b (equation (2.7)):

$$\int_{-\infty}^{\infty} \tilde{\theta}\Pi(\tilde{\theta})\Lambda(\tilde{\theta})\mathrm{d}\tilde{\theta} = \sigma^2\chi + \mu b, \tag{2.33}$$

which gives:

$$\begin{aligned}
\mathbb{E}(\tilde{\theta}|Y=y) &= \frac{1/2(\mu + (-1)^y(\sigma^2\chi + \mu b))}{\Pi(Y=y)} && (2.34)\\
&= \frac{\mu + (-1)^y(\sigma^2\chi + \mu b)}{1+(-1)^y b} && (2.35)\\
& && (2.36)
\end{aligned}$$

and with equations (2.13) and (2.21) gives:

$$\mathbb{E}(\tilde{\theta}|Y=y) = \frac{\mu + (-1)^y e^{-\frac{1}{2}\tilde{\lambda}^2\sigma^2}\big(\mu\cos(\tilde{\lambda}\mu) - \sigma^2\tilde{\lambda}\sin(\tilde{\lambda}\mu)\big)}{1+(-1)^y e^{-\frac{1}{2}\tilde{\lambda}^2\sigma^2}\cos(\tilde{\lambda}\mu)}. \tag{2.37}$$

Now for the posterior variance $\mathrm{Var}(\tilde{\theta}|Y=y)$, we only need $\mathbb{E}(\tilde{\theta}^2|Y=y)$, which is

$$\mathbb{E}(\tilde{\theta}^2|Y=y) = \frac{1}{\Pi(Y=y)}\int_{-\infty}^{\infty}\tilde{\theta}^2\Pi(\tilde{\theta})\mathcal{L}(y,\tilde{\theta})\mathrm{d}\tilde{\theta} \tag{2.38}$$

$$= \frac{1}{\Pi(Y=y)}\int_{-\infty}^{\infty}\tilde{\theta}^2\Pi(\tilde{\theta})\Big(\frac{1}{2}(1+(-1)^y\cos(\tilde{\lambda}\tilde{\theta}))\Big)\mathrm{d}\tilde{\theta} \tag{2.39}$$

$$= \frac{1/2}{\Pi(Y=y)}\left(\int_{-\infty}^{\infty}\tilde{\theta}^2\Pi(\tilde{\theta})\mathrm{d}\tilde{\theta} + (-1)^y\int\tilde{\theta}^2\Pi(\tilde{\theta})\Lambda(\tilde{\theta})\mathrm{d}\tilde{\theta}\right), \tag{2.40}$$

and using

$$\int_{-\infty}^{\infty}\tilde{\theta}^2\Pi(\tilde{\theta})\mathrm{d}\tilde{\theta} = \sigma^2+\mu^2, \tag{2.41}$$

this gives

$$\mathbb{E}(\tilde{\theta}^2|Y=y) = \frac{1/2}{\Pi(Y=y)}\left(\sigma^2+\mu^2+(-1)^y\underbrace{\int_{-\infty}^{\infty}\tilde{\theta}^2\Pi(\tilde{\theta})\Lambda(\tilde{\theta})\mathrm{d}\tilde{\theta}}_{\star}\right). \tag{2.42}$$

Let us focus on the last integral, which we define as $\star$:

$$\star = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\tilde{\theta}^2 e^{-\frac{(\tilde{\theta}-\mu)^2}{2\sigma^2}}\cos(\tilde{\lambda}\tilde{\theta})\mathrm{d}\tilde{\theta}, \tag{2.43}$$

which can be written as

$$\frac{1}{2\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\tilde{\theta}^2 e^{-\frac{(\tilde{\theta}-\mu)^2}{2\sigma^2}}\Big(e^{i\tilde{\lambda}\tilde{\theta}}+e^{-i\tilde{\lambda}\tilde{\theta}}\Big)\mathrm{d}\tilde{\theta}, \tag{2.44}$$

and that breaks up the problem in two problems:

$$\frac{1}{2\sqrt{2\pi\sigma^2}}\Big(\underbrace{\int_{-\infty}^{\infty}\tilde{\theta}^2 e^{-\frac{(\tilde{\theta}-\mu)^2}{2\sigma^2}}e^{i\tilde{\lambda}\tilde{\theta}}\mathrm{d}\tilde{\theta}}_{A_0}+\underbrace{\int_{-\infty}^{\infty}\tilde{\theta}^2 e^{-\frac{(\tilde{\theta}-\mu)^2}{2\sigma^2}}e^{-i\tilde{\lambda}\tilde{\theta}}\mathrm{d}\tilde{\theta}}_{B_0}\Big) \tag{2.45}$$

Here, A_0 and B_0 are very similar, except for the sign of the complex part.

