Problem 1

Given ground set , $X = \{x_1, x_2, x_3, \cdots, x_n\}$ and $A \subset X$.
Sol(a) $f_{max}(A) = max_{x_i \in A} w_i$.
Let's assume two subsets A and B such that $A \subset B \subset X$. We can choose an element $x \in X$ such that $x \in X \setminus B$. We can define,

$$\Delta(x|A) = f(A \cup \{x\}) - f(A) \tag{1}$$

$$= \begin{cases} 0, & \text{if } f_{max}(A) \geq w_x \\ w_x + f_{max}(A), & \text{otherwise} \end{cases} \tag{2}$$

where w_x is the weight of the element x.
As can be observed, for all possible cases for $A \subset B \subset X$,

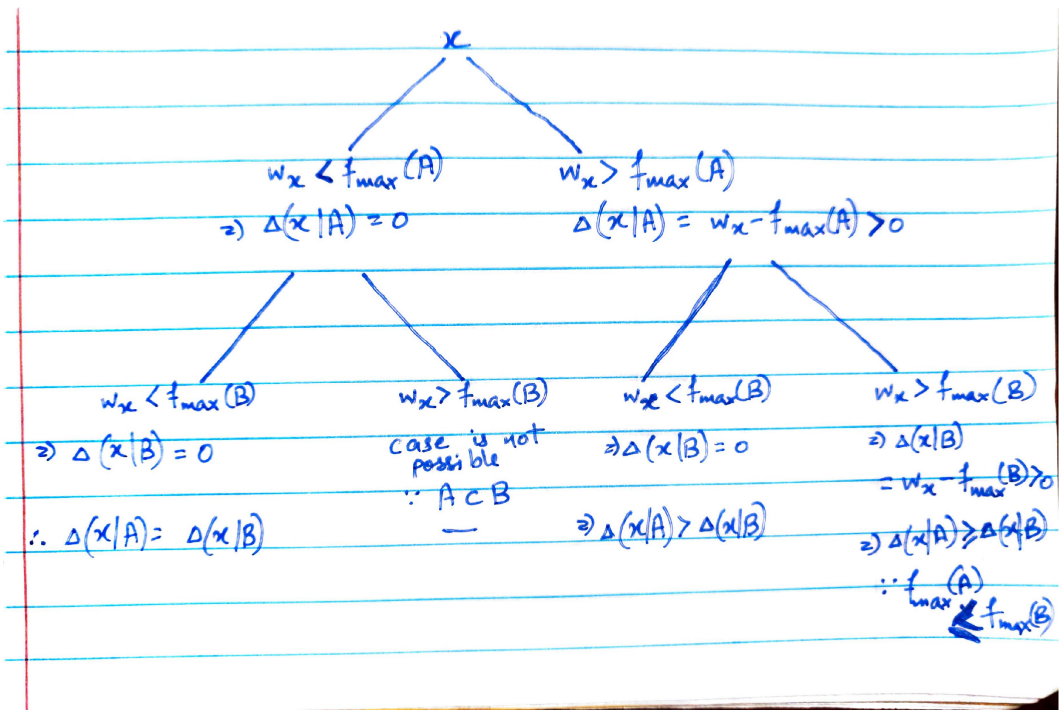

Figure 1:

$$f(A \cup \{x\}) - f(A) \geq f(B \cup \{x\}) - f(B) \tag{3}$$

This proves that the function $f_{max}(A)$ is submodular.

Sol(b) Let A ⊆ B ⊆ X. Given f_{min}, a submodular function, we can say that

$$f(A) \geq f(B) \tag{4}$$

$$\implies -f(A) \leq -f(B) \tag{5}$$

Therefore, for an $x \in X \backslash B$ we can assume that $w_x < f(B) < f(A)$ (weight of x is less than $f_{min}(B)$ which is lesser than $f_{min}(B)$), using 4

Problem 1 continued on next page. . .

$$f(A \cup \{x\}) - f(A) = w_x - f(A) < 0 \tag{6}$$
$$f(B \cup \{x\}) - f(B) = w_x - f(B) < 0 \tag{7}$$
$$\implies f(A \cup \{x\}) - f(A) \leq f(B \cup \{x\}) - f(B) \tag{8}$$
$$\tag{9}$$

This example contradicts the definition of sub-modularity and thus, the function $f_{min}(A)$ is not submodular.

Problem 2

Let's assume a very simple case of two vertices and one edge. S and $\bar{S}$ are two disjoint sets containing vertices such that $S \cup \bar{S} = V$. Given below are the different possible cases of S and $\bar{S}$ For the different

