Subject Outline

Subject Name	Essential Mathematics for Data Science
Subject Code	MA5801
Study Period	SP84
Study Mode	External
Campus	JCU Online
Subject Coordinator	Professor Ron White/Associate Professor Shaun Belward

We acknowledge the Traditional Owners of the lands and waters where our University is located and actively seek to contribute and support the JCU Reconciliation Statement, which exemplifies respect for Australian Aboriginal and Torres Strait cultures, heritage, knowledge and the valuing of justice and equity for all Australians.

Cairns
Singapore
Townsville

Pre-requisites

This subject outline has been prepared by Ron White and Shaun Belward for the College of Science and Engineering, Division of Tropical Environment and Societies, James Cook University. Updated 29/11/2017.

Q1. This subject is offered across more than one campus and/or mode and/or teaching period within the one calendar year.	Yes ☒	No ☐
Q2. If Yes (Q1), the design of all offerings of this subject ensure the same learning outcomes and assessment types and weightings.	Yes ☒	No ☐

Subject Outline Peer Reviewer

Name	Patrick Higgins
Position	Lecturer, Mathematics
Date Reviewed	30/11/17

Staff Contact Details

Teaching Team	Staff Member	Room	Phone	Email	Consultation times*
Subject Coordinator	Professor Ron White Assoc. Prof. Shaun Belward	N/A	0747814197 0747814297	Ronald.White@jcu.edu.au Shaun.Belward@jcu.edu.au	N/A
Tutor 1	Peter Stokes	N/A	N/A	peter.stokes@my.jcu.edu.au	Online
Tutor 2	Nathan Garland	N/A	N/A	Nathan.garland@jcu.edu.au	Online

*Other consultation times by appointment only.

Contents

1 Subject at a glance 4

1.1 Student participation requirements 4

1.2 Key dates 4

2 Subject details 4

2.1 Subject description 4

2.2 Subject learning outcomes 5

2.3 Learning and teaching in this subject 5

2.4 Student feedback on subject 5

2.5 Subject resources and special requirements 5

3 Assessment details 6

3.1 Requirements for successful completion of subject 6

3.2 Feedback on student learning 6

3.3 Assessment tasks 6

4 Submission and return of assessment 10

4.1 Submission of assessment 10

4.2 Return of assessment 10

5 Subject calendar 11

1 Subject at a glance

1.1 Student participation requirements

The JCU Learning, Teaching and Assessment Policy (4.3) indicates that, "a **3 credit point subject** will require a **130 hour work load** of study-related participation including class attendance over the duration of the study period, **irrespective of mode of delivery"**. This work load comprises **timetabled hours** and **other attendance requirements**, as well as **personal study hours,** including completion of online learning activities and assessment requirements. Note that "attendance at specified classes will be a mandatory requirement for satisfactory completion of some subjects" (Learning, Teaching and Assessment Policy, 5.10); and that additional hours may be required per week for those students in need of **English language, numeracy** or **other learning support.**

Key subject activities	Time	Day & Date	Room/Location
See subject calendar and assessment details below.			

1.2 Key dates

Key dates	Date
Census date	19/7/2018
Last date to withdraw without academic penalty	
Assessment task 1 [Quiz] [10 %]	Due: Sunday of Week 1
Assessment task 2 [MATLAB code workbook] [20 %]	Due: Sunday of Week 6
Assessment task 3 [Assignment - Microchip] [10 %]	Due: Sunday of Week 2
Assessment task 4 [Assignment - SVD] [20 %]	Due: Sunday of Week 5
Assessment task 5 [Assignment - VisualRank] [20 %]	Due: Sunday of Week 6
Assessment task 6 [Test – Discrete Maths] [20 %]	Due: Wednesday of Week 7

The dates and times are subject to change in special circumstances. Please always refer to LearnJCU for dates and times.

2 Subject details

2.1 Subject description

Data science is grounded in mathematics. This subject will provide you with the essential elements of mathematics required for data scientists.

