Math524 Final Exam by Dr. Samuel Shen, FA2024

Due 11:59 PM Tuesday/December 17, 2024

This is an open-book exam. You can use Google or ChatGPT to search for any helpful information. You can use a calculator, R, computer, or experiment. However, you CANNOT ask any natural person to help you. When done, please submit **a single pdf file and a single R file** *online via Canvas.*

1. $[20\ points]$ **SVD for a space-time data matrix**
   Make an SVD analysis of the January standardized anomalies data of sea level pressure (SLP) at Darwin and Tahiti from 1961 to 2000. You can download the SLP data from this Final Exam on Canvas. The two data files are `PSTANDdarwin.txt` and `PSTANDtahiti.txt`.
   (a) $[4\ points]$ Write an R code to organize the January data from 1961 to 2000 into a 2×40 space-time data matrix A. Put Darwin data in the first row and Tahiti in the second row. Explicitly print the first six columns of the data matrix A. Copy and paste the six columns of the data into your R code as comments indicated by #.
   (b) $[4\ points]$ Make the SVD calculation for this space-time matrix $A = UDV^t$, where V^t is the transpose matrix of V.
   (c) $[4\ points]$ Plot the first singular vector in the above temporal matrix **V** against time 1961 to 2000 as a time series curve, which is called the first principal component, denoted by PC1.
   (d) $[4\ points]$ Check the historical El Niño events between 1961 and 2000 from the Internet like

   `https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php`

   and interpret the three extreme values of PC1 in January 1983, January 1992, and January 1997 from the perspective of El Niño events. Here, extreme values mean either the local maxima or local minima for the PC1 curve in Part (c).
   (e) $[4\ points]$ Interpret the first singular column vector in **U**, which is called the first empirical orthogonal function (EOF1), as weights of Darwin and Tahiti stations. *Hint: Read some website materials or books on El Niño and check if the two values of the EOF1 vector have the same sign or different signs. Here, different signs mean that one value is positive and another is negative; the same sign means that both values are positive or negative.*

2. $[14\ points]$ **The singular values of SVD and the eigenvalues of a covariance matrix**
   Let A be a space-time anomaly data matrix. Let C be a square covariance matrix defined by

   $$C = AA^t, \tag{0.1}$$

   where A^t is the transpose matrix of A. Let d_1 be the first singular value of A, and λ_1 be the first eigenvalue of C. Show that

   $$\lambda_1 = d_1^2. \tag{0.2}$$

3. $[8\ points]$ **Linear equations**
   (a) $[6\ points]$ Use R to solve the following linear equations $Ax = b$ to find the vector x, where A and b are given as follows.

   $$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 0 & 9 \\ 3 & 1 & 2 & 9 \end{bmatrix}, \tag{0.3}$$

   $$b = \begin{bmatrix} 2 \\ 5 \\ 0 \\ 1 \end{bmatrix}, \tag{0.4}$$

(b) [2 $points$] Using text, create your own word problem for two linear equations with two variables. *For example, I have two brothers, John and Mike. The difference of John's age minus Mike's is 4, and the sum of their ages is 20. What are the ages of John and Mike?*

4. [12 $points$] **Dot product, cross product, angle, and R plotting**

(a) [2 points] Given the following two vectors

$$\mathbf{a} = (2, 1) \tag{0.5}$$
$$\mathbf{b} = (1, 2). \tag{0.6}$$

Use hand calculation to find the dot product of these two vectors.

(b) [2 points] Use R and hand calculation to calculate the angle θ between the two vectors, as shown in Fig. 1. Use degrees, not radians, for the angle.

(c) [2 points] Write an R code to reproduce Fig. 1.

(d) [2 points] Use hand calculation to find the cross product of the following two 3D vectors

$$\mathbf{u} = (2, 1, 0) \tag{0.7}$$
$$\mathbf{v} = (1, 2, 0). \tag{0.8}$$

That is, find the vector $\mathbf{w} = \mathbf{u} \times \mathbf{v}$. Use R to verify your result.

(e) [2 points] Use your hand to draw the three vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.

