1. Explain in detail how you would find the fit curve using techniques of this course. (Warning: To check your answer, you might want to make up four data points and work out the solution explicitly.)

Answer: We want to solve

$$\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & \cos x_1 & \cos^2 x_1 & \cos 2x_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \cos x_N & \cos^2 x_N & \cos 2x_N \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}.$$

Let A denote that big $N \times 4$ matrix made from the x_i. From our discussion of least squares we know that if A has rank 4, then $A^\top A$ is invertible and the least squares solution $\left(A^\top A\right)^{-1} A^\top \vec{y}$. However, our A is not of rank 4, because the set of functions $\{1, \cos x, \cos^2 x, \cos 2x\}$ is not linearly independent! For example, $\cos 2x = -1 + 2\cos^2 x$. So let's throw out the $\cos 2x$ term. Now we want to solve

$$\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & \cos x_1 & \cos^2 x_1 \\ \vdots & \vdots & \vdots \\ 1 & \cos x_N & \cos^2 x_N \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

The procedure is to let A be that big $N \times 3$ matrix (which is of rank 3, as needed) and compute the least squares solution as

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \left(A^\top A\right)^{-1} A^\top \vec{y}.$$

This gives us the coefficients a, b, c to use in the fit curve $y = a + b\cos x + c\cos^2 x$. This function satisfies the requirements of the problem, because it is of Ms. Ogunmola's requested form with $d = 0$.

[Remark: I expected that most students would not realize that the four functions are not independent. That is why I suggested working an example; in any example it becomes clear that $A^\top A$ is not invertible (although it may not be clear how to fix this).]

[Remark: Instead of throwing out $\cos 2x$ we could throw out 1 or $\cos^2 x$. We could not throw out $\cos x$ without truly making our wind speed model less expressive.]

2. Show that if $\vec{n}$ is perpendicular to a given polygon, then $\left(A^{-1}\right)^\top \vec{n}$ is perpendicular to the transformed polygon.

Answer: Let $\vec{v}$ be any vector lying in the transformed polygon. We wish to show that $\left(\left(A^{-1}\right)^\top \vec{n}\right) \cdot \vec{v} = 0$. Toward that end, let $\vec{y}_1$ and $\vec{y}_2$ be the points at the head and tail of $\vec{v}$, so that $\vec{v} = \vec{y}_1 - \vec{y}_2$. These points $\vec{y}_1$ and $\vec{y}_2$ are in the transformed polygon, so there must exist points $\vec{x}_1$ and $\vec{x}_2$ in the original polygon such that $A\vec{x}_1 = \vec{y}_1$ and $A\vec{x}_2 = \vec{y}_2$. Because $\vec{x}_1$ and

$\vec{x}_2$ are points in the original polygon, $\vec{x}_1 - \vec{x}_2$ is a vector lying in the original polygon, and so $\vec{n} \cdot (\vec{x}_1 - \vec{x}_2) = 0$. Then, using the basic facts that $\vec{a} \cdot \vec{b} = \vec{a}^\top \vec{b}$ and $(BC)^\top = C^\top B^\top$, we have

$$\begin{aligned}\left(\left(A^{-1}\right)^\top \vec{n}\right) \cdot \vec{v} &= \left(\left(A^{-1}\right)^\top \vec{n}\right)^\top (\vec{y}_1 - \vec{y}_2) \\ &= \vec{n}^\top A^{-1}(A\vec{x}_1 - A\vec{x}_2) \\ &= \vec{n}^\top(\vec{x}_1 - \vec{x}_2) \\ &= \vec{n} \cdot (\vec{x}_1 - \vec{x}_2) \\ &= 0.\end{aligned}$$

3. What happens in the special case when A is a rotation? Explain in detail.

Answer: If A is a rotation, then A preserves the length of any vector, so A is orthogonal. This implies that $A^{-1} = A^\top$, so that $\left(A^{-1}\right)^\top = A$. In this special case, normals transform by A. This makes sense, because an orthogonal transformation such as a rotation preserves angles; if $\vec{n}$ is perpendicular to a polygon, then after both are transformed orthogonally the results will still be perpendicular.

