QBUS1040: Foundations of Business Analytics
Homework 3

Semester 2, 2019

This homework consists of eight problems that require you to submit a written response and a coding component. The problems that require a written response are described in this paper. You need to print it and write your answers directly in this paper. You should use scratch paper (which you will not turn in) to do your rough work. You should submit a scanned copy of your written solution as a PDF via Canvas.

Please do not submit *photos* of your written solution as it is difficult for the marking system to recognise your submission. Also, please do not use a tablet to write your homework as it is very likely that your submission will not be processed correctly through the marking system. Please use a conventional scanner. You should double check your PDF before you submit the file. Your PDF submission must be a 5-page document. The file size of your PDF document should not exceed 128MB, or it will not be accepted by the submission system.

The coding components are described in a separate Jupyter Notebook file. You should also download the Jupyter Notebook file for Homework 3 and enter your code in the space provided. You should submit your code as a Jupyter notebook file via Canvas.

The homework is due by **5pm on Friday, the 27th of September. Late homework will not be accepted. Violation of the above submission instructions may incur a 30% penalty.**

For all problems where you are asked for a free-form answer, it must be written in the box below the problem. We won't read anything outside the boxes.

All problems have equal weight. Some are easy. Others, not so much.

Tutorial time: Friday 4 - 6pm

Tutor's name: Mr Kam Fung (Henry) Cheung

Your SID: 4 8 0 0 4 8 6 9 1 1

(For QBUS1040 staff only)

Question:	1	2	3	4	5	6	7	8	Total
Points:	10	10	10	10	10	10	10	10	80
Score:									

1. (10 points) *Centroid interpretations.* The n-vectors $x_1, \dots, x_N$ contain n attributes of N patients admitted to a hospital. The first component, $(x_i)_1$, is the age of patient i. The second component, $(x_i)_2$, is 1 if the patient is having trouble breathing, and 0 if not. The third component,$(x_i)_3$, is the body mass index of the patient. (The other components give other attributes.) A QBUS1040 graduate carries out k-means on this data set, with $k = 25$. She finds the 18th centroid or group representative is $z_{18} = (41.7, 0.36, 19.8, \dots, 29.6)$. Give a simple short interpretation in English of the first three components.

The average age of the patients in group 18 is 41.7.

The proportion of patients having trouble breathing in group 18 is 0.36.

The average body mass index of the patients in group 18 is 19.8.

2. (10 points) *Orthogonalizing vectors.* Suppose that a and b are any n-vectors. Show that we can always find a scalar γ so that $(a - \gamma b) \perp b$, and that γ is unique if $b \neq 0$. (Give a formula for the scalar γ.) In other words, we can always subtract a multiple of a vector from another one, so that the result is orthogonal to the original vector. The orthogonalization step in the Gram-Schmidt algorithm is an application of this. You should think about what happens when $b = 0$ and when $b \neq 0$.

Suppose γ is a scalar such that $(a - \gamma b) \perp b$

Since $(a - \gamma b) \perp b \Rightarrow (a - \gamma b)^T b = 0$

$$a^T b - \gamma b^T b = 0$$
$$a^T b = \gamma b^T b$$
$$a^T b = \gamma \|b\|^2$$
$$\gamma = \frac{a^T b}{\|b\|^2}, \quad b \neq 0$$

Suppose η is also a scalar such that $(a - \eta b) \perp b$

$\Rightarrow (a - \eta b)^T b = 0 \Leftrightarrow a^T b - \eta b^T b = 0 \Leftrightarrow a^T b = \eta b^T b$

$\Leftrightarrow a^T b = \eta \|b\|^2 \Leftrightarrow \eta = \dfrac{a^T b}{\|b\|^2}, \ b \neq 0$

Thus, $\eta = \gamma$. Therefore, γ is unique if $b \neq 0$.

3. (10 points) Let α, β, and η be scalars and let a, b, and c be pairwise orthogonal n-vectors. (This means that $a \perp b$, $a \perp c$, and $b \perp c$.) Express $\|\alpha a + \beta b + \eta c\|$ in terms of $\|a\|$, $\|b\|$, $\|c\|$, α, β, and η.

We have: $\|\alpha a + \beta b + \eta c\|^2$

$= (\alpha a + \beta b + \eta c)^T (\alpha a + \beta b + \eta c)$

$= \alpha a^T \alpha a + \alpha a^T \beta b + \alpha a^T \eta c + \beta b^T \alpha a + \beta b^T \beta b + \beta b^T \eta c + \eta c^T \alpha a + \eta c^T \beta b + \eta c^T \eta c$

$= \alpha^2 (a^T a) + (\alpha\beta) a^T b + (\alpha\eta) a^T c + \beta\alpha (b^T a) + \beta^2 (b^T b) + \beta\eta (b^T c) + \eta\alpha (c^T a) + \eta\beta (c^T b) + \eta^2 (c^T c)$

Since a, b, and c are pairwise orthogonal vectors

$\Rightarrow a^T b = a^T c = b^T c = 0$

So $\|\alpha a + \beta b + \eta c\|^2 = \alpha^2 (a^T a) + \beta^2 (b^T b) + \eta^2 (c^T c)$

Therefore, $\|\alpha a + \beta b + \eta c\| = \sqrt{\alpha^2 \|a\| + \beta^2 \|b\| + \eta^2 \|c\|}$

4. (10 points) *Running Gram-Schmidt algorithm twice.*

You should provide your answer to parts (a)-(c) in the Homework 3 Jupyter Notebook file and put your answer to part (d) in the space below.

