Consider the model $y_i = \mu + \beta \sin(x_i + \eta) + \epsilon_i$ where $x_i = \frac{2\pi i}{n}$ for $i = 0, 1, \ldots, n$, $\epsilon_i \sim \mathsf{N}(0, 1)$ independently for each i, and $\eta \in [0, \pi]$.

1. (25 pts) Assume for now that η is known. What is the maximum likelihood estimator for (μ, β)?

   Solution: The model can be written as $\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon}$, where

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \qquad \mathbf{X} = \begin{pmatrix} 1 & \sin(x_0 + \eta) \\ \vdots & \vdots \\ 1 & \sin(x_n + \eta) \end{pmatrix} \qquad \boldsymbol{\theta} = \begin{pmatrix} \mu \\ \beta \end{pmatrix} \qquad \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

   The MLE is simply $\boldsymbol{\theta} = \left\{\mathbf{X}^T\mathbf{X}\right\}^{-1} \mathbf{X}^T\mathbf{y}$ with

$$\mathbf{X}^T\mathbf{X} = \begin{pmatrix} n & \sum_{i=0}^n \sin(x_i + \eta) \\ \sum_{i=0}^n \sin(x_i + \eta) & \sum_{i=0}^n \sin^2(x_i + \eta) \end{pmatrix}$$

$$\mathbf{X}^T\mathbf{y} = \begin{pmatrix} \sum_{i=0}^n y_i \\ \sum_{i=0}^n y_i \sin(x_i + \eta) \end{pmatrix}$$

   so that

$$\hat{\mu} = \frac{\left\{\sum_{i=0}^n y_i\right\}\left\{\sum_{i=0}^n \sin^2(x_i + \eta)\right\} - \left\{\sum_{i=0}^n y_i \sin(x_i + \eta)\right\}\left\{\sum_{i=0}^n \sin(x_i + \eta)\right\}}{n\sum_{i=0}^n \sin^2(x_i + \eta) - \left\{\sum_{i=0}^n \sin(x_i + \eta)\right\}^2}$$

   and

$$\hat{\beta} = \frac{n\left\{\sum_{i=0}^n y_i \sin(x_i + \eta)\right\} - \left\{\sum_{i=0}^n y_i\right\}\left\{\sum_{i=0}^n \sin(x_i + \eta)\right\}}{n\sum_{i=0}^n \sin^2(x_i + \eta) - \left\{\sum_{i=0}^n \sin(x_i + \eta)\right\}^2}.$$

2. (35 pts) Assume now that η is unknown. Compute the maximum likelihood estimator for (μ, β, η) (Hint: You can use simple trigonometric identities to write $\beta \sin(x_i + \eta) = \beta \sin x_i \cos \eta + \beta \cos x_i \sin \eta$. Also, note that $\sum_{i=0}^n \sin x_i = 0$ and $\sum_{i=0}^n \sin x_i \cos x_i = 0$.)

   Solution: The proposed alternative representation is valid because of the well-known trigonometric identity $\sin(\alpha + \gamma) = \sin\alpha\cos\gamma + \cos\alpha\sin\gamma$. Note that the model can now be written as

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \qquad \mathbf{X} = \begin{pmatrix} 1 & \sin x_0 & \cos x_0 \\ \vdots & \vdots & \vdots \\ 1 & \sin x_n & \cos x_n \end{pmatrix} \qquad \boldsymbol{\theta} = \begin{pmatrix} \mu \\ \phi_1 \\ \phi_2 \end{pmatrix} \qquad \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

where $\phi_1 = \beta \cos \eta$ and $\phi_2 = \beta \sin \eta$. In this case

$$\mathbf{X}^T\mathbf{X} = \begin{pmatrix} n & \sum_{i=0}^n \sin x_i & \sum_{i=0}^n \cos x_i \\ \sum_{i=0}^n \sin x_i & \sum_{i=0}^n \sin^2 x_i & \sum_{i=0}^n \sin x_i \cos x_i \\ \sum_{i=0}^n \cos x_i & \sum_{i=0}^n \sin x_i \cos x_i & \sum_{i=0}^n \cos^2 x_i \end{pmatrix},$$

$$\mathbf{X}^T\mathbf{y} = \begin{pmatrix} \sum_{i=0}^n y_i \\ \sum_{i=0}^n y_i \sin x_i \\ \sum_{i=0}^n y_i \cos x_i \end{pmatrix}.$$

They can proceed with this general form, but it is easier to note that $\sum_{i=0}^n \sin x_i = 0$ and $\sum_{i=0}^n \sin x_i \cos x_i = 0$ (since both are odd functions around π and the x_is are equally spaced), so that

$$\mathbf{X}^T\mathbf{X} = \begin{pmatrix} n & 0 & \sum_{i=0}^n \cos x_i \\ 0 & \sum_{i=0}^n \sin^2 x_i & 0 \\ \sum_{i=0}^n \cos x_i & 0 & \sum_{i=0}^n \cos^2 x_i \end{pmatrix}.$$

