THE UNIVERSITY OF AUCKLAND

FIRST SEMESTER, 2011 – MID-SEMESTER TEST
Campus: City

COMPUTER SCIENCE

Computational Science

(Time allowed: 50 minutes)

NOTE: Attempt *all* questions

Use of calculators is NOT permitted.

Put your answers in the answer boxes provided below each question. You may use the blank page at the end of the test script for scratch work, which will not be marked.

Section:	**A**	**B**	**C**	**D**	Total
Possible marks:	15	10	12	13	50
Awarded marks:					

SURNAME:

FIRSTNAME:

ID:

ID: ______________________________

Section A: Linear Systems, SVD, PCA (10 marks)

1. Which types of matrices appear in the LU decomposition, $\mathbf{A} = \mathbf{LU}$, and QR decomposition, $\mathbf{A} = \mathbf{QR}$, of a square $m \times m$ matrix $\mathbf{A}$. [2 marks]

   $\mathbf{L}$: ($\times$) ______________________ matrix
   $\mathbf{U}$: ($\times$) ______________________ matrix
   $\mathbf{Q}$: ($\times$) ______________________ matrix
   $\mathbf{R}$: ($\times$) ______________________ matrix

2. What decomposition is done by Gauss elimination and Gram-Schmidt orthogonalisation? [2 marks]

   Gauss elimination: ______ decomposition
   Gram-Schmidt orthogonalisation: ______ decomposition

3. The Singular Value Decomposition (SVD) of an ordinary $m \times n$ matrix $\mathbf{A}$, $m \gg n$, is represented by a product, $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^\mathsf{T}$, of three matrices with special properties. In particular, $\mathbf{U}$ is an $m \times n$ column-orthogonal matrix with the following properties: its n columns are the n top-rank eigenvectors $\mathbf{u}_1$,…,$\mathbf{u}_n$ of the $m \times m$ symmetric matrix $\mathbf{A}\mathbf{A}^\mathsf{T}$, i.e. the only n eigenvectors that can have non-zero eigenvalues. Specify the properties of the matrices $\mathbf{D}$ and $\mathbf{V}$.

   Hint: Consider the eigenvectors and eigenvalues $\{(\mathbf{v}_i, \lambda_i) : i = 1, \ldots, n\}$, of the $n \times n$ matrix $\mathbf{A}^\mathsf{T}\mathbf{A}$. [3 marks]

4. Let $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_n]$; $m \ll n$, be an $m \times n$ matrix of n centred $m \times 1$ measurement vectors $\mathbf{a}_i$; $i = 1, \ldots, n$. Let $\mathbf{u}_1$ and $\mathbf{u}_2$ be two principal components, i.e. the eigenvectors of the $m \times m$ covariance matrix of the measurements, $\mathbf{\Sigma}_n = \frac{1}{n-1}\mathbf{A}\mathbf{A}^\mathsf{T}$, with the largest eigenvalues. Write down the projection matrix $\mathbf{P}_2$, which projects the measurement vectors (columns of $\mathbf{A}$) to the principal subspace of the eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$.

   [3 marks]

   $\mathbf{P}_2 =$

ID: ____________________

Section B: Root Finders, Optimisation, Least Squares (15 marks)

5. Explain how the root of the non-linear function $g(x) = x^4 + 2x^2 - 25$ is being searched for with the **modified Newton method** starting from the first approximation $x^{[0]} = 2$ and derive the approximate root value $x^{[1]}$ at the first step of the search.

   Hint: In this particular case the Jacobian $J(x)$ of the function $g(x)$ is the first derivative $J(x) = \frac{dg(x)}{dx}$ at point x. The **non-modified Newton method** chooses at each iteration t the root $x^{[t+1]}$ of the linear approximation $g_{\text{app}:t}(x) = g(x^{[t]}) + J(x^{[t]})\left(x - x^{[t]}\right)$, i.e. $g_{\text{ap}:t}(x^{[t+1]}) = 0$. Recall what is different in the **modified** method. [4 marks]

6. At each iteration of unconstrained gradient maximisation of a scalar bivariate function $f(x, y)$, the gradient, $\nabla f(x, y) = \left(\frac{\partial f(x,y)}{\partial x}, \frac{\partial f(x,y)}{\partial y}\right)$, is computed at the current point (x_i, y_i), and the next point, (x_{i+1}, y_{i+1}), is selected along a line $L_i(t) = (x_i + tu_i, y_i + tv_i)$ through (x_i, y_i) in the gradient direction $\left(u_i = \frac{\partial f(x,y)}{\partial x}\Big|_{x=x_i, y=y_i}, v_i = \frac{\partial f(x,y)}{\partial y}\Big|_{x=x_i, y=y_i}\right)$. Write down in the general form the condition for selecting the next steepest ascent point (x_{i+1}, y_{i+1}).

