MAT9004 Assignment 3 Solutions

1. To optimise a neural network can require selecting the best combination of a large set of parameters. For a simple case, suppose that a neural network has just two parameters x and y. The network is only feasible if $y \geq -1$, $x \leq 3$ and $y \leq x$. An analyst establishes that the performance function of the network is

$$f(x, y) = xy^2 + x^2 - 4xy.$$

(a) Find $\nabla f(x, y)$.

Solution: $\begin{pmatrix} y^2 + 2x - 4y \\ 2xy - 4x \end{pmatrix}$.

(b) Find the Hessian matrix $H(x, y)$ for f.

Solution: $H(x, y) = \begin{bmatrix} 2 & 2y - 4 \\ 2y - 4 & 2x \end{bmatrix}$.

(c) Locate and classify all stationary points of $f(x, y)$.

Solution: Setting the partials equal to zero,
$f_x = 0$ says $y^2 - 4y + 2x = 0$. (i)
$f_y = 0$ says $2xy - 4x = x(2y - 4) = 0$. (ii)
Now you can either solve (i) for x, which gives quadratic in y, and plug into (ii) (which gives a cubic equation in y), or solve (ii) in some way and plug into (i). We'll do the latter since it results only in quadratic equations.

For $x(2y - 4) = 0$, there are two cases.

Case 1. $x = 0$. Plug this into (i): $y^2 - 4y = 0$ or $y(y - 4) = 0$. This yields two stationary points: $(0, 0)$ and $(0, 4)$.

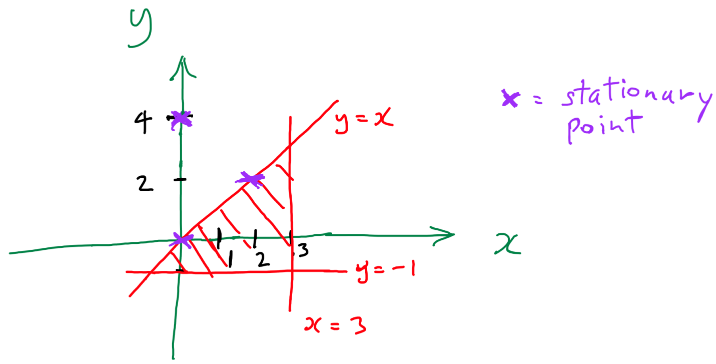
Case 2. $2y - 4 = 0$, or $y = 2$. Plug this into (i): $x = 2$. This yields only one stationary point: $(2, 2)$. To classify them we first compute the sign of $\det H(x, y)$ at each stationary point. Now $\det H(x, y) = 4x - (2y - 4)^2$, so

$$\det H(0, 0) = 0 - 16 < 0, \quad \det H(0, 4) = 0 - 16 < 0, \quad \det H(2, 2) = 8 - 0 > 0.$$

Thus $(0, 0)$ and $(0, 4)$ are saddle points, and $(2, 2)$ is an extremum whose nature is determined by $f_{xx} = 2 > 0$. So $(2, 2)$ is a local minimum.

(d) Draw a rough sketch showing the feasible region of parameters, along with the locations of the stationary points.

Solution:



(e) For what values of x and y is the maximum performance of the network achieved, and for what values is the minimum achieved?

Solution: We can first check the value at the local minimum: $f(2,2) = -4$. The saddle points cannot be an interior max or min, and anyway one of them is outside the feasible region of parameters.

There are no singular points, so we just have to compare -4 with the min on the boundary, and also check the max on the boundary. We notice that $(2,2)$ is actually ON the boundary, so our examination of the boundary will automatically take that into account. So if we completely ignore the stationary points, we will get the correct result in this case! To proceed methodically, we find the min and max along each of the three parts of the boundary by comparing the values at all stationary points and the endpoints of each part of the boundary, i.e. (see sketch) the three triangle vertices $(-1,-1)$, $(3,-1)$ and $(3,3)$. Since these three points each occur in two boundaries, we first find local extrema along each boundary interval and then compare with the triangle vertices at the end.

