TAKE-HOME CLASS QUIZ SOLUTIONS: DUE MONDAY NOVEMBER 25: STOCHASTIC MATRICES

MATH 196, SECTION 57 (VIPUL NAIK)

1. Performance review

22 people took this 10-question quiz. The score distribution was as follows:

- Score of 4: 2 people
- Score of 5: 2 people
- Score of 6: 5 people
- Score of 7: 4 people
- Score of 8: 5 people
- Score of 9: 1 person
- Score of 10: 3 people

The mean score was about 7.05.

The question-wise answers and performance review were as follows:

(1) Option (C): 11 people
(2) Option (C): 13 people
(3) Option (C): 18 people
(4) Option (E): 20 people
(5) Option (B): 21 people
(6) Option (D): 11 people
(7) Option (A): 16 people
(8) Option (C): 19 people
(9) Option (B): 15 people
(10) Option (B): 11 people

2. Solutions

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

This quiz can be viewed as a continuation of the quiz on linear dynamical systems. The book defines column-stochastic matrices using the jargon "transition matrix" on Page 53 (Definition 2.1.4) and uses them throughout the text when describing (a simplified version of) Google's PageRank algorithm. The quiz questions are self-contained and do not require you to read the book, but you may benefit from skimming through the book's discussion of PageRank to complement these questions. Note that the 4th Edition does not include the discussion of transition matrices and PageRank.

In this quiz, we discuss the dynamics of a very special type of linear transformation. A $n \times n$ matrix A is termed a *row-stochastic matrix* if all its entries are in the interval $[0, 1]$ and all the row sums are equal to 1. A $n \times n$ matrix is termed a *column-stochastic matrix* if all its entries are in the interval $[0, 1]$ and all the column sums are equal to 1. A $n \times n$ matrix A is termed a *doubly stochastic matrix* if it is both row-stochastic and column-stochastic, i.e., all the entries are in the interval $[0, 1]$, all the row sums are equal to 1, and all the column sums are equal to 1.

(1) Suppose A and B are two $n \times n$ row-stochastic matrices. Which of the following is *guaranteed* to be row-stochastic? Please see Options (D) and (E) before answering.
   (A) $A + B$
   (B) $A - B$
   (C) AB

(D) All of the above
(E) None of the above

Answer: Option (C)

Explanation: Suppose we are trying to compute the $(ik)^{th}$ entry of AB. This is the sum:

$$\sum_{j=1}^{n} a_{ij} b_{jk}$$

We now want to sum up all such entries in the i^{th} row of AB. Thus, the sum is:

$$\sum_{k=1}^{n} \sum_{j=1}^{n} a_{ij} b_{jk}$$

The sum can be rearranged as:

$$\sum_{j=1}^{n} \left(a_{ij} \sum_{k=1}^{n} b_{jk} \right)$$

Each of the inner sums is 1, on account of being a row sum of B. Thus, the sum simplifies to:

$$\sum_{j=1}^{n} a_{ij}$$

This is 1, on account of being a row sum of A.

Thus, every row sum of AB is 1. Further, because of the way we define matrix multiplication, all the entries of AB are nonnegative. Combined with the condition on sums, we get that all entries are in $[0, 1]$ with all row sums 1. Thus, the matrix AB is row-stochastic.

As for Options (A) and (B), note that for Option (A), the row sums will become 2 and for Option (B), the row sums will become 0.

Performance review: 11 out of 22 got this. 10 chose (E), 1 chose (A).

Historical note (last time): 14 out of 24 got this. 9 chose (E), 1 chose (D).

(2) Suppose A and B are two $n \times n$ column-stochastic matrices. Which of the following is *guaranteed* to be column-stochastic? Please see Options (D) and (E) before answering.

(A) $A + B$
(B) $A - B$
(C) AB
(D) All of the above
(E) None of the above

Answer: Option (C)

Explanation: Suppose we are trying to compute the $(ik)^{th}$ entry of AB. This is the sum:

$$\sum_{j=1}^{n} a_{ij} b_{jk}$$

We now want to sum up all such entries in the k^{th} column of AB. Thus, the sum is:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{jk}$$

The sum can be rearranged as:

$$\sum_{j=1}^{n} \left(b_{jk} \sum_{i=1}^{n} a_{ij} \right)$$

Each of the inner sums is 1, on account of being a row sum of A. Thus, the sum simplifies to:

$$\sum_{j=1}^{n} b_{jk}$$

This is 1, on account of being a column sum of B.

Thus, every column sum of AB is 1. Further, because of the way we define matrix multiplication, all the entries of AB are nonnegative. Combined with the condition on sums, we get that all entries are in $[0,1]$ with all row sums 1. Thus, the matrix AB is column-stochastic.

As for Options (A) and (B), note that for Option (A), the column sums will become 2 and for Option (B), the column sums will become 0.

Performance review: 13 out of 22 got this. 8 chose (E), 1 chose (A).

Historical note (last time): 12 out of 24 got this. 11 chose (E), 1 chose (A).