   [3 marks]

ID: ____________________

7. You have to maximise the function $f(x,y) = x^3 - x^2y + 4xy^2 + y^3$ subject to the constraint $x^2 + y^2 = 4$. Write down the Lagrangian $F(x,y,\lambda)$ and derive the system of equations that give all the stationary (i.e. maximum, minimum, and saddle) points of the Lagrangian. You need not solve this system. [4 marks]

8. You have to solve an overdetermined system $\mathbf{Ax} = \mathbf{a}$ of m linear equations with $n \times 1$; $n \ll m$, vector $\mathbf{x}$ of unknowns, provided that the $m \times n$ matrix $\mathbf{A}$ and the $m \times 1$ vector $\mathbf{a}$ are known and the matrix $\mathbf{A}$ has n linearly independent columns.

   Derive the "normal" equation and the least squares solution for this system.

   Hint: Consider the residual error vector $\mathbf{e_x} = \mathbf{Ax} - \mathbf{a}$ and recall that the least squares solution $\mathbf{x}^*$ minimises the total squared error $E(\mathbf{x}) = |\mathbf{e_x}|^2 \equiv \mathbf{e_x^\mathsf{T}} \mathbf{e_x}$. [4 marks]

ID: ______________________________

Section C: Probability Models, Finite Differences (12 marks)

9. You have two probability models for scalar random variables $x_1, x_2, x_3, x_4, x_5, , x_6$:

   (a) $\Pr(x_1, \dots, x_6) = p_1(x_1)p_2(x_2|x_1)p_3(x_3|x_2)p_4(x_4|x_3)p_5(x_5|x_4)p_6(x_6|x_5)$ and

   (b) $\Pr(x_1, \dots, x_6) = p_1(x_1)p_2(x_2)p_3(x_3)p_4(x_4|x_1, x_2)p_5(x_5|x_1, x_2, x_3)p_6(x_6|x_4, x_5)$

   in terms of the known unconditional and conditional probability distributions $p_j(x_j)$, $j = 1, \dots, 6$, $p_j(x_j|x_{j-1})$; $j = 2, \dots, 6$, $p_4(x_4|x_1, x_2)$, $p_6|x_4, x_5)$, and $p_5(x_j|x_1, x_2, x_3)$.
   Name each model and draw a digraph showing causal relationships between its variables.

   [5 marks]

(a) $\Pr(x_1, \dots, x_6) = p_1(x_1)p_2(x_2|x_1)p_3(x_3|x_2)p_4(x_4|x_3)p_5(x_5|x_4)p_6(x_6|x_5)$

Name of the model:

Digraph:

(b) $\Pr(x_1, \dots, x_6) = p_1(x_1)p_2(x_2)p_3(x_3)p_4(x_4|x_1, x_2)p_5(x_5|x_1, x_2, x_3)p_6(x_6|x_4, x_5)$

Name(s) of the model:

Digraph:

10. Which nodes in the digraph showing the causal relationships in the above probability model (9b) form the Markov blanket of the node for the variable x_4?

    Hint: The Markov blanket of a node is the set of its parents, children, and co-parents of its children. [3 marks]

ID: ______________________________

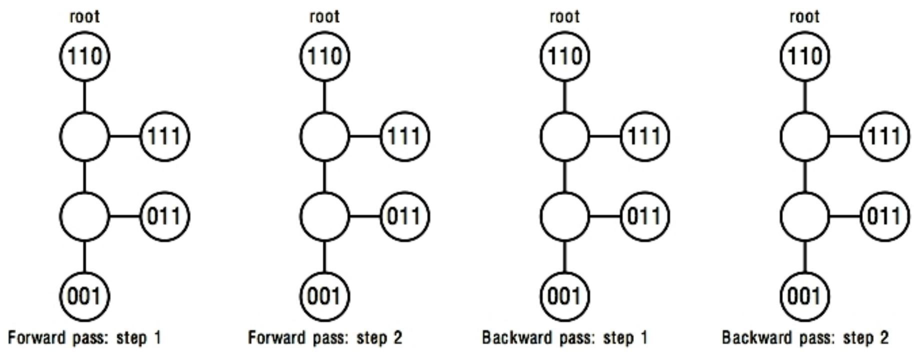
11. Assuming only integer values of x, derive the centred finite difference approximation for the first derivative of function $u(x) = x^4$ and compare it to the exact derivative $\frac{d}{dx}u(x)$ at point $x = 2$.