(i) $x = 3$. Along this boundary, $f(x,y) = f(3,y) = 3y^2 - 12y + 9$. The function $g_1(y) = 3y^2 - 12y + 9$ has $g_1'(x) = 6y - 12$ which is 0 when $y = 2$. At this point $g_1(y) = f(3,2) = -3$.

(ii) $y = -1$. Here $f(x,y) = f(x,-1) = x^2 + 5x$. The function $g_2(x) = x^2 + 5x$ has $g_2'(x) = 2x + 5$ which is 0 when $x = -5/2$. This is outside the boundary $(-1 \leq x \leq 3)$ so we can safely ignore it.

(iii) $y = x$. Here $f = g_3(x) = x^3 - 3x^2$, $g_3'(x) = 3x(x-2)$ and so there are stationary points at $x = 0$ and $x = 2$. Now $g_3(0) = 0$ and $g_3(2) = -4$ (which is actually the local min that we found earlier for f).

Finally, the values of f at the three triangle vertices are $f(3,3) = 0$, $f(3,-1) = 24$ and $f(-1,-1) = -4$.

The minimum value encountered, hence the global minimum of $f(x,y)$ on the feasible region, is -4, which occurred both at $(2,2)$ (local min as well as boundary point) and $(-1,-1)$. These are the points of minimum network performance. The maximum value found was $f(3,-1) = 24$. So the maximum network performance only occurs at $(3,-1)$.

[Question 1 total: **15** marks]

2. The planner in charge of raw materials operations at a well known beer-producing company needs to obtain 5 truckloads of fresh water. The five trucks available can go to any of 11 sources of water (mainly dams and reservoirs). The basic question is how many possible ways this can be achieved, using each truck once. The 11 sources all have different kinds of minerals in the water, so it is always important how many truckloads come from which source, but there are different conditions as follows. Find the number of ways in each case. You may leave your answer expressed in terms of binomials, factorials, integers raised to integer powers, or products and quotients of these, unless otherwise requested.(*Important: make sure you explain what formulae you use, and why.*

   (a) It does not matter which truck brings which kind of water, and any sources can be used by any number of trucks. Express your answer as a single integer.

   Solution: Here for each of $r = 5$ trucks, the planner selects a water source from $n = 11$ possible ones, and the choices are unordered because which truck brings water is irrelevant. The selections (sources) can be repeated, so this is unordered selection with repetition. The answer is

$$\binom{11+5-1}{5} = \binom{15}{5} = 3003.$$

   (b) It does not matter which truck brings which kind of water, but no source can be used by more than one truck. In this case, give the answer as a single integer.

   Solution: Again for each truck select a water source, the choices are again unordered, but the selections cannot be repeated, so the answer is

$$\binom{11}{5} = \frac{11 \cdot 10 \cdot 9 \cdot 8 \cdot 7}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 11 \cdot 3 \cdot 2 \cdot 7 = 462.$$

   (c) The company needs to keep track of how far each truck goes, so two arrangements are counted as different if they send different trucks to any given source. Still, no source can be used by more than one truck.

   Solution: Again for each truck select a water source, but the choices are now ordered since it's important which truck goes to each source. So ordered selection with no repetition. Hence the answer is

$$\frac{11!}{6!} \quad \text{or} \quad 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7.$$

   (d) The conditions in (c) hold, but additionally the company needs to record the order that the trucks return with the water. How many possible outcomes are there?

   Solution: For each of the possibilities in (c) there are now 5! possible orders in which the trucks return. So by the product rule,

$$\frac{11!}{6!} \times 5! = \frac{11!}{6}.$$

(e) The conditions in (a) hold, but additionally there is one of the sources that may only be used at most twice due to lack of supply. (The other sources are unrestricted). Express your answer as a single integer.

Solution:

Method 1. Count the ones using that particular source 0, 1 or 2 times separately.

Using it 0 times: select only from the other 10 sources.

$$\binom{10+5-1}{5} = \binom{14}{5} = 2002.$$

Using it 1 time: select 4 times from the other 10 sources.

$$\binom{10+4-1}{4} = \binom{13}{4} = 715.$$

Using it twice: select 3 times from the other 10 sources.

$$\binom{10+3-1}{3} = \binom{12}{3} = 220.$$

By the rule of sum, the number is the total of these: 2937.