Let us simplify the exponent. The full exponent for A_0 is

$$-\frac{(\tilde{\theta}-\mu)^2}{2\sigma^2}+i\tilde{\lambda}\tilde{\theta} \tag{2.46}$$

$$= \frac{-1}{2\sigma^2}(\tilde{\theta}^2-2\tilde{\theta}\mu+\mu^2-2i\tilde{\lambda}\tilde{\theta}\sigma^2) \tag{2.47}$$

$$= \frac{-1}{2\sigma^2}(\tilde{\theta}^2-2(\mu+i\tilde{\lambda}\sigma^2)\tilde{\theta}+\mu^2) \tag{2.48}$$

$$= \frac{-1}{2\sigma^2}\big((\tilde{\theta}-(\mu+i\tilde{\lambda}\sigma^2))^2-(\mu^2+2\mu i\tilde{\lambda}\sigma^2-\tilde{\lambda}^2\sigma^4)+\mu^2\big) \tag{2.49}$$

$$= -\frac{(\tilde{\theta}-(\mu+i\tilde{\lambda}\sigma^2))^2}{2\sigma^2}+\tilde{\lambda}\mu i-\frac{1}{2}\tilde{\lambda}^2\sigma^2 \tag{2.50}$$

where in the penultimate line, we completed the square. Analogously, for B_0 the exponent is

$$-\frac{(\tilde{\theta}-(\mu-i\tilde{\lambda}\sigma^2))^2}{2\sigma^2}-\tilde{\lambda}\mu i-\frac{1}{2}\tilde{\lambda}^2\sigma^2. \tag{2.51}$$

In totally, $\star$ now becomes

$$\frac{1}{2\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2}\tilde{\lambda}^2\sigma^2}\Big(e^{\tilde{\lambda}\mu i}\underbrace{\int_{-\infty}^{\infty}\tilde{\theta}^2e^{-\frac{(\tilde{\theta}-(\mu+i\tilde{\lambda}\sigma^2))^2}{2\sigma^2}}\mathrm{d}\tilde{\theta}}_{A_1}+e^{-\tilde{\lambda}\mu i}\underbrace{\int_{-\infty}^{\infty}\tilde{\theta}^2e^{-\frac{(\tilde{\theta}-(\mu-i\tilde{\lambda}\sigma^2))^2}{2\sigma^2}}\mathrm{d}\tilde{\theta}}_{B_1}\Big). \tag{2.52}$$

We can solve A_1 and B_1 by using substitution:

$$u=\tilde{\theta}-(\mu+i\tilde{\lambda}\sigma^2) \qquad v=\tilde{\theta}-(\mu-i\tilde{\lambda}\sigma^2) \tag{2.53}$$

$$\mathrm{d}u=\mathrm{d}\tilde{\theta} \qquad \mathrm{d}v=\mathrm{d}\tilde{\theta} \tag{2.54}$$

$$\tilde{\theta}^2=u^2+2u(\mu+i\tilde{\lambda}\sigma^2)+(\mu+i\tilde{\lambda}\sigma^2)^2 \quad \tilde{\theta}^2=v^2+2v(\mu-i\tilde{\lambda}\sigma^2)+(\mu-i\tilde{\lambda}\sigma^2)^2 \tag{2.55}$$

Now A_1 is

$$A_1=\int_{-\infty}^{\infty}\tilde{\theta}^2e^{-\frac{(\tilde{\theta}-(\mu+i\tilde{\lambda}\sigma^2))^2}{2\sigma^2}}\mathrm{d}\tilde{\theta} \tag{2.56}$$

$$=\int_{-\infty}^{\infty}(u^2+2u(\mu+i\tilde{\lambda}\sigma^2)+(\mu+i\tilde{\lambda}\sigma^2)^2)e^{-\frac{u^2}{2\sigma^2}}\mathrm{d}\tilde{\theta} \tag{2.57}$$

$$=\underbrace{\int_{-\infty}^{\infty}u^2e^{-\frac{u^2}{2\sigma^2}}\mathrm{d}u}_{=\sqrt{2\pi\sigma^6}}+2(\mu+i\tilde{\lambda}\sigma^2)\underbrace{\int_{-\infty}^{\infty}ue^{-\frac{u^2}{2\sigma^2}}\mathrm{d}u}_{=0}+(\mu+i\tilde{\lambda}\sigma^2)^2\underbrace{\int_{-\infty}^{\infty}e^{-\frac{u^2}{2\sigma^2}}\mathrm{d}u}_{=\sqrt{2\pi\sigma^2}} \tag{2.58}$$

