STAT542 Homework 6

Yu-Ching Liao

April 20, 2023

1 Problem 1

1.1 Q1

To compute the integer moments of Z, we first need to find the expectation $E[Z^n]$ for any non-negative integer n. Recall that $Z = e^{-F}$, where F is a standard normal random variable with mean 0 and variance 1.

The n-th moment of Z can be written as:

$$E[Z^n] = E[e^{-nF}]$$

Now we can use the moment-generating function (MGF) of the standard normal distribution to compute the expectation. The MGF of a standard normal random variable X is given by:

$$MGF_X(t) = E[e^{tX}] = e^{t^2/2}$$

We can rewrite the expectation $E[e^{-nF}]$ as:

$$E[e^{-nF}] = MGF_F(-n)$$

Plugging in the MGF of the standard normal distribution:

$$E[e^{-nF}] = e^{(-n)^2/2} = e^{n^2/2}$$

So, the n-th moment of Z is given by $E[Z^n] = e^{n^2/2}$ for any non-negative integer n.

1.2 Q2

To compute the integer moments of Z_θ, we need to find the expectation $E[Z_\theta^n]$ for any non-negative integer n, where $Z_\theta = e^{-F}$ and F follows the distribution with the given density $p_\theta(F)$.

The n-th moment of Z_θ can be written as:

$$E[Z_\theta^n] = E[e^{-nF_\theta}]$$

We can compute this expectation using the given density function $p_\theta(F)$:

$$E[Z_\theta^n] = \int e^{-nF} \times p_\theta(F)\, dF$$

Now, substitute the density function $p_\theta(F) = p(F)[1 + \theta sin(2\pi F)]$, where $p(F)$ is the standard normal density:

$$E[Z_\theta^n] = \int e^{-nF} \times p(F)[1 + \theta sin(2\pi F)]\, dF$$

This integral can be split into two parts:

$$E[Z_\theta^n] = \int e^{-nF} \times p(F) \times 1\, dF + \theta \int e^{-nF} \times p(F) \times sin(2\pi F)\, dF$$

The first part of the integral is just the n-th moment of Z (where F follows the standard normal distribution):

$$E[Z^n] = e^{n^2/2}$$

The second part of the integral is the expectation of e^{-nF} multiplied by $sin(2\pi F)$ with respect to the standard normal distribution. This expectation is 0 because $sin(2\pi F)$ is an odd function and the standard normal distribution is symmetric around 0. Mathematically, this can be shown as:

$$\int e^{-nF} \times p(F) \times sin(2\pi F)\, dF = 0$$

So, the n-th moment of Z_θ is given by:

$$E[Z_\theta^n] = e^{n^2/2} + \theta \times 0 = e^{n^2/2}$$

This result shows that the integer moments of Z_θ are the same as the integer moments of Z, which are $e^{n^2/2}$ for any non-negative integer n.

1.3 Q3

From our previous calculations, we found that the integer moments of Z_θ are given by $E[Z_\theta^n] = e^{n^2/2}$, for any non-negative integer n. We can see that these moments do not depend on θ. This means that the integer moments of Z_θ are the same for any value of θ, and therefore, they do not contain any information about the parameter θ.

Since the empirical moments are estimates of the true moments, they will also not depend on θ. This implies that we cannot estimate θ using the empirical moments of samples from $Z_\theta = e^{-F_\theta}$, where $F_\theta \sim p_\theta$. The reason for this is that the moments do not carry any information about θ, making it impossible to estimate θ from the empirical moments alone.

2 Problem 2

Given the hint, we will use the *Dirichlet Distribution* as the conjugate prior for the *Multinomial Distribution*. The prior distribution is uniform on the probability simplex, which corresponds to a *Dirichlet Distribution* with all parameters equal to 1. That is, the prior is $Dir(1, 1, 1)$.

Now, let's incorporate the observed data. The counts of the words 'machines', 'have', and 'conscious' are 25, 25, and 50 respectively. To find the posterior distribution, we will update the parameters of the *Dirichlet* prior by adding the observed counts to the corresponding parameters:

$$Posterior = Dir(1 + 25, 1 + 25, 1 + 50) = Dir(26, 26, 51)$$

So, the posterior distribution of [$p1$, $p2$, $p3$] is a *Dirichlet Distribution* with parameters 26, 26, and 51.