    Hint: Recall that $(a \pm b)^4 = a^4 \pm 4a^3b + 6a^2b^2 \pm 4ab^3 + b^4$ and $\frac{dx^n}{dx} = nx^{n-1}$. [4 marks]

Section D: Dynamic Programming, Search Methods (13 marks)

12. Consider a collection $C = \{001, 011, 111, 110\}$ of 4 strings of length 3 over the binary alphabet $\{0, 1\}$, which are the leaves of the tree T below. Using the leave 110 as the root, perform the forward and backward passes of the Fitch's dynamic programming algorithm to find sets of the best candidate strings and pick the strings, respectively, for each of the two internal (parent) nodes, calculate the Hamming distances between the nodes, and find the optimum parsimony score for T.

    Hint: The forward pass computes a set of best candidate strings for each internal (non-leaf) node and the backward pass picks a label from the set for each internal node while computing an optimum parsimony score. Use the notation $\alpha[\beta : \gamma][\delta : \epsilon]$ for the set of strings $\alpha\beta\delta$ and $\alpha\gamma\epsilon$. [4 marks]

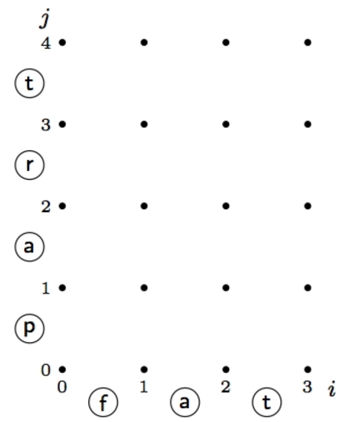


Optimum parsimony score =

ID: ____

13. Draw a weighed digraph for computing the edit distance $D(C_1, C_2)$ between two strings, $C_1 = \mathrm{fat}$ and $C_2 = \mathrm{part}$, assuming the unit weight, $\delta = 1$, of inserting or deleting a character; the unit weight, $\alpha_{i,j} = 1$, for substituting a j-th character $c_{2:j}$ of C_2 for a different i-th character $c_{1:i}$ of C_1, such that $c_{2:j} \neq c_{1:i}$, and zero weight, $\alpha_{i,j} = 0$, for substituting the same character, $c_{2:j} = c_{1:i}$. Show the unit and zero weights of the graph edges with solid and dashed lines, respectively.

    Specify recurrent computations of the edit distance $d_{i,j} = D\left(C_{1:[i]}, C_{2:[j]}\right)$ for transforming a substring $C_{1:i} = c_{1:0} \cdots c_{1:i}$ into a substring $C_{2:j} = c_{2:0} \cdots c_{2:j}$. The indices $i = 0$ and $j = 0$ relate to empty substrings, and $i = 1, 2, 3$ and $j = 1, 2, 3, 4$ correspond to the substrings ending on the characters $c_{1:i}$ and $c_{2:j}$, respectively, e.g. $C_{1:1} = \mathrm{f}$ or $C_{2:3} = \mathrm{par}$. [4 marks]



$$d_{i,j} = \begin{cases} 0 & \text{if} \quad i = j = 0 \\ & \text{if} \quad i > 0;\ j = 0 \\ & \text{if} \quad i = 0;\ j > 0 \\ & \text{if} \quad i > 0;\ j > 0 \end{cases}$$

14. Which types of search through an implicit graph of possible solutions are performed by backtracking and branch-and-bound methods? [2 marks]

    Backtracking:

    Branch-and-bound:

15. You have to minimise a function $f(x_1, x_2, \ldots, x_n)$ of discrete variables $x_i \in \{0, 1, \ldots, K\}$, which has for any fixed subsequence $(x_1^\circ, \ldots, x_j^\circ)$; $j = 1, \ldots, n-1$, easily computable lower and upper bounds $L(x_1^\circ, \ldots, x_j^\circ) \leq f(x_1^\circ, \ldots, x_j^\circ, x_{j+1}, \ldots, x_n)\} \leq U(x_1^\circ, \ldots, x_j^\circ)$, which do not depend on the remaining variables $(x_{j+1}, \ldots, x_n)$.

    Explain in brief how these bounds can be used in the branch-and-bound search to prune the tree of all the K^n candidate solutions for discarding large subsets of fruitless candidates. [3 marks]

ID: ____________________

Blank page — will not be marked