Bottom-Up Computing and Discrete Mathematics

Answer THREE questions out of FIVE

Time allowed: 2 hours

Turn over . . .

1. (a) An n-component vector $v = (v_1, v_2, ..., v_n)$ is 'probablistic' if element $v_i \geq 0$ for all i and $\sum_i v_i = 1$. Carefully explain the sense in which such a vector might represent a candidate for the PageRank importance ranking of web pages. Given an initial guess v for this ranking, the PageRank algorithm uses the formula

$$v_i' = \frac{1-p}{\sigma} + p \sum_n \left(\frac{W_{ni} v_n}{\sum_m W_{nm}} \right)$$

to determine a new guess. Explain the meaning of all the terms in this formula. Under what conditions can we be sure that repeated use of this algorithm will result in convergence to a stable vector?

(b) Explain the sense in which the world wide web is represented by a digraph for the purposes of the PageRank importance ranking algorithm.

(c) Explain the meaning of the term *adjacency matrix* for a digraph. Explain how the adjacency matrix of the web is turned into a stochastic matrix for use in the PageRank algorithm.

(d) Explain how the stochastic matrix in (1c) above is combined with another stochastic matrix which models the possibility of a random jump to any point on the web with equal probability, to form the final stochastic matrix appropriate for the PageRank algorithm.

(e) Give the definition of the term *invariant measure* for a stochastic matrix.

(f) Compute the PageRank invariant measure associated to the model of the web with adjacency matrix

$$W = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

in case the PageRank parameter p is very close to 1.

Turn over ...

2. (a) Show that if two vectors are eigenvectors of a square matrix A with distinct eigenvalues then these eigenvectors are necessarily linearly independent.

   (b) Show that if two vectors are eigenvectors of a square matrix A with the same eigenvalue then any nonzero linear combination of these eigenvectors is also an eigenvector.

   (c) Assume that A is an $n \times n$ matrix with n distinct eigenvalues. Show that if B is another matrix and $AB = BA$ then every eigenvector of A is also an eigenvector of B.

   (d) Let D_n denote the $n \times n$ matrix with all entries equal to 1. For example

$$D_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

   Note that

$$\sum_{j=0}^{n-1} \exp(2k\pi i j/n) = 0$$

   for any $k \in \{1, 2, ..., n-1\}$ (here $i = \sqrt{-1}$). Using this result, or otherwise, determine a set of 5 pairwise linearly independent eigenvectors (and their associated eigenvalues) for D_5.

   (e) Compute the eigenvalues and eigenvectors of the matrix

$$M(p) = \frac{1-p}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + p \begin{pmatrix} 0 & +1 \\ +1 & 0 \end{pmatrix}$$

   as functions of p. Comment on your answer in terms of the PageRank importance ranking algorithm (that is, say what is the model of the web here; say what is the invariant measure, and how it ascribes importance to the web pages). What happens when $p = 1$?

   (f) Compute the eigenvalues and eigenvectors of each of

$$M_1 = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

   Comment, as in part (2e), on your answers in terms of PageRank.

Turn over ...

(g) Compute PageRank for

$$W = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

that is, compute the PageRank invariant measure vector for a suitable value of PageRank parameter p (if you are careful you can determine the ranking by considering $p = 1$ only).

3. (a) Prove that if M, N are $n \times n$ stochastic matrices then so is MN.

(b) Explain what is meant by the Jordan canonical form of a square matrix.

(c) Recall that the *spectral radius* of a square matrix is the magnitude of the eigenvalue with largest magnitude. Determine the spectral radius of

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -2i \end{pmatrix}$$

(d) By considering the Jordan form, or otherwise, show that the spectral radius of every stochastic matrix is 1.

4. (a) Give the set theoretic characterisation of a *directed graph*, and explain how this might be represented by a picture consisting of vertices and directed edges.

(b) Explain the sense in which the following statement is true: A *graph* is a special kind of directed graph.

(c) Let G and G' be graphs having no vertices in common. Then let $G \cup G'$ denote the natural disconnected composite graph (that is, the graph whose vertex set is the union of the vertex sets of G and G', and similarly for the edge sets).

Explain how one might construct the adjacency matrix of $G \cup G'$ from those of G and G'.

(d) Let $\mathcal{G}$ denote the set of finite simple loop–free graphs. Show that the operation $\cup$ defined above is a closed binary operation on $\mathcal{G}$ (i.e. if $G, G' \in \mathcal{G}$ then $G \cup G' \in \mathcal{G}$).

(e) Suppose there is a map C from $\mathcal{G}$ to the set of polynomials of finite degree in an indeterminate v, with the property that for all $Q \in \mathbb{N}$, $C(G)|_{v=Q}$ (that is, the polynomial *evaluated* at $v = Q$) is the number of ways of colouring G with Q colours. Explain why such a map, if it exists, is unique; and explain why each of the following must be true:

$$C(K_1) = v$$

$$C(G \cup G') = C(G).C(G')$$

(f) Let K_2 denote the complete graph on 2 vertices. Determine $C(K_2)$, showing all your working.

Turn over ...

5. (a) Recall that a *model* of the world wide web is any notional set of web pages together with the links between them.

Suppose some web author simply concatenated three web pages into a single page (combining all the html together) with the URL of the first of these pages. (So that the second and third URL ceased to contain a web page.) What would this do to the graph of the web (a) in the short term; (b) after other authors had adjusted their links accordingly?

Suppose that the PageRank invariant measure vector is already known for the web before this change. What might be a sensible starting point vector to begin iterating towards PageRank for the web *after* the change?

(b) A matrix is called *positive* if all its entries are positive. You may assume the following (from the Perron–Frobenius theorem): Every positive $n \times n$ matrix has a unique eigenvalue of largest magnitude; this eigenvalue is positive; and there is a corresponding eigenvector which also has all positive entries.

Show that the matrix

$$L = \begin{pmatrix} 1 & 0 & 2 & 0 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

has a largest magnitude eigenvalue which is positive. (N.B., You do *not* need to show that this largest magnitude eigenvalue is nondegenerate.)

(c) Using any appropriate notation, draw a graph showing how the following subset of modular components might be physically connected in a network of two PCs (denoted PC1 and PC2):

CPU1; CPU2; MOTHERBOARD1; MOTHERBOARD2; NETWORK-CARD1; NETWORK-CARD2; NETWORK-HUB; HARD-DISK1; HARD-DISK2; VIDEO-CARD1; MONITOR1; KEYBOARD1; KEYBOARD2.

Turn over . . .

Internal Examiner:	Professor P.P. Martin
External Examiners:	Professor D.J. Needham
	Professor M.E. O'Neill