(a) Run the Gram-Schmidt algorithm on the given set of vectors described in the Jupyter Notebook file and return the vectors $q_1, \dots, q_{15}$.

(b) Run the Gram-Schmidt algorithm on vectors $q_1, \dots, q_{15}$ obtained from the previous part and return the vectors $z_1, \dots, z_{15}$.

(c) Compute the distances between vectors q_1 and z_1, q_2 and z_2, ..., q_{15} and z_{15}.

(d) Explain the magnitude of the distances obtained in the previous part. You may want to start by saying what vectors $q_1, \dots, q_{15}$ and $z_1, \dots, z_{15}$ are.

Given that we ran the GS algorithm successfully the first time from part a), we can conclude that the vectors $q_1, \dots, q_{15}$ are orthonormal $\Rightarrow \begin{cases} q_i^T q_j = 0 & \text{if } i \neq j \\ q_i^T q_j = 1 & \text{if } i = 1 \end{cases}$

Onto running the algorithm the second time:

- $\tilde{z}_1 = q_1 \Rightarrow z_1 = \frac{q_1}{\|q_1\|} = q_1$
- $\tilde{z}_2 = q_2 - (z_1^T q_2) z_1 = q_2 - (q_1^T q_2) q_1 = q_2 \Rightarrow z_2 = \frac{q_2}{\|q_2\|} = q_2$

similarly to the 15th term

- $\tilde{z}_{15} = q_{15} - (z_1^T q_{15}) z_1 - \dots - (z_{14}^T q_{15}) z_{15}$
  $= q_{15} - (q_1^T q_{15}) q_1 - \dots - (q_{14}^T q_{15}) z_{15} = q_{15}$

Therefore, $z_i = q_i$, with $i = 1, \dots, 15$.

Thus, the magnitude of the distances obtained in part c are zeroes or very close to zero (due to floating-point errors during computation).

5. (10 points) *Matrix sizes.* Suppose A, B, and C are matrices that satisfy $A + BB^T = C$. Determine which of the following statements are necessarily true. (There may be more than one true statement.)

   (a) A is square.
   ☑ True ○ False

   (b) A and B have the same dimensions.
   ○ True ☑ False

   (c) A, B, and C have the same number of rows.
   ☑ True ○ False

   (d) B is a tall matrix.
   ○ True ☑ False

6. (10 points) *Matrix dimensions.* Suppose A is a 5×10 matrix, B is a 20×10 matrix, and C is a 10×10 matrix. Determine whether each of the following expressions make sense. If the expression makes sense, give its dimensions.

   (a) $A^TA + C$

   We have A^T is a 10×5 matrix $\Rightarrow$ A^TA is a 10×10 matrix
   Thus $A^TA + C$ makes sense, its dimension is 10×10

   (b) BC^3

   $C^2 = C.C$ is a 10×10 matrix $\Rightarrow$ $C^3 = C^2.C$ is also a 10×10 matrix
   Thus BC^3 makes sense, its dimension is 20×10.

   (c) $I + BC^T$

   C^T is a 10×10 matrix $\Rightarrow$ BC^T is a 20×10 matrix.
   Since I is ~~has to be~~ a square matrix $\Rightarrow$ $I + BC^T$ does not make sense

   (d) $B^T - [C \quad I]$

   $[C \quad I]$ is a 10×20 matrix, so is B^T.
   Thus $B^T - [C \quad I]$ makes sense, and its dimension is 10×20.

   (e) $B\begin{bmatrix} A \\ A \end{bmatrix}C$

   $\begin{bmatrix} A \\ A \end{bmatrix}$ is a 10×10 matrix $\Rightarrow$ $B\begin{bmatrix} A \\ A \end{bmatrix}$ is a 20×10 matrix
   $B\begin{bmatrix} A \\ A \end{bmatrix}C$ makes sense, its dimension is 20×10.

7. (10 points) *Portfolio sector exposures.* The n-vector h denotes a portfolio of investments in n assets, with h_i the dollar value invested in asset i. We consider a set of m industry sectors, such as pharmaceuticals or consumer electronics. Each asset is assigned to one of these sectors. (More complex models allow for an asset to be assigned to more than one sector.) The *exposure* of the portfolio to sector i is defined as the sum of investments in the assets in that sector. We denote the sector exposures using the m-vector s, where s_i is the portfolio exposure to sector i. (When $s_i = 0$, the portfolio is said to be *neutral* to sector i.) An investment advisor specifies a set of desired sector exposures, given as the m-vector s^{des}. Express the requirement $s = s^{des}$ as a set of linear equations of the form $Ah = b$. (You must describe the matrix A and the vector b.) *Remark.* A typical practical case involves $n = 1000$ assets and $m = 50$ sectors. An advisor might specify$s_i^{des} = 0$ if she does not have an opinion as how companies in that sector$i$ will do in the future; she might specify a positive value for s^{des} if she thinks the companies i in that sector will do well (i.e., generate positive returns) in the future, and a negative value if she thinks they will do poorly.

With m sectors and n assets, we can express the requirement $s = s^{des}$ of the form $Ah = b$, with b the s^{des} vector and A the $m \times n$ matrix with each row a_i be a n-vector that encodes whether each asset is in some specific sector.

$$\begin{cases} a_{ij} = 1 \text{ if the asset } j \text{ belongs to sector } i \\ a_{ij} = 0 \text{ if otherwire} \end{cases}$$

8. (10 points) *Orthogonality.* **Please refer to the Jupyter Notebook file for this problem.**