UNIVERSITY OF CANTERBURY

End-of-Year Examination 2015

Prescription Number:	**EMTH211-15S2**
Paper Title:	**Engineering Linear Algebra and Statistics**

Time allowed: **3 HOURS**

- Answer ALL questions in the answer book provided.
- There is a *total* of 100 marks.
- Marks will be lost for poorly presented or incomplete answers.
- UC approved calculators are allowed.

PART ONE (24 marks)

1. Consider the matrix

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix}$$

   (a) Compute the LU decomposition of A.

   (b) Use the LU decomposition from above to solve

$$A\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

2. The set of all 2×2 matrices $M_{2\times 2}$ with real entries is vector space with the usual matrix addition and matrix scalar multiplication. Are the following subsets W and U subspaces? (Give reasons)

   - $$W = \left\{ A \in M_{2\times 2} \;:\; A = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \right\}$$

   - $$U = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2} \;:\; a + c = 1 \right\}$$

3. Consider the following matrix

$$A = \begin{bmatrix} 3 & 1 & 1 & 4 \\ 3 & 2 & 0 & 1 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 5 & 3 \end{bmatrix}$$

   (a) Determine the row rank (that is the dimension of the row space) of the matrix A.

   (b) Find a basis for the null-space

$$\text{null}(A) = \{\mathbf{x} \in V : A\mathbf{x} = \mathbf{0}\}$$

   of A.

   (c) How are the row rank and the nullity (the dimension of the null space) of an $n \times n$ matrix related in general?

PART TWO (32 marks)

4. (a) What is an eigenvector? What is an eigenvalue?

(b) Let A be a 2×2 matrix with eigenvectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, corresponding respectively to eigenvalues $\lambda_1 = 1, \lambda_2 = 4$. Find $A^4\mathbf{x}$, where $\mathbf{x} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$.

(c) A 7×7 matrix has three eigenvalues. One eigenspace is 2-dimensional, and another is 3- dimensional. Is it possible that the matrix is not diagonalizable? Explain carefully.

5. (a) i. Draw the Gerschgorin row-related disks to estimate the location of the eigenvalues of A in the complex plane, where

$$A = \begin{bmatrix} 7 & 0 & 1 \\ 2 & -6 & -1 \\ -1 & 2 & 4 \end{bmatrix}.$$

ii. Now on a separate diagram draw the Gerschgorin column-related disks.

iii. Given the placement of the disks and the comparison between the two diagrams, what can you conclude about the eigenvalues?

(b) i. Use the shifted power method to compute the second eigenvalue of $C = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$, if the first eigenvalue is $\lambda_1 = 2$.

(Hint: use $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ as your intial guess and two iterations will suffice).

ii. Alternatively, set up the equations and explain (but do not solve!) how you would find the second eigenvalue of matrix C by using the inverse power method (but with no shifts) instead.

6. Consider the two matrices:

$$A = \begin{bmatrix} 5 & -9 \\ 1 & -3 \\ -3 & 7 \\ 1 & -3 \end{bmatrix} \quad Q = \begin{bmatrix} 5/6 & 1/2 \\ 1/6 & -1/2 \\ -3/6 & 1/2 \\ 1/6 & -1/2 \end{bmatrix}$$

where the columns of Q have been obtained by applying the Gram-Schmidt process to the columns of A and then normalizing.

(a) What is the purpose of the Gram-Schmidt process? Give a brief **geometric** description of how it achieves this purpose.

(b) Find the QR factorisation of the matrix A.

(c) Use the QR factorization to solve the system of equations

$$A\mathbf{x} = \begin{bmatrix} 11 \\ 1 \\ -5 \\ 1 \end{bmatrix}$$

and check your answer.

(d) What would you expect to happen if you used the QR factorisation to solve the system of equations:

$$A\mathbf{x} = \begin{bmatrix} 11 \\ 1 \\ -5 \\ 2 \end{bmatrix}?$$

(You do not need to solve this system!)

7. (a) Find a symmetric matrix with eigenvalues $\lambda_1 = \lambda_2 = 1, \lambda_3 = -2$, and eigenspaces:

$$E_1 = \text{span}\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right), \quad E_2 = \text{span}\left(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right)$$

(Hint: Recall the spectral decomposition of $A = \lambda_1 \mathbf{q}_1\mathbf{q}_1^T + \lambda_2 \mathbf{q}_2\mathbf{q}_2^T + \ldots + \lambda_n \mathbf{q}_n\mathbf{q}_n^T$)

(b) Find a singular value decomposition for $B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

PART THREE (44 marks)

8. (a) Let $\mathbf{x}$ be a vector of n measurements of a variable x with mean $\bar{x}$. Furthermore, let $\tilde{\mathbf{x}}$ denote the centered measurements of x and $\mathbf{1}$ be a column vector of n ones.

    i. What does the process of centering do to the variable vector $\mathbf{x}$?

    ii. Show that $\tilde{\mathbf{x}}$ and $\mathbf{1}$ are orthogonal.

    iii. Let s_x be the standard deviation of x. Show that centering does not change the spread of the variable (*Hint: show that* $s_{\tilde{x}} = s_x$).

