

# QBUS1040: Foundations of Business Analytics

## Homework 2

Semester 2, 2019

This homework consists of seven 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.

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.

The coding components are described in a separate Jupyter Notebook file. You should also download the Jupyter Notebook file for Homework 2 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 30th of August. **Late homework will not be accepted.**

You must provide your SID in the space below.

The problems have unequal weight. Some are easy. Others, not so much.

**Tutorial time:** \_\_\_\_\_

**Tutor's name:** \_\_\_\_\_

**Your SID:**

--	--	--	--	--	--	--	--	--	--

(For QBUS1040 staff only)

Question:	1	2	3	4	5	6	7	Total
Points:	10	10	10	12	10	8	10	70
Score:								

1. (10 points) *Linear combinations of cash flows.* We consider cash flow vectors over  $T$  time periods, with a positive entry meaning a payment received, and negative meaning a payment made. A (unit) single period loan, at time period  $t$ , is the  $T$ -vector  $l_t$  that corresponds to a payment received of \$1 in period  $t$  and a payment made of  $$(1 + r)$ in period  $t + 1$ , with all other payments zero. Here  $r > 0$  is the interest rate (over one period). Let  $c$  be a $1  $T - 1$  period loan, starting at period 1. This means that $1 is received in period 1,  $$(1 + r)^{T-1}$  is paid in period  $T$ , and all other payments (i.e.,  $c_2, \dots, c_{T-1}$ ) are zero. Express  $c$  as a linear combination of single period loans.$

2. (10 points) *Reverse triangle inequality.* Suppose  $a$  and  $b$  are vectors of the same size. The triangle inequality states that  $\|a + b\| \leq \|a\| + \|b\|$ . Show that we also have  $\|a + b\| \geq \|a\| - \|b\|$ . *Hints.* Draw a picture to get the idea. To show the inequality, apply the triangle inequality to  $(a + b) + (-b)$ .

3. (10 points) *Norm identities.* Verify that the following identities hold for any two vectors  $a$  and  $b$  of the same size.

(a)  $(a + b)^T(a - b) = \|a\|^2 - \|b\|^2$ .

- (b)  $\|a + b\|^2 + \|a - b\|^2 = 2(\|a\|^2 + \|b\|^2)$ . This is called the *parallelogram law*.

4. (12 points) *k-means with nonnegative, proportions, or Boolean vectors.* Suppose that the vectors  $x_1, \dots, x_N$  are clustered using *k-means*, with group representatives  $z_1, \dots, z_k$ .

- (a) Suppose the original vectors  $x_i$  are nonnegative, i.e., their entries are nonnegative. Explain why the representatives  $z_j$  are also nonnegative.

- (b) Suppose the original vectors  $x_i$  represent proportions, i.e., their entries are nonnegative and sum to one. (This is the case when  $x_i$  are word count histograms, for example.) Explain why the representatives  $z_j$  also represent proportions, i.e., their entries are nonnegative and sum to one.

- (c) Suppose the original vectors  $x_i$  are Boolean, i.e., their entries are either 0 or 1. Give an interpretation of  $(z_j)_i$ , the  $i$ th entry of the  $j$  group representative.