(3) Suppose A and B are two $n \times n$ doubly stochastic matrices. Which of the following is *guaranteed* to be doubly stochastic? Please see Options (D) and (E) before answering.
(A) $A + B$
(B) $A - B$
(C) AB
(D) All of the above
(E) None of the above

Answer: Option (C)

Explanation: This follows by combining the two preceding questions.

Performance review: 18 out of 22 got this. 3 chose (E), 1 chose (A).

Historical note (last time): 11 out of 24 got this. 8 chose (D), 4 chose (E), 1 chose (A).

We now consider the case $n = 2$. In this case, the doubly stochastic matrices have the form:

$$\begin{bmatrix} a & 1-a \\ 1-a & a \end{bmatrix}$$

where $a \in [0,1]$. Denote this matrix by $D(a)$ for short.

(4) Suppose $a, b \in [0,1]$ (they are allowed to be equal). The product $D(a)D(b)$ equals $D(c)$ for some $c \in [0,1]$. What is that value of c?
(A) $a + b$
(B) ab
(C) $2ab + a + b$
(D) $(1-a)(1-b)$
(E) $1 - a - b + 2ab$

Answer: Option (E)

Explanation: We carry out the multiplication:

$$\begin{bmatrix} a & 1-a \\ 1-a & a \end{bmatrix} \begin{bmatrix} b & 1-b \\ 1-b & b \end{bmatrix} = \begin{bmatrix} ab+(1-a)(1-b) & a(1-b)+(1-a)b \\ a(1-b)+(1-a)b & ab+(1-a)(1-b) \end{bmatrix} = D(ab+(1-a)(1-b)) = D(1-a-b+2ab)$$

Performance review: 20 out of 22 got this. 1 each chose (A) and (D).

Historical note (last time): 19 out of 24 got this. 2 chose (B), 1 each chose (A), (C), and (D).

(5) For what value(s) of a is the matrix $D(a)$ non-invertible? Note that when judging invertibility, we do not insist that the inverse matrix also be doubly stochastic.
(A) $a = 0$ only
(B) $a = 1/2$ only
(C) $a = 1$ only
(D) $0 < a < 1$ (i.e., $D(a)$ is invertible only at $a = 0$ and $a = 1$)
(E) $a \neq 1/2$

Answer: Option (B)

Explanation: For the matrix to be non-invertible, we need the rows to be scalar multiples of each other. Since both row sums are 1, this can happen only if the rows are identical, which happens iff $a = 1/2$. Equivalently, we can note that invertibility requires a nonzero determinant, and the determinant is $a^2 - (1-a)^2 = 2a - 1$, which is 0 iff $a = 1/2$.

Performance review: 21 out of 22 got this. 1 chose (A).

Historical note (last time): 15 out of 24 got this. 4 chose (A), 2 each chose (C) and (D), 1 chose (E).

(6) For what value(s) of a is it true that the matrix $D(a)$ does not have an inverse that is a doubly stochastic matrix? In other words, either $D(a)$ should be non-invertible or it should be invertible but the inverse is not a doubly stochastic matrix.
(A) $a = 0$ only
(B) $a = 1/2$ only
(C) $a = 1$ only
(D) $0 < a < 1$ (i.e., $D(a)$ has an inverse that is also doubly stochastic only if $a = 0$ or $a = 1$)
(E) $a \neq 1/2$

Answer: Option (D)

Explanation: The inverse of $D(a)$, for $a \neq 1/2$, is:

$$\begin{bmatrix} a/(2a-1) & (a-1)/(2a-1) \\ (a-1)/(2a-1) & a/(2a-1) \end{bmatrix}$$

For this to be doubly stochastic, we need that $0 \leq a/(2a-1) \leq 1$. We make cases:

- $a = 0$: In this case, $D(a)$ is its own inverse.
- $a \neq 0$ (excluding $a = 1/2$): In this case, since $0 \leq a/(2a-1)$, we obtain that $2a - 1 > 0$. So, starting with $a/(2a-1) \leq 1$ we get $a \leq 2a - 1$, which simplifies to $1 \leq a$, forcing $a = 1$ (since we are constrained by $0 \leq a \leq 1$). The choice $a = 1$ works (in the sense that $D(1)$ is its own inverse).

The upshot is that the only cases where the inverse is doubly stochastic are the cases $a = 0$ and $a = 1$. Otherwise, the inverse either does not exist (case $a = 1/2$) or exists but is not doubly stochastic.

Performance review: 11 out of 22 got this. 9 chose (B), 2 chose (C).

Historical note (last time): 9 out of 24 got this. 9 chose (A), 2 each chose (B), (C), and (E).

For the next few questions, denote by T_a the linear transformation whose matrix is $D(a)$. For any vector $\vec{x} \in \mathbb{R}^2$, we can consider the sequence:

$$\vec{x}, T_a(\vec{x}), T_a^2(\vec{x}), \dots$$

Note that if we were to start with a vector $\vec{x} \in \mathbb{R}^2$ with both coordinates equal, it would be invariant under T_a.

Thus, for the questions below, assume that we start with a nonzero vector $\vec{x} \in \mathbb{R}^2$ for which the two coordinates are not equal to each other.