Method 2. Subtract the ones using that particular source 3 or more times from the total in (a), using the complement rule.

We can count the ones using it at least 3 times by selecting the other two destinations for trucks, from all 11 sources (with repetition). That is,

$$\binom{11+2-1}{2} = \binom{12}{2} = 66.$$

So by the rule of complement, the total is $3003 - 66 = 2937$.

[Question 2 total: **11** marks]

3. A password is a sequence of letters (a–z) and digits (0–9). Find the number of passwords of length 10 under the constraints in (a), (b) or (c) (three separate problems). Express your answer using factorials and integers, products and ratios of them, and/or sums of such things.

(a) There are 3 letters and 7 digits, and at most one '9'.

Solution: With no '9': first select (unordered, no repetitions) the positions for the 3 letters from the ten positions in $\binom{10}{3}$ ways. Then select (ordered, with repetitions) which of the 26 letters to put in each of the three positions in 26^3 ways. Finally (ordered, with repetitions) which of the 7 digits from 0–8 to place in the remaining positions, in 9^6 ways. Since the number of choices of each action does not depend on how we did the previous actions, the rule of product says the number of ways to perform all three actions is $\binom{10}{3}26^3 9^7$.

With one '9': argue as above, but now when we come to filling in the digits, first choose where the '9' goes (in $\binom{7}{1} = 7$ ways) and then make an ordered choice for the remaining 6 digits in 9^6 ways. Thus for this case we get $\binom{10}{3}26^3 \times 7 \times 9^6$.

By the rule of sum, since these two cases are mututally exclusive, the answer is the sum of the two above quantities, i.e.

$$\binom{10}{3}26^3(9^7 + 7 \times 9^6) = \frac{10 \times 9 \times 8}{3 \times 2 \times 1}26^3 9^6 \times 16.$$

The answer can also be written as $120 \times 26^3 9^6 \times 16$.

(b) There are 6 letters and 4 digits, and no digit occurs twice.

Solution: Here the method is similar to (a), with $\binom{10}{4}$ ways to choose the positions of the digits. However, when we come to filling in the digits, since no digit occurs twice, the ordered selection of 4 digits from 10 is with no repetitions. The number of such selections is $10!/6!$. Hence the answer here is

$$\binom{10}{4}26^6 \times 10!/6! = \frac{10!}{6!4!}26^6 \cdot 10 \cdot 9 \cdot 8 \cdot 7.$$

Alternate answers $26^6(10 \cdot 9 \cdot 8 \cdot 7)^2/24$ or similar, since $4! = 24$.

(c) No letters are used BUT the first four digits are not all the same AND the last four digits are not all the same, AND the 4th, 5th, 6th and 7th digits are not all the same.

Solution: Use the principle of inclusion-exclusion. Consider properties P_1 (first four digits the same), P_2 (last four digits the same), P_3 (middle four the same). Consider $S \subseteq \{1, 2, 3\}$. Let A_S denote the set of passwords satisfying all properties P_i for $i \in S$.

$|U| = 10^{10}$ where U is the set of all passwords using integers only.

$|A_1| = 10^7 = |A_2| = |A_3|$.

$|A_1 \cap A_2| = 10^4$ (since we can choose the digit for the first four, then for the last four, then the two in the middle, using the rule of product); $|A_1 \cap A_3| = |A_2 \cap A_3| = 10^4$ (similar calculation).

$|A_1 \cap A_2 \cap A_3| = 10$ (easy).

So by the principle of inclusion-exclusion, the answer is

$$10^{10} - 3 \cdot 10^7 + 3 \cdot 10^4 - 10 = 9,970,029,990.$$

Alternate solutions There are several alternate solutions. Here is one.

We can first choose the middle digits, then after choosing them choose the first three, and then the last three. In choosing the middle four there are 10^4 choices, except for the 10 choices where one digit is repeated. Hence 9990 possibilities for them.

The next action is to choose the first three digits. We just have to make sure they are not all the same as the already-chosen fourth digit. By a similar argument there are 999 ways to do this.

Finally choose the last three digits. Again, just avoid choosing three that are all the same as the seventh digit — in 999 ways. So by the rule of product the answer is 9990×999^2.