It includes elements of discrete mathematics including logics, sets, proof, functions, relations, graphs and trees; as well as elements of linear algebra including linear systems and matrix formulation, vector spaces, eigenvalues/eigenvectors, singular value decomposition, optimisation and numerical methods. Computational aspects of this course will be developed in MATLAB.

2.2 Subject learning outcomes

Students who successfully complete this subject will be able to:

- Identify and apply concepts of set theory, arithmetic, logic, proof techniques, binary relations, graphs and trees to solve problems in data science
- Apply linear algebra and numerical mathematics concepts for optimisation and dimensionality reduction in data science problems
- Apply and implement concepts in discrete mathematics, optimisation and linear algebra in data science using MATLAB.

These outcomes will contribute to your overall achievement of **course learning outcomes.** Your course learning outcomes can be located in the entry for your course in the electronic JCU Course Handbook 2018 (see *Academic Requirements for Course Completion*).

2.3 Learning and teaching in this subject

Students are required to progress through the weekly material on Learn Ultra. Each week consists of reading materials including worked examples and exercises, some video demonstrations and different types of assessment. Please see below for further details of topics covered in each week and the assessment for the subject.

Some weeks will include sections labelled “For Interest”. These are interesting concepts or commentary that go beyond the essential knowledge content you need to know for a certain topic. While you will not be assessed on these, you might find them useful and insightful.

2.4 Student feedback on subject

As part of our commitment at JCU to improving the quality of our courses and teaching, we regularly seek feedback on your learning experiences. Student feedback informs evaluation of subject and teaching strengths and areas that may need refinement or change. ***Your JCU Subject and Teaching Surveys*** provide a formal and confidential method for you to provide feedback about your subjects and the staff members teaching within them. You will receive an email invitation when the survey opens. We value your feedback and ask that you provide constructive feedback about your learning experiences for each of your subjects, in accordance with responsibilities outlined in the Student Charter. Refrain from providing personal feedback on topics that do not affect your learning experiences. Malicious comments about staff are deemed unacceptable by the University.

2.5 Subject resources and special requirements

All subject resources will be available through Learn Ultra. Some links to extra online resources are made available in the weekly readings.

Students will require the MATLAB software.

3 Assessment details

3.1 Requirements for successful completion of subject

In order to pass this subject, you must:

- Achieve an overall percentage of 50% or more

Assessment items and final grades will be reviewed through moderation processes (Learning, Teaching and Assessment Policy, 5.13-5.18). It is important to be aware that assessment "is always subject to final ratification following the examination period and that no single result represents a final grade in a subject" (Learning, Teaching and Assessment Policy, 5.22.).

3.1.1 Inherent requirements

Inherent requirements are the fundamental abilities, attributes, skills and behaviours needed to achieve the learning outcomes of a course while preserving the academic integrity of the university's learning, assessment and accreditation processes. Students and prospective students must be able to demonstrate that they have acquired or have the ability to acquire the inherent requirements for their degree.

Reasonable adjustments may be made to assist students manage additional circumstances impacting on their studies provided these do not change the academic integrity of a degree. Reasonable adjustments do not alter the need to be able to demonstrate the inherent requirements of the course. Students who believe they will experience challenges completing their degree or course because of their disability, health condition or other reason should discuss their concerns with an AccessAbility Services team member or a member of College staff, such as the Course Coordinator. In the case where it is determined that inherent requirements cannot be met with reasonable adjustments, the University staff can provide guidance regarding other study options.

3.2 Feedback on student learning

Feedback will be provided to students on all assessment items. Some exercises that are built into the weekly readings have automatic feedback provided. The submitted assessment will be marked by tutors and returned to students with feedback.

3.3 Assessment tasks

ASSESSMENT TASK 1: QUIZ

Aligned subject learning outcomes	• Apply linear algebra and numerical mathematics concepts for optimisation and dimensionality reduction in data science problems
Group or individual	Individual
Weighting	10%
Due date	Sunday of Week 1

ASSESSMENT TASK 1: DESCRIPTION

Answer quiz questions on Linear Algebra topics covered in Week 1, including linear equations, matrices, Gaussian elimination, eigenvalues, eigenvectors and inverses.

