PROJECT 4

Math 441

Due: Thursday, October 2

All "by hand" calculations require that you show your work. All MATLAB code and output must be turned in.

1. **Positive Definite Matrices have positive eigenvalues:** Let $A \in \Re^{n\times n}$ be a symmetric matrix. Prove that A has positive eigenvalues if and only if $x^T Ax > 0$ for every $x \in \Re^n$.

    (a) To get the "right direction" where you assume that A has positive eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$, write $A = W\Lambda W^T$ where the columns of W are the orthonormal eigenvectors of A and Λ is a diagonal matrix with the corresponding eigenvalues on the diagonal. Now consider $x^T Ax$ and let $y = W^T x$ ...

    (b) To get the "left direction," where you assume that $x^T Ax > 0$ for every $x \in \Re^n$, consider $x = w_i$ where w_i is the eigenvector corresponding to λ_i (make sure you know what this correspondence means).

2. Do exercise 1.4.15 by hand.

3. Do exercise 1.4.21 using MATLAB for the calculations in (a) and (b). Recall that **R=chol(A)** gives the Cholesky factorization. Now answer the question: use the result of your calculations to say whether A is positive definite.

4. Do exercise 1.4.35. *Hint*: Write

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \text{and} \quad R = \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix}$$

    where R_{11} and R_{22} are upper triangular.

5. Do exercise 1.5.3.

6. Do exercise 1.5.6.

7. Do exercise 1.6.5. Use **p=symrcm(A); A=A(p,p);** to do part (b). Use **p=symamd(A); A=A(p,p);** to do part (c).

8. **Principle Component Analysis:** Five individuals took part in a study which studied the effect of a new drug on weekly alcohol consumption over three weeks. The data matrix for these five individuals is:

$$D = \begin{bmatrix} 20 & 10 & 8 \\ 12 & 6 & 6 \\ 7 & 4 & 0 \\ 10 & 5 & 5 \\ 16 & 12 & 13 \end{bmatrix}$$

    where each row contains the three weekly measurements (number of drinks per week) for a different individual. You will calculate a single number, or statistic, which is the "best" in terms capturing the most variability in the data for each individual:

(a) Enter the matrix into MATLAB by entering: **D=[20 10 8;12 6 6;7 4 0;10 5 5;16 12 13]**

(b) Plot the "response curves for each individual. In MATLAB, enter **plot(D')**. Notice that MATLAB colors each response differently. Note that each reponse represents a point in $\Re^3$.

(c) Based on this plot, is the drug effective in reducing alcohol consumption? Why or why not?

(d) Calculate the covariance matrix for this data set. In MATLAB, **C=cov(D)**.

(e) Calculate the eigenvectors and eigenvalues of the covariance matrix C via **[W Lambda]=eig(C)**.

(f) If $0 < \lambda_1 < \lambda_2 < \lambda_3$ are the three eigenvalues of C in increasing order, and w_1, w_2 and w_3 are the corresponding eigenvectors (columns of W), calculate the proportion of the variance of the data described by the eigenvector w_3 by

$$\frac{\lambda_3}{\sum_{i=1}^{3} \lambda_i}.$$

(g) Project the data onto the eigenvector w_3 by entering **ProjW3D = W(:,3)'*D'**. Note that this operation is performing a weighted average of the three data points for each individual. Which two data measurements (at week 1, 2 or 3) are most important (i.e. which ones are weighed the most)?

(h) Project the data onto the eigenvector w_2. What type of operation is this projection doing (averaging or differencing) and which measurements are most important for this calculation?

(i) What is the proportion of variance described by w_2?

(j) What single number do you recommend to use which summarizes the data for each individual. Why is this number the "best"?

EXTRA PROBLEMS:

These last 4 problems are optional. For each problem you do, you will receive extra credit. Extra problems 1-3 give the supporting theory which we relied upon to prove our theorem in class: A is symmetric positive definite if and only if $A = R^T R$ for some upper triangular R. Extra problem 4 is an inductive proof about the sum of the series $\sum_{i=1}^{n} i^2$, which we have used repeatedly in class when counting flops.

1. Do exercise 1.4.54.
2. Do exercise 1.4.56.
3. Do exercise 1.4.71.
4. Do exercise 1.5.9.