(f) [2 points] Use text, the right-hand rule, and a real-life example to explain the meaning of the cross product of two vectors. *[Limited to 30 - 100 words.]*

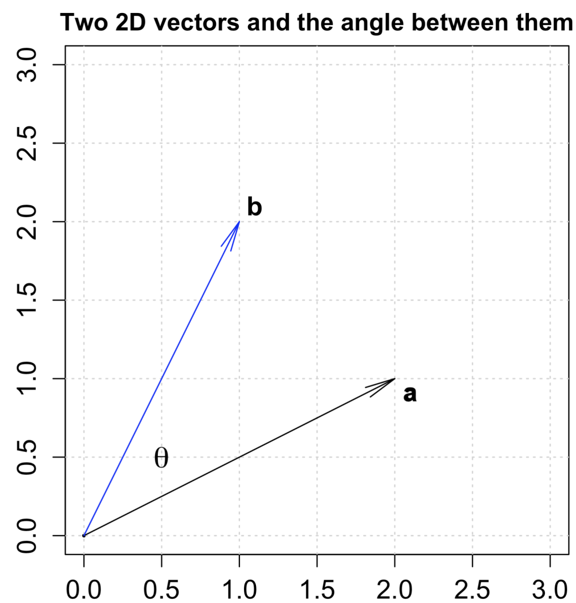

Figure 1 Figure for Problem #4: Two 2D vectors $\mathbf{a}$ and $\mathbf{b}$ and the angle θ between them.

5. [16 $points$] **Concept problems**

(a) [3 points] Use your own words and hand-draw a diagram to explain what it means for two 2D vectors that are linearly dependent. Limited to 20-50 words.

(b) [3 points] If $\mathbf{u}$ is an eigenvector of a matrix A, and if $\mathbf{v}$ is also an eigenvector of the matrix A, then is the vector $\mathbf{u} + \mathbf{v}$ an eigenvector of the matrix A? Use a mathematical proof to justify your answer.

(c) [10 points] Given that

$$A = \begin{bmatrix} 5 & 1 \\ 0 & 3 \end{bmatrix} \tag{0.9}$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{0.10}$$

(i) [2 points] Show by hand, not by R, that vector $(1, 1)$ is not an eigenvector of this matrix A.
(ii) [3 points] According to the following matrix expression

$$P(x_1, x_2) = \mathbf{x}^t A^t A \mathbf{x}, \tag{0.11}$$

write down by hand the second-order polynomial of x_1 and x_2 in the following form

$$P(x_1, x_2) = a x_1^2 + b x_1 x_2 + c x_2^2. \tag{0.12}$$

You need to find a, b and c.
(iii) [2 points] Use R to compute all the eigenvectors and eigenvalues of $A^t A$.
(iv) [3 points] Let (x_1, x_2) be equal to the first unit eigenvector of $A^t A$ computed in (ii). Use R to compute the numerical value of $P(x_1, x_2)$.

6. [10 $points$] **From a photo to a data matrix**
(a) [4 points] Use R package `imager` to read Sam's photo file `sam.png` and produce a grayscale data matrix `graydat`. Print the first four rows and first three columns of the data matrix. Put this part of the print out on your code as comments marked by #. The figure can be downloaded from the exam site on Canvas.
(b) [2 points] Use R to find the maximum and minimum values of `graydat`.
(c) [4 points] Plot a grayscale photo of Sam. According to the graydat matrix' rows and columns, use R command `points()` to mark the location of the 100th row and 200th column with a blue solid round point on the grayscale photo.

7. [20 $points$] **Machine learning**
(a) [7 points] K-means: Based on the K-means principle of minimal tWCSS described in Section 3.1 of our textbook, use both hand-calculation and R to determine the two clusters from the following three points:

$$P_1(1, 1), \quad P_2(1, 0), \quad P_3(3, 4). \tag{0.13}$$

(b) [3 points] Use R to plot the three points and the two centers on a figure similar to Fig. 3.1 in the textbook.
(c) [7 points] Support vector machine: Use R to plot the two hyperplanes $\mathbf{w} \cdot \mathbf{x} - b = \pm 1$ and the separating plane $\mathbf{w} \cdot \mathbf{x} - b = 0$ for the maximum separation between the two clusters in (a). Use Fig. 3.6 in the textbook as a reference.
(d) [3 points] What are the supporting vectors in (c)? Calculate the distance D_m between the positive hyperplane and the negative hyperplane according to the formula $D_m = 2/|\mathbf{w}|$?