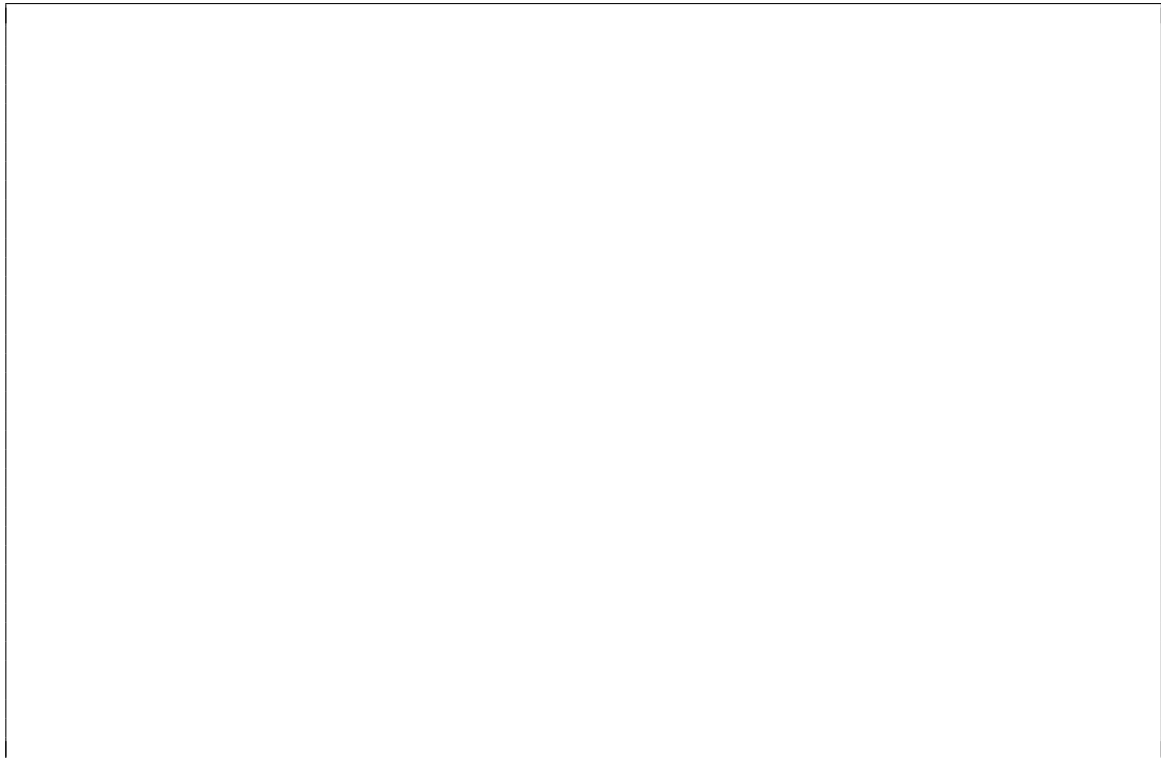
5. (10 points) *Weighted norm.* On page 51 of the textbook we discuss the importance of choosing the units or scaling for the individual entries of vectors, when they represent heterogeneous quantities. Another approach is to use a weighted norm of a vector  $x$ , defined as

$$\|x\|_w = \sqrt{w_1 x_1^2 + \cdots + w_n x_n^2},$$

where  $w_1, \dots, w_n$  are given positive weights, used to assign more or less importance to the different elements of the  $n$ -vector  $x$ . If all the weights are one, the weighted norm reduces to the usual ('unweighted') norm. It can be shown that the weighted norm is a general norm, i.e., it satisfies the four norm properties listed on page 46. Following the discussion on page 51, one common rule of thumb is to choose the weight  $w_i$  as the inverse of the typical value of  $x_i^2$  in the application. A version of the Cauchy-Schwarz inequality holds for weighted norms: For any  $n$ -vector  $x$  and  $y$ , we have

$$|w_1 x_1 y_1 + \cdots + w_n x_n y_n| \leq \|x\|_w \|y\|_w.$$

(The expression inside the absolute value on the left-hand side is sometimes called the weighted inner product of  $x$  and  $y$ .) Show that this inequality holds. *Hint.* Consider the vectors  $\tilde{x} = (x_1 \sqrt{w_1}, \dots, x_n \sqrt{w_n})$  and  $\tilde{y} = (y_1 \sqrt{w_1}, \dots, y_n \sqrt{w_n})$ , and use the (standard) Cauchy-Schwarz inequality.



6. (8 points) **Please refer to the Jupyter Notebook file for this problem.**
7. (10 points) **Please refer to the Jupyter Notebook file for this problem.**