40 marks, with one hour working time. The quiz consists of some questions for which students have to provide numeric responses and several true/false questions.

ASSESSMENT TASK 1: CRITERIA SHEET

For all online quiz questions, no working is required. Multiple choice or numeric answers will be marked immediately as correct or incorrect by comparing to pre-determined answers to the questions.

ASSESSMENT TASK 2: MATLAB CODE WORKBOOK

Aligned subject learning outcomes	• Apply and implement concepts in discrete mathematics, optimisation and linear algebra in data science using MATLAB.
Group or individual	Individual
Weighting	20%
Due date	2A: Sunday of Week 2 2B & 2C: Sunday of Week 3 2D: Sunday of Week 5 2E: Sunday of week 6

ASSESSMENT TASK 2: DESCRIPTION

Produce MATLAB code by either writing complete scripts (functions or programs) or adapting supplied scripts to complete the following tasks:

a) 2A - Creating functions - Write Matlab scripts to complete a program. Given a master file, write two function files.

b) 2B - Image Compression - Write Matlab scripts to complete a program, based on existing code. Given an image perform a SVD and use that reduce the quantity of data stored. We examine the ability of the SVD to retain the significant features in the image as the quantity of data retained is reduced.

c) 2C - Image Comparison - Use the SVD as a means to compare two images. The goal is to have the students run the code and to have them link the code with the outputs it produces.

d) 2D - MATLAB Logic - Use Matlab to evaluate a proposition in logic.

e) 2E - MATLAB Relations - Use Matlab to characterise relations.

ASSESSMENT TASK 2: CRITERIA SHEET

For all assessment questions, students are expected to show all necessary working. It is insufficient to write only the answer, even if it is correct. Also, while marks are not awarded for neatness, students may be penalised for poorly written or extremely untidy work.

For all Matlab coding questions, students are expected to comment each part of their code. Students will be penalised for poorly written or uncommented code. Main programs and all subroutines and functions should include a description of its purpose to be easily read by someone that doesn't necessarily know Matlab.

All assessment questions (written and coding) will be marked by comparing student answers to a model set of solutions and marking scheme. The model solutions will be written to a standard that is expected by a mathematics student that has gained enough knowledge about the lecture material and is able to demonstrate this knowledge.

All assessment will consist of a variety of questions covering topics from most areas of the subject. Assessment of your answers to these questions will depend on the type of question, but will fundamentally be on the basis of:

Correctness – is the answer provided correct, or the most correct answer that a mathematics student could be expected to provide?

Completeness – does the response provided answer the question completely, within the context of information provided (or reasonably expected to have been discovered) during the study of this subject?

Likelihood of arriving at the appropriate result – when a response is incomplete, is it likely that the 'train of thought' indicated in the partial response would lead toward an appropriate outcome (based on the knowledge and skills embodied by this subject)?

Optimisation – does the response provided represent the highest order of knowledge and skill that could be developed solely through the study of this subject?

ASSESSMENT TASK 3: ASSIGNMENT - MICROCHIP

Aligned subject learning outcomes	• Apply linear algebra and numerical mathematics concepts for optimisation and dimensionality reduction in data science problems • Apply and implement concepts in discrete mathematics, optimisation and linear algebra in data science using MATLAB.
Group or individual	Individual
Weighting	10%
Due date	Sunday of Week 2

ASSESSMENT TASK 3: DESCRIPTION

Track a power surge in a microchip which is modelled by a linear system of equations. This involves writing MATLAB code, generate images and interpret output.

ASSESSMENT TASK 3: CRITERIA SHEET

See CRITERIA SHEET for ASSESSMENT TASK 2

ASSESSMENT TASK 4: ASSIGNMENT - SVD

Aligned subject learning outcomes	• Apply linear algebra and numerical mathematics concepts for optimisation and dimensionality reduction in data science problems • Apply and implement concepts in discrete mathematics, optimisation and linear algebra in data science using MATLAB.
Group or individual	Individual
Weighting	20%
Due date	Sunday of Week 5

ASSESSMENT TASK 4: DESCRIPTION

Use the Singular Value Decomposition (SVD) to complete a principal component analysis (PCA) on a data set. An example of a fictitious data set is supplied within the work for week three. This is then used as a model for students to use in doing a PCA on some demographic and socio-economic data from countries in the tropics. The students are to describe the relationship between the variables in the data set.