3 Problem 3

3.1 Q1

To prove that the rank of the given matrix M is at most 3, we need to express its elements M_{ij} as the sum of products of functions $f_k(i)$ and $g_k(j)$ for $k = 1, 2, 3$. Let's start by examining the given element $(i - j)^2$:

$$M_{ij} = (i - j)^2 = i^2 - 2ij + j^2$$

Now, let's try to find suitable functions f_k andg_k for $k = 1, 2, 3$:

$$f_1(i) = i, g_1(j) = -2j$$

$$f_2(i) = i^2, g_2(j) = 1$$

$$f_3(i) = 1, g_3(j) = j^2$$

Now we can express M_{ij} as the sum of products of these functions:

$$\begin{aligned} M_{ij} &= f_1(i)g_1(j) + f_2(i)g_2(j) + f_3(i)g_3(j) \\ &= i(-2j) + (i^2)1 + 1(j^2) \\ &= i^2 - 2ij + j^2 \end{aligned}$$

Since we can express M_{ij} as a sum of products of functions f_k and g_k for $k = 1, 2, 3$, the rank of M is at most 3. This shows that the *nonnegative* rank of bounded-rank matrices may be unbounded.

3.2 Q2

To show that $\text{rank}(M) \geq 3$ for $n \geq 3$, we need to find a 3×3 submatrix of M with rank 3. Let's construct a 3×3 submatrix M' by selecting three distinct rows and columns of M. For simplicity, let's choose the first three rows and columns, i.e., $i = 1, 2, 3$ and $j = 1, 2, 3$.

The submatrix M' is:

$$M' = \begin{bmatrix} (1-1)^2 & (1-2)^2 & (1-3)^2 \\ (2-1)^2 & (2-2)^2 & (2-3)^2 \\ (3-1)^2 & (3-2)^2 & (3-3)^2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 4 \\ 1 & 0 & 1 \\ 4 & 1 & 0 \end{bmatrix}$$

Now, let's compute the determinant of M':

$$\det(M') = 0 \cdot (0 \cdot 0 - 1 \cdot 1) - 1 \cdot (1 \cdot 0 - 1 \cdot 4) + 4 \cdot (1 \cdot 1 - 0 \cdot 4) = 0 + 4 + 4 = 8.$$

Since the determinant of M' is non-zero, it implies that M' has full rank, which is 3 in this case. Therefore, the rank of M is at least 3, as required. So, for $n \geq 3$, $\text{rank}(M) \geq 3$.

3.3 Q3

Consider the nonnegative rank $\text{rank}^+(M) \geq \log_2 n$. To show this, we can focus on the locations of zeros in the matrix M. Suppose that $M = AB^\top$ for some entrywise nonnegative matrices $A, B \in R^{n \times r}$.

We need to analyze whether two rows of A can have the same zero pattern. Let A_i and A_j be two distinct rows of A. Suppose that they have the same zero pattern, i.e., $A_{i,k} = 0$ if and only if $A_{j,k} = 0$ for $k = 1, \ldots, r$. Then, for any column k of B, $B_{k,i} = 0$ if and only if $B_{k,j} = 0$. This would imply that the rows M_i and M_j of M are equal:

$$M_{i,p} = \sum_{k=1}^{r} A_{i,k} B_{k,p} = \sum_{k=1}^{r} A_{j,k} B_{k,p} = M_{j,p}$$

for all $p = 1, \ldots, n$. However, we know that the diagonal elements of M are all zeros, while the off-diagonal elements are positive. This contradicts the assumption that M_i and M_j are equal, since $M_{i,i} = 0$ and $M_{j,j} = 0$ while $M_{i,j} > 0$ and $M_{j,i} > 0$.

Therefore, no two rows of A can have the same zero pattern. Since there are n rows in A, there must be at least $\log_2 n$ different zero patterns (binary encoding). This implies that $r \geq \log_2 n$, and thus the nonnegative rank $\text{rank}^+(M) \geq \log_2 n$.