$$=\sqrt{2\pi\sigma^6}+(\mu+i\tilde{\lambda}\sigma^2)^2\sqrt{2\pi\sigma^2} \tag{2.59}$$

$$=\sqrt{2\pi\sigma^2}(\sigma^2+(\mu+i\tilde{\lambda}\sigma^2)^2) \tag{2.60}$$

and for B_1

$$B_1=\sqrt{2\pi\sigma^2}(\sigma^2+(\mu-i\tilde{\lambda}\sigma^2)^2). \tag{2.61}$$

The prefactor $\sqrt{2\pi\sigma^2}$ cancels with the prefactor in $\star$ to give

$$\star=\frac{1}{2}e^{-\frac{1}{2}\tilde{\lambda}^2\sigma^2}\Big(e^{\tilde{\lambda}\mu i}(\sigma^2+(\mu+i\tilde{\lambda}\sigma^2)^2)+e^{-\tilde{\lambda}\mu i}(\sigma^2+(\mu-i\tilde{\lambda}\sigma^2)^2)\Big). \tag{2.62}$$

We can do some accounting to arrive at the final form:

$$\star = \frac{1}{2}e^{-\frac{1}{2}\tilde{\lambda}^2\sigma^2}\Big(e^{\tilde{\lambda}\mu i}(\sigma^2+\mu^2-\tilde{\lambda}^2\sigma^4+2\mu i\tilde{\lambda}\sigma^2)+e^{-\tilde{\lambda}\mu i}(\sigma^2+\mu^2-\tilde{\lambda}^2\sigma^4-2\mu i\tilde{\lambda}\sigma^2)\Big) \tag{2.63}$$

$$= \frac{1}{2}e^{\frac{1}{2}\tilde{\lambda}^2\sigma^2}\Big((\sigma^2+\mu^2-\tilde{\lambda}^2\sigma^4)(\underbrace{e^{\tilde{\lambda}\mu i}+e^{-\tilde{\lambda}\mu i}}_{=2\cos\tilde{\lambda}\mu})+2\tilde{\lambda}\mu\sigma^2(\underbrace{ie^{\tilde{\lambda}\mu i}-ie^{-\tilde{\lambda}\mu i}}_{=2\sin\tilde{\lambda}\mu})\Big) \tag{2.64}$$

$$= \sigma^2 e^{-\frac{1}{2}\tilde{\lambda}^2\sigma^2}\Big((1+\frac{\mu^2}{\sigma^2}-\tilde{\lambda}^2\sigma^2)\cos\tilde{\lambda}\mu+2\tilde{\lambda}\mu\sin\tilde{\lambda}\mu\Big). \tag{2.65}$$

And with equation (2.13) and (2.21), we can also write it as

$$\star = \sigma^2\Big((1+\frac{\mu^2}{\sigma^2}-\tilde{\lambda}^2\sigma^2)b-2\mu\chi\Big) \tag{2.66}$$

Plugging this back in equation (2.42) gives for the second moment,

$$\mathbb{E}(\tilde{\theta}^2|Y=y) = \frac{\sigma^2+\mu^2+(-1)^y\sigma^2\Big((1+\frac{\mu^2}{\sigma^2}-\tilde{\lambda}^2\sigma^2)b-2\mu\chi\Big)}{1+(-1)^y b}, \tag{2.67}$$

or when fully expanded:

$$\mathbb{E}(\tilde{\theta}^2|Y=y) = \frac{\sigma^2+\mu^2+(-1)^y\sigma^2 e^{-\frac{1}{2}\tilde{\lambda}^2\sigma^2}\Big((1+\frac{\mu^2}{\sigma^2}-\tilde{\lambda}^2\sigma^2)\cos\tilde{\lambda}\mu+2\tilde{\lambda}\mu\sin\tilde{\lambda}\mu\Big)}{1+(-1)^y e^{-\frac{1}{2}\tilde{\lambda}^2\sigma^2}\cos(\tilde{\lambda}\mu)}. \tag{2.68}$$

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

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

Bibliography

[1] D. E. Koh, G. Wang, P. D. Johnson, and Y. Cao. Foundations for bayesian inference with engineered likelihood functions for robust amplitude estimation. *Journal of Mathematical Physics*, 6 2022.