   (b) Suppose that we have 37 measurements of a variable x and 82 measurements of a variable y. Let $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ denote the centered measurements of x and y, respectively. If the length of $\tilde{\mathbf{x}}$ is 25 and the length of $\tilde{\mathbf{y}}$ is 36, which variable has greater spread? Justify your answer.

   (c) Suppose that we have the same number of measurements for each of three variables, x, y, and z. Show clearly whether it is possible for the correlations between the variables to simultaneously be $r_{xy} = 0.7, r_{xz} = 0.2$, and $r_{yz} = -0.1$.

   (d) Use the vector diagram in Figure 1 to answer this question. Suppose that we have the same number, n, of measurements for each of three variables, x, y, and z. Given that the standard deviation of y is 3, find

    i. n

    ii. s_x and s_z

    iii. r_{xy}, r_{xz}, and r_{yz}.

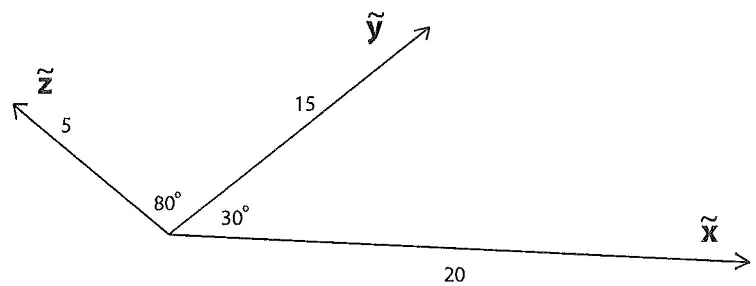
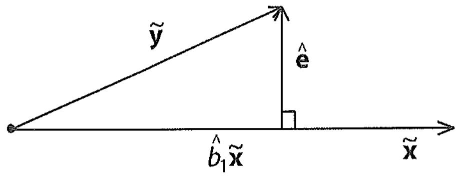


Figure 1: The angles between the vectors are as shown. The number on each vector is its length. The vectors all lie in the same the same plane.

9. Consider the simple linear regression problem involving predictor variable x, and response variable y, with a fitted slope $\hat{b}_1$. Let $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ denote the centered measurements of x and y, respectively.

   (a) Use the following vector diagram to show that $\hat{b}_1 = \tilde{\mathbf{x}} \cdot \tilde{\mathbf{y}} / \|\tilde{\mathbf{x}}\|^2$.



   (b) Show that $\hat{b}_1 = r_{xy}\|\tilde{\mathbf{y}}\|/\|\tilde{\mathbf{x}}\|$, where r_{xy} is the correlation between x and y.

   (c) List the three assumptions about the regression errors (residuals) that must be satisfied when using the F statistic to determine whether the fitted model is a good fit.

   (d) Suppose that $n = 32$ and $R = 0.8$, and recall that the goodness of fit F statistic is
$$F = \frac{(n-2)R^2}{1-R^2}.$$
   Assuming the three assumptions above are satisfied, use F to determine, at level $q = 0.05$, whether the fitted model is a good fit. A table of the F distribution is given at the end of the exam paper.

   (e) The level q mentioned in part (d) is the probability of making a type 1 error. What is a type 1 error in the context of this question?

   (f) What is a type 2 error in the context of this question?

   (g) If the correlation between x and y is $r_{xy} = 10$, what is the correlation between $-2x$ and y? Justify your answer.

10. Consider the problem of fitting a multiple linear regression (MLR) model with response variable y and $p = 3$ predictor variables, x_1, x_2, and x_3.

 (a) Pinpoint, in a single sentence, the root cause of instability in a MLR model.

 (b) Describe clearly, with the aid of vector diagrams, how the volume of the parallelepiped formed by the three predictor vectors can indicate whether the fitted model is stable.

 Now suppose that the number of measurements of each variable is $n = 34$. Let C be the correlation matrix for the predictor vectors and suppose that the eigenvalues of C are 0.1, 1, and 3.

 (c) Find the normalized volume of the parallelepiped formed by the three predictor vectors.

 (d) Recall that the Haitovsky statistic for stability of a MLR model is given by

$$H = \left(1 + \frac{2p+5}{6} - n\right) \ln(1 - \det C),$$

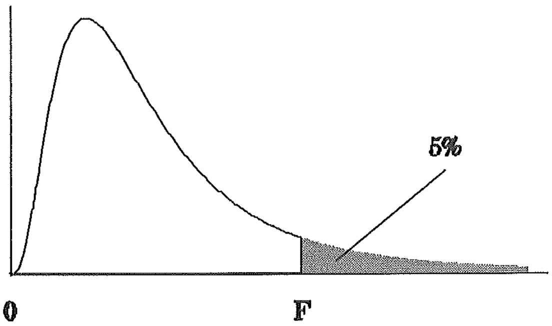
 where p is the number of predictor variables. Use H to determine, at level $q = 0.05$, whether the MLR model will be stable. You may assume that H has chi-squared distribution with parameter $p(p-1)/2$. A table of the chi-square distribution is given at the end of the exam paper.

