THE UNIVERSITY OF AUCKLAND

FIRST SEMESTER, 2014 — MID-SEMESTER TEST
Campus: City

COMPUTER SCIENCE

Computational Science

(Time allowed: 50 minutes)

NOTE: Attempt *all* questions. There are 51 Marks in total.

Use of calculators is NOT permitted.

Put your answers in the answer boxes provided below each question.

SURNAME:

FIRSTNAME:

ID:

ID: ____________________

1. Consider the bisection method for finding the roots of $f(x) = x^2 + \frac{x}{2} - 7$. [5 marks total]

   (a) Check that $a = 2$ and $b = 6$ are valid starting points for the algorithm. [**3** marks]

   To be valid, need $sign(f(a)) \neq sign(f(b))$. Here, $f(a) = 4 + 1 - 7 = -2$ and $f(b) = 36 + 3 - 7 = 32$ so ok.

   (b) What are the values of a and b after one iteration of the bisection method? [**2** marks]

   Find the mid-point $c = \frac{a+b}{2} = \frac{2+6}{2} = 4$ and calculate $f(c) = 16 + 2 - 7 = 11$ which has the same sign as b, so $a = 2$ and $b = 4$.

2. Consider the system of linear equations $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A}$ is an $n \times n$ matrix. [7 marks total]

   (a) The system is considered easy to solve when $\mathbf{A}$ is diagonal, lower triangular or orthonormal. Define the terms diagonal, lower triangular and orthonormal in this context. [**3** marks]

   Diagonal means the only non-zero elements of $\mathbf{A}$ are the diagonal terms (that is $A_{ij} = 0$ for $i \neq j$). Lower triangular manes that $A_{ij} = 0$ if $i < j$. Orthonormal means that the columns of A are orthogonal and have magnitude 1 (so $A_i^T A_j = 0$ if $i \neq j$ or $= 1$ if $i = j$).

   (b) Suppose $\mathbf{A}$ is diagonal. What simple test determines whether a solution to $\mathbf{Ax} = \mathbf{b}$ exists? [**1** mark]

   A solution exists if $A_{ii} \neq 0$ for all i.

   (c) Suppose $\mathbf{A}$ is orthonormal. What is the solution to $\mathbf{Ax} = \mathbf{b}$? [**2** marks]

   $\mathbf{x} = \mathbf{A^T b}$.

ID: ______________________________

(d) Name a method that can coerce the equation $\mathbf{Ax} = \mathbf{b}$ into the form $\mathbf{Cx} = \mathbf{d}$ where $\mathbf{C}$ is triangular. [**1** mark]

Gaussian elimination or row reduction.

3. Let $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 2 \end{bmatrix}$. [15 marks total]

(a) The eigenvectors of $\mathbf{A}^\mathsf{T}\mathbf{A}$ are

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Find the corresponding eigenvalues. [**6** marks]

$$\mathbf{A}^\mathsf{T}\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 5 & 6 \end{bmatrix}$$

If $\mathbf{u}$ is an eigenvector, $\mathbf{Au} = \lambda\mathbf{u}$ where λ is the associated eigenvalue. So

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 6 & 5 \\ 5 & 6 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{\lambda}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \implies 11 = \lambda.$$

Similarly, the eigenvalue associated with $\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is 1.

ID: ______________________________

(b) Given that the eigenvectors of $\mathbf{A}\mathbf{A}^\mathsf{T}$ corresponding to the two largest eigenvalues are (in order of descending eigenvalues)

$$\frac{1}{\sqrt{22}}\begin{bmatrix}2\\3\\3\end{bmatrix} \text{ and } \frac{1}{\sqrt{2}}\begin{bmatrix}0\\-1\\1\end{bmatrix},$$

find the singular value decomposition of $\mathbf{A}$. [**3** marks]

The singular value decomposition is $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^\mathsf{T}$ where $\mathbf{D}$ is the diagonal matrix of singular values ordered from largest to smallest, $\mathbf{U}$ is the matrix with columns the normalised eigenvectors of $\mathbf{A}\mathbf{A}^\mathsf{T}$ and $\mathbf{V}$ is the matrix with columns the normalised eigenvectors of $\mathbf{A}\mathbf{A}^\mathsf{T}$. The ith singular value is the square root of the ith eigenvalue of $\mathbf{A}^\mathsf{T}\mathbf{A}$. So

$$\mathbf{A} = \underbrace{\begin{bmatrix}2/\sqrt{22} & 0\\ 3/\sqrt{22} & -1/\sqrt{2}\\ 3/\sqrt{22} & 1/\sqrt{2}\end{bmatrix}}_{\mathbf{U}}\underbrace{\begin{bmatrix}\sqrt{11} & 0\\ 0 & 1\end{bmatrix}}_{\mathbf{D}}\underbrace{\begin{bmatrix}-1/\sqrt{2} & 1/\sqrt{2}\\ 1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}}_{\mathbf{V}^\mathsf{T}}.$$

