**Problem 1.** Show that if $A$ is diagonalizable, then $\det(A)$ is equal to the product of its eigenvalues. Show that $\det(A) = \lambda_1 \lambda_2 \ldots \lambda_n$ if the characteristic polynomial is factored as $\chi(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \ldots (\lambda_n - \lambda)$.

If $A$ is diagonalizable, there is an invertible matrix $P$ and a diagonal matrix $D$ so that $A = PD P^{-1}$ where the diagonal entries $d_1, \ldots, d_n$ of $D$ are the eigenvalues of $A$. Since $\det(D) = d_1 d_2 \ldots d_n$ and $\det(A) = \det(P) \det(D) \det(P^{-1}) = \det(D)$, the determinant of $A$ is equal to the product of the eigenvalues of $A$.

In general, the characteristic polynomial is defined as $\chi(\lambda) = \det(A - \lambda I)$. Hence,

$$\det(A) = \det(A - 0I) = \chi(0) = (\lambda_1 - 0) \ldots (\lambda_n - 0) = \lambda_1 \lambda_2 \ldots \lambda_n.$$

**Problem 2.** If $A$ is positive definite, show that $A^{-1}$ is also positive definite.

$A$ is positive definite iff all of the eigenvalues of $A$ are positive. Thus, there exists a diagonal matrix $\Lambda$ with diagonal entries $\lambda_1, \ldots, \lambda_n > 0$ and an orthogonal matrix $V$ such that $A = V \Lambda V^T$. Then $A^{-1} = V \Lambda^{-1} V^T$, so the eigenvalues of $A^{-1}$ are $\lambda_1^{-1}, \ldots, \lambda_n^{-1}$ which are positive. Hence, $A^{-1}$ is positive definite.

**Problem 3.** Consider a line $L$ through $x_0$ minimizing $\sum_{i=1}^n |\text{distance}(x_i, L)|^2$. Show that the line goes through $\bar{x} = \left(\sum_{i=1}^n x_i\right)/n$ and the slope of the closest line is the first principal component.

Let $L^* = \{x_0^* + tw^*\}$ be the closest line. Let $y_0^* = x_0^* - (w^* \cdot x_0^*)w^*$ be the orthogonal projection of $x_0^*$ to span$\{w^*\}^\perp$, and let $y_i = x_i - (w^* \cdot x_i)w^*$ for $1 \leq i \leq n$ be the orthogonal projections of $x_1, \ldots, x_n$ to span$\{w^*\}^\perp$. We have

$$\text{distance}(x_i, L^*) = \|y_i - y_0^*\|,$$

so

$$\sum_{i=1}^n |\text{distance}(x_i, L^*)|^2 = \sum_{i=1}^n \|y_i - y_0^*\|^2 = n \left\|y_0^* - \frac{\sum_{i=1}^n y_i}{n}\right\|^2 + \sum_{i=1}^n \|y_i\|^2 - \frac{1}{n} \sum_{i=1}^n \|y_i\|^2.$$

Note that $y_i$ only depends on $x_i$ and $w^*$. For fixed $w^*$, the above expression is minimized when $y_0^* = \frac{\sum_{i=1}^n y_i}{n} = \bar{x} - (w \cdot \bar{x})w$, then $L^* = \{y_0^* + tw^*\} = \{\bar{x} + tw^*\}$ goes through $\bar{x}$. Thus, $L^*$ goes through $\bar{x}$.

We now optimize over $w^*$ knowing that $L^*$ goes through $\bar{x}$. The distance from $x_i$ to $L^*$ can be written as

$$\text{distance}(x_i, L^*)^2 = \|x_i - \bar{x}\|^2 - |w^* \cdot (x_i - \bar{x})|^2,$$
so
\[ \sum_{i=1}^{n} |\text{distance}(x_i, \mathcal{L}^*)|^2 = \sum_{i=1}^{n} \|x_i - \bar{x}\|^2 - \sum_{i=1}^{n} |w^* \cdot (x_i - \bar{x})|^2. \]

Since \( \sum_{i=1}^{n} \|x_i - \bar{x}\|^2 \) does not depend on \( w^* \), the above expression is minimized when \( \sum_{i=1}^{n} \|w^* \cdot (x_i - \bar{x})\|^2 \) is maximized. This is exactly the variance of the orthogonal projections of \( x_i \) on \( \text{span}\{w^*\} \), which is maximized by the first principal component. Thus the slope of the line \( w^* \) is the first principal component.

**Problem 4.** (a) Suppose \( A \) is positive definite. Is \( \begin{bmatrix} A & I \\ I & A^{-1} \end{bmatrix} \succeq 0 \)?

Recall that if \( A \succ 0 \) and the Schur complement \( C = D - B^T A^{-1} B \succeq 0 \), then \( \begin{bmatrix} A & B \\ B^T & D \end{bmatrix} \succeq 0 \).

Apply this with \( A = A, B = I, D = A^{-1} \), since \( C = A^{-1} - I^T A^{-1} I = 0 \succeq 0 \), we have \( \begin{bmatrix} A & I \\ I & A^{-1} \end{bmatrix} \succeq 0 \).

(b) Computational question.

The solution can be found in the separate file.