 (e) Now suppose $F = 3$. Use F to determine, at level $q = 0.05$, whether the fitted model is a good fit. You may assume F has a $F(p, n-p-1)$ distribution. A table of the F distribution is given at the end of the exam paper.

 (f) Would increasing the level to $q = 0.10$ change your decision in part (e)? Explain.

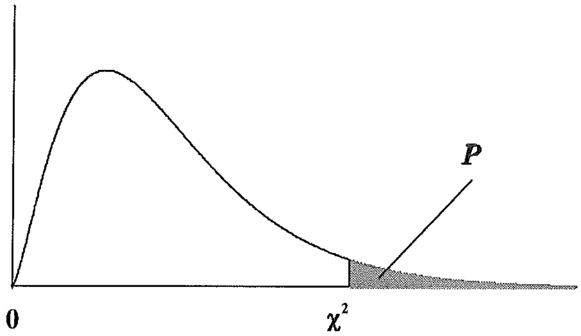
END OF PAPER

(b) 5% Points of the F distribution



ν_1	1	2	3	4	5	6	7	8	10	12	24
$\nu_2 = 2$	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.79	8.74	8.64
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	5.96	5.91	5.77
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.74	4.68	4.53
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.06	4.00	3.84
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.64	3.57	3.41
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.35	3.28	3.12
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.14	3.07	2.90
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	2.98	2.91	2.74
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.85	2.79	2.61
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.75	2.69	2.51
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.67	2.60	2.42
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.60	2.53	2.35
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.54	2.48	2.29
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.49	2.42	2.24
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.45	2.38	2.19
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.41	2.34	2.15
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.38	2.31	2.11
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.35	2.28	2.08
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.30	2.23	2.03
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.25	2.18	1.98
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.22	2.15	1.95
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.19	2.12	1.91
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.16	2.09	1.89
35	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.11	2.04	1.83
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.08	2.00	1.79
45	4.06	3.20	2.81	2.58	2.42	2.31	2.22	2.15	2.05	1.97	1.76
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.03	1.95	1.74
55	4.02	3.16	2.77	2.54	2.38	2.27	2.18	2.11	2.01	1.93	1.72
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	1.99	1.92	1.70

Percentage Points of the Chi-Square Distribution



This table gives percentage points of the chi-square distribution on ν d.f. These are the values of χ^2 for which a given percentage, P, of the chi-square distribution is greater than χ^2.

P	97.5	95	50	10	5	2.5	1	0.1
$\nu = 1$	.000982	0.00393	0.45	2.71	3.84	5.02	6.64	10.8
2	0.0506	0.103	1.39	4.61	5.99	7.38	9.21	13.8
3	0.216	0.352	2.37	6.25	7.81	9.35	11.3	16.3
4	0.484	0.711	3.36	7.78	9.49	11.1	13.3	18.5
5	0.831	1.15	4.35	9.24	11.1	12.8	15.1	20.5
6	1.24	1.64	5.35	10.6	12.6	14.5	16.8	22.5
7	1.69	2.17	6.35	12.0	14.1	16.0	18.5	24.3
8	2.18	2.73	7.34	13.4	15.5	17.5	20.1	26.1
9	2.70	3.33	8.34	14.7	16.9	19.0	21.7	27.9
10	3.25	3.94	9.34	16.0	18.3	20.5	23.2	29.6
11	3.82	4.57	10.3	17.3	19.7	21.9	24.7	31.3
12	4.40	5.23	11.3	18.5	21.0	23.3	26.2	32.9
13	5.01	5.89	12.3	19.8	22.4	24.7	27.7	34.5
14	5.63	6.57	13.3	21.1	23.7	26.1	29.1	36.1
15	6.26	7.26	14.3	22.3	25.0	27.5	30.6	37.7
16	6.91	7.96	15.3	23.5	26.3	28.8	32.0	39.3
17	7.56	8.67	16.3	24.8	27.6	30.2	33.4	40.8
18	8.23	9.39	17.3	26.0	28.9	31.5	34.8	42.3
19	8.91	10.1	18.3	27.2	30.1	32.9	36.2	43.8
20	9.59	10.9	19.3	28.4	31.4	34.2	37.6	45.3
22	11.0	12.3	21.3	30.8	33.9	36.8	40.3	48.3
24	12.4	13.9	23.3	33.2	36.4	39.4	43.0	51.2
26	13.8	15.4	25.3	35.6	38.9	41.9	45.6	54.1
28	15.3	16.9	27.3	37.9	41.3	44.5	48.3	56.9
30	16.8	18.5	29.3	40.3	43.8	47.0	50.9	59.7
35	20.6	22.5	34.3	46.1	49.8	53.2	57.3	66.6
40	24.4	26.5	39.3	51.8	55.8	59.3	63.7	73.4
45	28.4	30.6	44.3	57.5	61.7	65.4	70.0	80.1
50	32.4	34.8	49.3	63.2	67.5	71.4	76.2	86.7
55	36.4	39.0	54.3	68.8	73.3	77.4	82.3	93.2