Students must determine the relationships between the variables reported in the spreadsheet and the strength of those relationships. Students should also identify the outlier countries, those for which the relationships the PCA identifies are not present.

ASSESSMENT TASK 4: CRITERIA SHEET

See CRITERIA SHEET for ASSESSMENT TASK 2

ASSESSMENT TASK 5: ASSIGNMENT - VISUALRANK

Aligned subject learning outcomes	• Identify and apply concepts of set theory, arithmetic, logic, proof techniques, binary relations, graphs and trees to solve problems in data science • Apply and implement concepts in discrete mathematics, optimisation and linear algebra in data science using MATLAB.
Group or individual	Individual
Weighting	20%
Due date	Sunday of Week 6

ASSESSMENT TASK 5: DESCRIPTION

Apply Google's VisualRank algorithm to a dataset of images using the SVD-based image similarity measure introduced in Week 3.

ASSESSMENT TASK 5: CRITERIA SHEET

See CRITERIA SHEET for ASSESSMENT TASK 2

ASSESSMENT TASK 6: TEST

Aligned subject learning outcomes	• Identify and apply concepts of set theory, arithmetic, logic, proof techniques, binary relations, graphs and trees to solve problems in data science
Group or individual	Individual
Weighting	20%
Due date	Wednesday of Week 7

ASSESSMENT TASK 6: DESCRIPTION

Demonstrate knowledge of all covered aspects of discrete mathematics, including propositional and predicate logic, set theory, graph theory and combinatorics.

The test is 75 minutes in length and requires short answers to questions.

ASSESSMENT TASK 6: CRITERIA SHEET

See CRITERIA SHEET for ASSESSMENT TASK 2

4 Submission and return of assessment

4.1 Submission of assessment

All assessments are submitted through Learn Ultra.

Note that the Learning, Teaching and Assessment Policy (5.22.3) outlines a uniform formula of penalties that will be imposed for submission of an assessment task after the due date. **This formula is 5% of the total possible marks for the assessment item per day including part-days, weekends and public holidays**. After 20 days, the assessment item thus would be awarded 0 marks (i.e. 5% x 20 = 100% of total possible marks in penalties).

4.2 Return of assessment

Feedback on marked assessments will be available in the Gradebook in Learn Ultra.

Please see The Learning Centre website for other important student information pertaining to plagiarism and referencing, examinations advice and student support services.

5 Subject calendar

Please note, the sequence of some topics may change due to staff availability, resourcing, or due to unforeseen circumstances.

Week/Date/Module		Topics Covered
O	Orientation	
1	Matrix notation, matrix equations, and Gaussian elimination	Linear equations, Gaussian elimination, matrices, operations on matrices, existence of a solution, types of matrices, inverse of a matrix, applications of matrices, eigenvalues, the eigenvalue problem
2	Vectors, linear combinations and the Gramm-Schmidt procedure	Fundamentals of MATLAB programming, Structuring MATLAB programs, vector spaces, basis of a vector space, Orthogonality, Gram-Schmidt procedure, Vector Spaces of Matrices, Condition Numbers in MATLAB, Loading and Saving files in MATLAB
3	Singular Value Decomposition (SVD)	What is an SVD, Applications of SVDs - pseudo inverse and least squares fitting, Applications of SVDs - image compression, Applications of SVDs - image comparison
4	Graph theory	Graph theory fundamentals, Paths and circuits, Directed and weighted graphs
5	Logic	What is logic, Propositional logic, Connective definitions, Complex propositions, Truth tables, Laws of logic, Predicate logic, Quantification, Proofing methods
6	Set theory	Set basics, operations, and laws, Cardinality and Cartesian product, Relation notation and representation, Types of relations, Equivalence relations, Inclusion, exclusion and multiplication, Permutations and combinations
7	Study week to finish assessment	