(7) For what value of a is it the case that $\lim_{r \to \infty} T_a^r(\vec{x})$ does *not* exist?
(A) $a = 0$ only
(B) $a = 1/2$ only
(C) $a = 1$ only
(D) $0 < a < 1$
(E) $a \neq 1/2$

Answer: Option (A)

Explanation: In this case, applying T_a interchanges the coordinates. A second application interchanges them back. The sequence thus cycles between the vector $\vec{x}$ and the vector obtained by interchanging its coordinates.

The failure of the remaining options will become clear from the answers to the rest of the questions.

Performance review: 16 out of 22 got this. 2 chose (B), 1 each chose (C), (D), and (E). 1 left the question blank.

Historical note (last time): 7 out of 24 got this. 11 chose (B), 4 chose (C), 1 each chose (D) and (E).

(8) For what value of a is it the case that the sequence

$$\vec{x}, T_a(\vec{x}), T_a^2(\vec{x}), \dots$$

is a constant sequence?
(A) $a = 0$ only
(B) $a = 1/2$ only
(C) $a = 1$ only
(D) $0 < a < 1$
(E) $a \neq 1/2$

Answer: Option (C)

Explanation: $a = 1$ gives the identity transformation. Any other choice of a replaces the first coordinate by a combination of the two coordinates that cannot be equal to it unless the two coordinates were equal to begin with. Explicitly:

$$D(a)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} ax_1 + (1-a)x_2 \\ (1-a)x_1 + ax_2 \end{bmatrix}$$

For the output to equal the input, we need that:

$$\begin{aligned} x_1 &= ax_1 + (1-a)x_2 \\ x_2 &= (1-a)x_1 + ax_2 \end{aligned}$$

Solving the first equation alone gives $x_1 = x_2$ or $a = 1$. By assumption, $x_1 \neq x_2$, so we get $a = 1$. Note that the second equation yields a similar conclusion.

Performance review: 19 out of 22 got this. 2 chose (B), 1 chose (A).

Historical note (last time): 18 out of 24 got this. 3 chose (D), 2 chose (E), 1 chose (B).

(9) For what value of a is it the case that the sequence

$$\vec{x}, T_a(\vec{x}), T_a^2(\vec{x}), \dots$$

is not a constant sequence but becomes constant from $T_a(\vec{x})$ onward?
(A) $a = 0$ only
(B) $a = 1/2$ only
(C) $a = 1$ only
(D) $0 < a < 1$
(E) $a \neq 1/2$

Answer: Option (B)

Explanation: If the sequence is not constant, then $a \neq 1$. However, it becomes constant from $T_a(\vec{x})$ onward. Thus, $T_a(\vec{x})$ has both coordinates equal by the previous question. Hence, we get that:

$$ax_1 + (1-a)x_2 = (1-a)x_1 + ax_2$$

This simplifies to:

$$(2a-1)(x_1 - x_2) = 0$$

Thus, either $a = 1/2$ or $x_1 = x_2$. Since $x_1 \neq x_2$ by assumption, we get $a = 1/2$.

Performance review: 15 out of 22 got this. 3 chose (A), 2 each chose (C) and (D).

Historical note (last time): 8 out of 24 got this. 10 chose (A), 4 chose (C), 1 each chose (D) and (E).

(10) For a other than 0, 1/2, or 1, what is the limit $\lim_{r \to \infty} (D(a))^r$? Here, when we talk of taking the limit of a sequence of matrices, we are taking the limit entry-wise.
(A) The matrix $D(0)$
(B) The matrix $D(1/2)$
(C) The matrix $D(1)$
(D) The matrix $D(a)$
(E) The matrix $D(1 - a)$

Answer: Option (B)

Explanation: Intuitively, what's happening is that we start off with the point:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

We now make the coordinates "come closer to each other" by converting to:

$$\begin{bmatrix} ax_1 + (1-a)x_2 \\ (1-a)x_1 + ax_2 \end{bmatrix}$$

We then iterate. Each time, the coordinates are coming closer to each other, but note also that the sum of the coordinates remains fixed. Thus, we hope to eventually converge to the vector with both coordinates $(x_1 + x_2)/2$. This is $D(1/2) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

Here is a more formal demonstration: By one of the previous questions, we can verify that if $D(a)D(b) = D(c)$, then $(2a-1)(2b-1) = 2c-1$. An immediate corollary is that if $(D(a))^r = D(u)$, then $(2a-1)^r = 2u-1$. As a result, if $\lim_{r\to\infty}(D(a))^r = D(u)$, then $2u - 1 = \lim_{r\to\infty}(2a-1)^r = 0$, forcing $u = 1/2$.

The geometric intuition for why we look at $2a - 1$ is because that is the ratio of the signed difference between the output coordinates to the signed difference between the input coordinates. It describes an contraction factor, and the contraction factors multiply when we compose.

Performance review: 11 out of 22 got this. 5 chose (A), 4 chose (E), 1 each chose (C) and (D).

Historical note (last time): 3 out of 24 got this. 10 chose (D), 7 chose (E), 2 each chose (A) and (C).