(c) Write an expression for an approximation, $\widehat{\mathbf{A}}_1$, of $\mathbf{A}$ (you can leave it in the form of a matrix product). [**3** marks]

$$\widehat{\mathbf{A}}_1 = \sigma_1\mathbf{u}_1\mathbf{v}_1^\mathsf{T} = \sqrt{11}\cdot\frac{1}{\sqrt{22}}\begin{bmatrix}2\\3\\3\end{bmatrix}\cdot\frac{1}{\sqrt{2}}\begin{bmatrix}-1 & 1\end{bmatrix} = \frac{1}{2\sqrt{11}}\begin{bmatrix}2\\3\\3\end{bmatrix}\begin{bmatrix}-1 & 1\end{bmatrix}$$

(d) Write down the pseudo-inverse of $\mathbf{A}$, $\mathbf{A}^+$ (you can leave it in the form of a matrix product). [**3** marks]

ID: ______________________________

The pseudo-inverse is $\mathbf{A}^{+} = \mathbf{V}\mathbf{D}^{+}\mathbf{U}^{\mathsf{T}}$ where $\mathbf{D}^{+} = \mathbf{D}^{-1} = \begin{bmatrix} \frac{1}{\sqrt{11}} & 0 \\ 0 & 1 \end{bmatrix}$

ID: ____________________

4. Let $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$. [8 marks total]

   (a) Why do we call the linear system $\mathbf{Au} = \mathbf{b}$, where $\mathbf{u}$ is unknown, over-determined for the given matrix $\mathbf{A}$? [**1** mark]

   $m > n$. That is, there are more columns (equations) than rows (parameters).

   (b) Let $\mathbf{A} = \mathbf{QR}$ be a QR-decompostion of $\mathbf{A}$. What are the dimensions and properties of $\mathbf{Q}$ and $\mathbf{R}$? [**3** marks]

   $\mathbf{Q}$ is an $m \times n$ matrix with n orthonormal columns and $\mathbf{R}$ is an $n \times n$ upper triangular matrix.

   (c) Find the first column of the matrix $\mathbf{Q}$ using the Gram-Schmidt process. [**2** marks]

   The first column of $\mathbf{Q}$ is $\mathbf{Q}_1 = \frac{\mathbf{v}_1}{||v_1||}$ where $v_1 = \mathbf{A}_1$. Since $||v_1|| = \sqrt{1+1+1+1} = 2$, $\mathbf{Q}_1 = [0.5\ 0.5\ 0.5\ 0.5]^T$.

   (d) Given $\mathbf{Q}$, describe how you could find the matrix $\mathbf{R}$. [**2** marks]

   $\mathbf{A} = \mathbf{QR}$, $\mathbf{R} = \mathbf{Q^{-1}A} = \mathbf{Q^TA}$ since $\mathbf{Q}$ is orthonormal.

ID: ____________________

5. Suppose we wish to study the annual mutation rate in some basic organism so we run five independent trials for one year each and count the number of mutations we see. The resulting data vector is $\mathbf{x} = (x_1, \ldots, x_5) = (15, 9, 16, 19, 16)$. [8 marks total]

   (a) Explain why a Poisson distribution could be an appropriate model for each trial. [**2** marks]

   Mutations are rare events that happen independently of one another

   (b) Given the data and the probability density function for a Poisson random variable with rate parameter λ is

   $$f(x) = \exp(-\lambda)\frac{\lambda^x}{x!},$$

   write down the likelihood for λ. [**3** marks]

   Since each trial is independent, we can write the likelihood as a product of the individual densities, so

   $$L(\lambda; \mathbf{x}) = P(\mathbf{x}|\lambda) = \prod_{i=1}^{5} \exp(-\lambda)\frac{\lambda^{x_i}}{x_i!} = \exp(-5\lambda)\lambda^{\sum_i x_i} \prod_{i=1}^{5} \frac{1}{x_i!}$$

   (c) Explain why we often prefer to work with the log-likelihood and write down an expression for the log-likelihood. [**3** marks]

   We work with log-likelihoods to avoid numerical underflow and to use simpler arithmetic operations.

   $$\log L(\lambda; \mathbf{x}) = \prod_{i=1}^{5} \exp(-\lambda)\frac{\lambda^{x_i}}{x_i!} = \sum_{i=1}^{5} (-\lambda + x_i \log(\lambda) - \log(x_i!))$$

ID: ______________________________

6. Suppose we have a method, called `randunif()`, that returns a uniform random variate between 0 and 1. [6 marks total]

   (a) Write the body of pseudo-code method, `randbern(p)`, that returns a Bernoulli random variate with parameter p. Include a check that the input parameter is legal. [**3** marks]

```
function randbern(p)
% check parameter
if (p < 0 or p > 1) error(p must be between 0 and 1)
% generate value if randunif() < p
return 1
else
return 0
```

   (b) Write the body of pseudo-code method, `randbin(n,p)`, that returns a binomial random variate with parameters n and p (you can assume the `randbern(p)` method). Include checks that the input parameters are legal. [**3** marks]

```
function randbin(p)
% check parameter
if (n < 0) error(n must be a positive integer)
%sum n Bernoullis
count = 0
for (i in 1:n)
count = count + randbern(p)

return count
```

7. Suppose that the sequence of random variables $X_0, X_1, X_2, \ldots$ has the Markov property. Then simplify the following: [**2** marks]

$$P(X_{n+1}|X_n, X_{n-1}, \ldots, X_0) = P(X_{n+1}|X_n)$$

ID: ______________________________

Blank page — will not be marked