[Question 3 total: **10** marks]

4. A fair coin is flipped six times. The outcomes of the coin flips form a palindrome if the sequence of T's and H's reads the same forwards and backwards, e.g. THTTHT. Let A denote the event that the first, second and fourth flips are all 'T'. Let Z denote the event that the six flips form a palindrome.

   (a) Is A independent of Z?

   Solution: For these repeated trials we always assume the outcomes of the trials are independent. Thus $\Pr(A)$ is the product of the probabilities that the first, second and fourth flips are each 'T', i.e. $(1/2)^3$.

   To find $\Pr(Z)$ we note that once the first three flips are given, the 4th–6th must copy them in reverse order. The probability of this copying is $1/2$ for each flip, and since they are independent, the probability is $(1/2)^3$ once again.

   To find $\Pr(A \cap Z)$, note that for a palindrome the 6th, 5th and 3rd flips must equal the 1st, 2nd and 4th respectively. So the only palindrome satisfying A is TTTTTT, and $\Pr(TTTTTT) = (1/2)^6$. So $\Pr(A \cap Z) = (1/2)^6$.

   Since $\Pr(A \cap Z) = \Pr(A)\Pr(Z)$, A and Z are independent.

   (b) Is $\overline{A}$ independent of Z?

   Solution: Yes.To justify this, you can either do a computation like the above, but with $\overline{A}$ instead of A, or quote the result from Lecture 19 that tells us that A and Z being independent is equivalent to $\overline{A}$ and Z being independent. So the result follows from (a).

   (c) A fair coin flipped six times and a certain property, Q, is being studied. Let Z be the event that the first three flips are all 'heads'. It is found that $\Pr(Q \mid Z) = 1/4$ and $\Pr(Q \mid \overline{Z}) = 2/7$. Show how to use *Bayes' Theorem* to find $\Pr(Z \mid Q)$.

   Solution: Since it is fair, $\Pr(Z) = 1/8$ for reasons as in (a). Thus $\Pr(\overline{Z}) = 7/8$. Bayes' theorem tells us that

$$\Pr(Z \mid Q) = \frac{\Pr(Q \mid Z)\Pr(Z)}{\Pr(Q \mid Z)\Pr(Z) + \Pr(Q \mid \overline{Z})\Pr(\overline{Z})} = \frac{(1/4)(1/8)}{(1/4)(1/8) + (2/7)(7/8)}$$

   Answer: 1/9.

Reminder: as with all assignment questions, you must provide justification for your answers.

[Question 4 total: **9** marks]

5. An integer N is chosen from 1 to 10 uniformly at random.

Two random variables are defined:

X is 1 plus the remainder on division of N by 3. So e.g. when $N = 5$, the remainder on division by 3 is 2, so $X = 3$.

Y is $\lceil N/3 \rceil$. So e.g. when $N = 5$, $Y = 2$.

(a) Find $\mathrm{E}[X]$, $\mathrm{E}[Y]$ and $\mathrm{Var}[Y]$.

Solution:

a	1	2	3	4
$\Pr(X = a)$	.3	.4	.3	0
$\Pr(Y = a)$	.3	.3	.3	.1

$\mathrm{E}[X] = 1 \times .3 + 2 \times .4 + 3 \times .3 + 4 \times 0 = 2.$

$\mathrm{E}[Y] = 1 \times .3 + 2 \times .3 + 3 \times .3 + 4 \times .1 = 2.2.$

Let's use $\mathrm{Var}[Y] = \mathrm{E}[Y^2] - (\mathrm{E}[Y])^2$ since that is simpler than working out $\mathrm{E}[(Y - \mathrm{E}[Y])^2]$.

$\mathrm{E}[Y^2] = 1 \times .3 + 4 \times .3 + 9 \times .3 + 16 \times .1 = 5.8.$

So $\mathrm{Var}[Y] = 5.8 - 2.2^2 = 0.96.$

(b) Are X and Y independent?

Solution: No since e.g. $\Pr(X = 1 \text{ and } Y = 4) = 0$ but $\Pr(X = 1)(Y = 4) = .03 \neq 0$.(Many other examples exist.)

[Question 5 total: **7** marks]