Now, if we work with (ϕ_1, μ, ϕ_2) instead of (μ, ϕ_1, ϕ_2) we obtain a block diagonal matrix that is easier to invert and

$$\hat{\mu} = \frac{\left\{\sum_{i=0}^n y_i\right\}\left\{\sum_{i=0}^n \cos^2 x_i\right\} - \left\{\sum_{i=0}^n y_i \cos x_i\right\}\left\{\sum_{i=0}^n \cos x_i\right\}}{n\sum_{i=0}^n \cos^2 x_i - \left\{\sum_{i=0}^n \cos x_i\right\}^2},$$

$$\hat{\phi_1} = \frac{\sum_{i=0}^n y_i \sin x_i}{\sum_{i=0}^n \sin^2 x_i},$$

and

$$\hat{\phi_2} = \frac{n\left\{\sum_{i=0}^n y_i \cos x_i\right\} - \left\{\sum_{i=0}^n y_i\right\}\left\{\sum_{i=0}^n \cos x_i\right\}}{n\sum_{i=0}^n \cos^2 x_i - \left\{\sum_{i=0}^n \cos x_i\right\}^2}.$$

Because of invariance, the MLE of η is simply $\hat{\eta} = \arctan\left\{\frac{\hat{\phi}_2}{\hat{\phi}_1}\right\}$.

3. (40 pts) Assume that you are interested in testing the hypotheses $H_0 : \eta = \pi/4$ vs. $H_a : \eta \neq \pi/4$. Derive the likelihood ratio test for this this pair of hypotheses. Please keep in mind that we are assuming that the variance is known!

   Solution: Under the null, the model is equivalent to setting $\phi_1 = \phi_2 = \beta$. Hence, testing $H_0 : \eta = \pi/4$ vs. $H_a : \eta \neq \pi/4$ is equivalent to testing the general linear hypotheses:

$$H_0 : (1, 0, -1)\begin{pmatrix}\phi_1 \\ \mu \\ \phi_2\end{pmatrix} = 0 \qquad \text{vs.} \qquad H_a : (1, 0, -1)\begin{pmatrix}\phi_1 \\ \mu \\ \phi_2\end{pmatrix} \neq 0,$$

Note that

$$(1,0,-1)\begin{pmatrix}\hat{\phi}_1\\ \hat{\mu}\\ \hat{\phi}_2\end{pmatrix} = \frac{\sum_{i=0}^n y_i \sin x_i}{\sum_{i=0}^n \sin^2 x_i} - \frac{n\left\{\sum_{i=0}^n y_i \cos x_i\right\} - \left\{\sum_{i=0}^n y_i\right\}\left\{\sum_{i=0}^n \cos x_i\right\}}{n\sum_{i=0}^n \cos^2 x_i - \left\{\sum_{i=0}^n \cos x_i\right\}^2},$$

and that

$$\begin{aligned}(1,0,-1)&\begin{pmatrix}\frac{1}{\sum_{i=0}^n \sin^2 x_i} & 0 & 0\\ 0 & \frac{\sum_{i=0}^n \cos^2 x_i}{n\sum_{i=0}^n \cos^2 x_i - \left\{\sum_{i=0}^n \cos x_i\right\}} & \frac{-\sum_{i=0}^n \cos x_i}{n\sum_{i=0}^n \cos^2 x_i - \left\{\sum_{i=0}^n \cos x_i\right\}}\\ 0 & \frac{-\sum_{i=0}^n \cos x_i}{n\sum_{i=0}^n \cos^2 x_i - \left\{\sum_{i=0}^n \cos x_i\right\}} & \frac{n}{n\sum_{i=0}^n \cos^2 x_i - \left\{\sum_{i=0}^n \cos x_i\right\}}\end{pmatrix}\begin{pmatrix}1\\0\\-1\end{pmatrix}\\ &= \frac{1}{\sum_{i=0}^n \sin^2 x_i} + \frac{n}{n\sum_{i=0}^n \cos^2 x_i - \left\{\sum_{i=0}^n \cos x_i\right\}}.\end{aligned}$$

Therefore

$$U = \frac{\left(\frac{\sum_{i=0}^n y_i \sin x_i}{\sum_{i=0}^n \sin^2 x_i} - \frac{n\left\{\sum_{i=0}^n y_i \cos x_i\right\} - \left\{\sum_{i=0}^n y_i\right\}\left\{\sum_{i=0}^n \cos x_i\right\}}{n\sum_{i=0}^n \cos^2 x_i - \left\{\sum_{i=0}^n \cos x_i\right\}^2}\right)^2}{\frac{1}{\sum_{i=0}^n \sin^2 x_i} + \frac{n}{n\sum_{i=0}^n \cos^2 x_i - \left\{\sum_{i=0}^n \cos x_i\right\}}} \sim \chi_1^2$$

So the test proceeds by computing U_{obs} using the sample and rejecting H_0 if $U_{obs} > \chi_1^2(1-\alpha)$ where $\chi_1^2(1-\alpha)$ denotes the $1-\alpha$ quantile of the chi squared distribution with one degree of freedom.