4. Using only the definitions of trace and matrix multiplication, prove that for any two matrices A and B,

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

Answer: For any $n \times n$ matrices A and B,

$$\begin{aligned}\operatorname{tr}(AB) &= \sum_{i=1}^{n} (AB)_{ii} \\ &= \sum_{i=1}^{n} \sum_{k=1}^{n} A_{ik} B_{ki} \\ &= \sum_{k=1}^{n} \sum_{i=1}^{n} B_{ki} A_{ik} \\ &= \sum_{k=1}^{n} (BA)_{kk} \\ &= \operatorname{tr}(BA).\end{aligned}$$

5. Using Problem 4, prove that for any matrix C and any invertible matrix S,

$$\operatorname{tr}(SCS^{-1}) = \operatorname{tr} C.$$

Answer: Let $A = S$ and $B = CS^{-1}$. Then, using Problem 4,

$$\text{tr}(SCS^{-1}) = \text{tr}(AB) = \text{tr}(BA) = \text{tr}(CS^{-1}S) = \text{tr}\, C.$$

6. Prove that $\bar{\mathcal{B}}$ is really a basis for $\bar{V}$.

Answer: First we show that $\bar{\mathcal{B}} = \{f_1, \ldots, f_n\}$ is linearly independent. Suppose that $c_1 f_1 + c_2 f_2 + \cdots + c_n f_n = 0$. Apply this function to v_1:

$$\begin{aligned} 0 &= (c_1 f_1 + c_2 f_2 + \cdots + c_n f_n)(v_1) \\ &= c_1 f_1(v_1) + c_2 f_2(v_1) + \cdots c_n f_n(v_1) \\ &= c_1 \cdot 1 + c_2 \cdot 0 + \cdots + c_n \cdot 0 \\ &= c_1. \end{aligned}$$

So $c_1 = 0$. Similarly, applying the function to any v_j shows that $c_j = 0$. Thus $c_1 = \cdots = c_n = 0$. This shows that $\bar{B}$ is linearly independent. To show that it spans $\bar{V}$, let $f : V \to \mathbb{R}$ be an arbitrary linear transformation. Let $c_1 = f(v_1), \ldots, c_n = f(v_n)$. I claim that $f = c_1 f_1 + \cdots + c_n f_n$. To see this, let v_j be any element of $\mathcal{B}$. Then

$$\begin{aligned} (c_1 f_1 + \cdots + c_n f_n)(v_j) &= c_1 f_1(v_j) + \cdots + c_n f_n(v_j) \\ &= c_1 \cdot 0 + \cdots + c_{j-1} \cdot 0 + c_j \cdot 1 + c_{j+1} \cdot 0 + \cdots + c_n \cdot 0 \\ &= f(v_j). \end{aligned}$$

So the functions $c_1 f_1 + \cdots + c_n f_n$ and f agree on every element of $\mathcal{B}$, and hence on all of V.

7. What is the relationship between $[T]_{\mathcal{B}}$ and $[\bar{T}]_{\bar{\mathcal{B}}}$?

Answer: In order to simplify the notation, let $A = [T]_{\mathcal{B}}$ and $B = [\bar{T}]_{\bar{\mathcal{B}}}$. These mean that

$$\begin{aligned} T(v_j) &= \sum_{k=1}^{n} A_{kj} v_k, \\ \bar{T}(f_i) &= \sum_{k=1}^{n} B_{ki} f_k. \end{aligned}$$

We now compute $(\bar{T}(f_i))(v_j) = f_i(T(v_j))$ in two different ways. On the one hand,

$$\begin{aligned} (\bar{T}(f_i))(v_j) &= \left(\sum_{k=1}^{n} B_{ki} f_k \right)(v_j) \\ &= \sum_{k=1}^{n} B_{ki} f_k(v_j) \\ &= B_{ji} \end{aligned}$$

(because only the $k = j$ term survives). On the other hand, using the fact that f_i is linear we have

$$\begin{aligned} f_i(T(v_j)) &= f_i\left(\sum_{k=1}^{n} A_{kj} v_k\right) \\ &= \sum_{k=1}^{n} A_{kj} f_i(v_k) \\ &= A_{ij} \end{aligned}$$

(because only the $k = i$ term survives). Thus $B_{ji} = A_{ij}$. We conclude that $[\bar{T}]_{\bar{\mathcal{B}}} = [T]_{\mathcal{B}}^{\top}$.

[Remark: Some students were skeptical of this problem; they seemed to suspect that I made it up just to irritate them! I did not; the *dual space* (I substituted the term "mirror space" to throw off potential cheaters) is a foundational concept used throughout linear algebra and its applications. For example, you can't do general relativity without it.]

[Remark: Remember that matrices are used to represent linear transformations (among other things). Multiplying matrices corresponds to composing transformations; that's why matrix multiplication exists. Adding matrices corresponds to adding transformations. So to what does transposing matrices correspond? Now you know: dualizing transformations.]