Feedback — Week5A (Advanced)

Help Center

You submitted this quiz on **Wed 4 Mar 2015 8:13 PM PST**. You got a score of **3.00** out of **3.00**.

Question 1

Consider the diagonal matrix M =

1	0	0
0	2	0
0	0	0

. Compute its Moore-Penrose pseudoinverse, and then identify, in the list below, the true statement about the elements of the pseudoinverse.

Your Answer		Score	Explanation
There is one element with value -1.			
There is one element with value 1.	✔	1.00	
There are seven elements with value infinity.			
There is one element with value 2.			
Total		1.00 / 1.00	

Question Explanation

The pseudoinverse has 0's off the diagonals. The diagonal elements are each 1 divided by the corresponding diagonal element of the given matrix M, with the exception of those diagonal elements of M that are 0. For those elements, the pseudoinverse has 0, rather than infinity. Thus, the pseudoinverse of M is

1	0	0
0	1/2	0
0	0	0

.

Question 2

An ad publisher selects three ads to place on each page, in order from the top. Click-through rates (CTR's) at each position differ for each advertiser, and each advertiser has a different CTR for each position. Each advertiser bids for click-throughs, and each advertiser has a daily budget, which may not be exceeded. When a click-through occurs, the advertiser pays the amount they bid. In one day, there are 101 click-throughs to be auctioned.

Here is a table of the bids, CTR's for positions 1, 2, and 3, and budget for each advertiser.

Advertiser	Bid	CTR1	CTR2	CTR3	Budget
A	$.10	.015	.010	.005	$1
B	$.09	.016	.012	.006	$2
C	$.08	.017	.014	.007	$3
D	$.07	.018	.015	.008	$4
E	$.06	.019	.016	.010	$5

The publisher uses the following strategy to allocate the three ad slots:

1. Any advertiser whose budget is spent is ignored in what follows.
2. The first slot goes to the advertiser whose expected yield for the first slot (product of the bid and the CTR for the first slot) is the greatest. This advertiser is ignored in what follows.
3. The second slot goes to the advertiser whose expected yield for the second slot (product of the bid and the CTR for the second slot) is the greatest. This advertiser is ignored in what follows.
4. The third slot goes to the advertiser whose expected yield for the third slot (product of the bid and the CTR for the third slot) is the greatest.

The same three advertisers get the three ad positions until one of two things happens:

1. An advertiser runs out of budget, or
2. All 101 click-throughs have been obtained.

Either of these events ends one *phase* of the allocation. If a phase ends because an advertiser ran out of budget, then they are assumed to get all the clicks their budget buys. During the same phase, we calculate the number of click-throughs received by the other two advertisers by assuming that all three received click-throughs in proportion to their respective CTR's for their positions (round to the nearest integer). If click-throughs remain, the publisher reallocates all three slots and starts a new phase.

If the phase ends because all click-throughs have been allocated, assume that the three advertisers received click-throughs in proportion to their respective CTR's (again, rounding if necessary).

Your task is to simulate the allocation of slots and to determine how many click-throughs each of the five advertisers get.

Your Answer	Score	Explanation

C gets 22 click-throughs.

C gets 37 click-throughs.

A gets 10 click-throughs. ✔ 1.00

B gets 0 click-throughs.

Total 1.00 / 1.00

Question Explanation

To begin, we compare the product of the bid and CTR1 for each of A through E, and we get .0015, .00144, .00136, .00126, and .00114, respectively. Thus, A gets the first slot. A similar comparison of the product of bids and CTR2 for B through E tells us C gets the second slot, with a product of .00112. Then, among B, D, and E, the best product of bid and CTR3 is E, with a product of .0006. Thus, the first phase can be summarized as follows:

Slot	Advertiser	CTR	Click-throughs
1	A	.015	10
2	C	.014	9
3	E	.010	7

The first phase ends when A gets 10 click-throughs and runs out of budget. We see in the table above the number of clicks that C and E get. For example, C, whose CTR is 14/15-ths of the CTR of A will get 14*10/15 = 9.33, or 9 (rounded).

For the second phase, A is no longer eligible, and B wins the first slot. However, C and E retain the second and third slots, respectively. The second phase ends when B runs out of budget, after getting 22 click-throughs. Note that B retains 2 cents of his budget, but that is not enough for another click-through, so they are effectively out of budget. Here is the summary of the second phase.

Slot	Advertiser	CTR	Click-throughs
1	B	.016	22
2	C	.014	19
3	E	.010	14

At this point, 81 of the 101 click-throughs have been allocated. For the third phase, A and B are out. The winners for the three slots are C, D, and E in that order. The third phase ends when the 20 clicks are allocated. This phase is summarized as follows:

Slot	Advertiser	CTR	Click-throughs
1	C	.017	8
2	D	.015	7
3	E	.010	5

Summing up the click-throughs for the three phases, we find the total for each advertiser to be A:10, B:22, C:36, D:7, and E:26.

Question 3

In certain clustering algorithms, such as CURE, we need to pick a representative set of points in a supposed cluster, and these points should be as far away from each other as possible. That is, begin with the two furthest points, and at each step add the point whose minimum distance to any of the previously selected points is maximum.

Suppose you are given the following points in two-dimensional Euclidean space: x = (0,0); y = (10,10), a = (1,6); b = (3,7); c = (4,3); d = (7,7), e = (8,2); f = (9,5). Obviously, x and y are furthest apart, so start with these. You must add five more points, which we shall refer to as the first, second,..., fifth points in what follows. The distance measure is the normal Euclidean L_2-norm. Which of the following is true about the order in which the five points are added?

Your Answer		Score	Explanation
f is added second			
b is added second	✔	1.00	
b is added first			
c is added fifth			
Total		1.00 / 1.00	

Question Explanation

It helps to construct the table of distances between each pair of points. Since we are only looking for minimum distances, rather than the exact distances, we shall tabulate the squares of the distances for convenience. The square of the distance between two points under the L_2 norm is just the sum of the squares of the differences of the components in each dimension. Here is the table of differences:

	a	b	c	d	e	f
x	37	58	25	98	68	106
y	97	58	85	18	68	26
a		5	18	37	65	65
b			17	16	50	40
c				25	17	29
d					26	8
e						10

It helps to keep track of the shortest distance between each unselected point and any selected point. When we select a point p, adjust the minimum distance from the remaining unselected points q, if p is the closest to q of any selected point. Initially, we have the following minimum distances:

a:37 b:58 c:25 d:18 e:68 f:26

These numbers are read off the table by taking the smaller of the distance to x and the distance to y. Thus, e, with the maximum minimum distance is the first selected. The table of minimum distances becomes:

a:37 b:50 c:17 d:18 f:10

For example, the minimum distance for b was lowered from 58 to 50, because e is closer to b, at distance sqrt{50}, than either x or y. Now, b has the maximum minimum distance, so it is selected second, and the table of minimum distances becomes:

a:5 c:17 d:16 f:10

Thus, c is third selected, and the minimum distances become a:5, d:16, f:10. Now, d is selected fourth, and the minimum distances are a:5, f:8. Thus f is selected fifth.