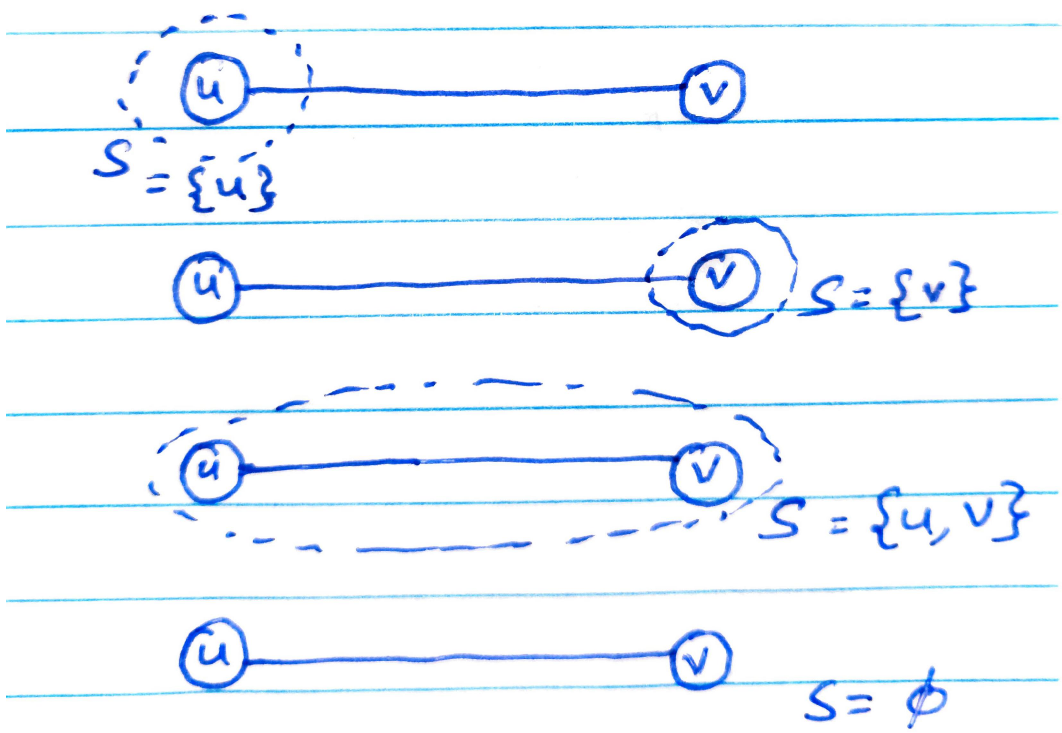

Figure 2: Possible sets S

possible cases the values of the function $c(S)$ and $c(\bar{S})$ can be written as:

1. $c(S) = 1$ and $c(\bar{S}) = 1$. Also, $c(S \cup \bar{S}) = 0$ and $c(S \cap \bar{S}) = 0$
2. $c(S) = 1$ and $c(\bar{S}) = 1$. Also, $c(S \cup \bar{S}) = 0$ and $c(S \cap \bar{S}) = 0$
3. $c(S) = 0$ and $c(\bar{S}) = 0$. Also, $c(S \cup \bar{S}) = 0$ and $c(S \cap \bar{S}) = 0$
4. $c(S) = 0$ and $c(\bar{S}) = 0$. Also, $c(S \cup \bar{S}) = 0$ and $c(S \cap \bar{S}) = 0$

Problem 2 continued on next page. . .

From all these cases it can be observed that for two sets S and T, $c(S)$ satisfies

$$c(S) + c(T) \geq c(S \cup T) + c(S \cap T) \tag{10}$$

Therefore, we can say that the cut of one edge is submodular. We know that a larger graph is a positive linear combination of edges. Also positive linear combination of submodular functions is also submodular. Using these results we can say that $c(S)$ is a submodular function for any set $S \subset V$.

Problem 3

Sol

Let the quality of each paper reviewed by a reviewerr_j be defined as $q_i(r_j)$. The total quality of the reviews then becomes

$$Q(R) = \sum_{\forall i} max_j q_i(r_j) \tag{11}$$

where R is the set of reviewers chosen such that $|R| < k$.

Problem 4

Sol(a)

The given problem is modelled as a Gaussian Process. The training input and output data is used fit a mapping from the input to the output data which is then used to predict the output mean and standard deviation values for the test input data. The kernel function used for the gaussian process is a mean-squared kernel with added white noise i.e.,

$$k(x, x*) = \sigma_f^2 (e^{-\frac{(x-x*)^2}{2l^2}}) + \sigma_n^2 \delta_{x,x*}$$

The parameters σ_f, length-scale l and white noise covariance σ_n were set to {1, 1, 1} initially. The optimal values learned for these parameters during optimization were {0.701, 0.448, 0.00858} respectively. The output plot is shown in the figure(3)
The mean-squared error computed for the predicted mean values vs the ground truth is : 0.2365

I also tried using the 'Exponential Sine Squared Kernel' which is usually used to model periodic functions. It is defined as

$$k(x_i, x_j) = exp(-2(sin(pi/periodicity * d(x_i, x_j))/length_scale)^2) \tag{12}$$

The results after using this kernel are shown below:
The mean-squared error computed for the predicted mean values vs the ground truth is : 0.0355
This shows that **Sol(b)**

The given problem is also modelled as a Gaussian Process. The training input(4-dimensional) and output data(1-dimensional) is used fit a mapping from the input to the output data which is then used to predict the output mean and standard deviation values for the test input data. The kernel function used for the gaussian process is a mean-squared kernel with added white noise i.e.,

$$k(x, x*) = \sigma_f^2 (e^{-\frac{(x-x*)^2}{2l^2}}) + \sigma_n^2 \delta_{x,x*}$$

The parameters σ_f, length-scale l and white noise covariance σ_n were set to {1, 1, 1} initially. The training was carried out for all data samples and the kernel values learned were respectively were {294, 110, 16.9}.

Problem 4 continued on next page...

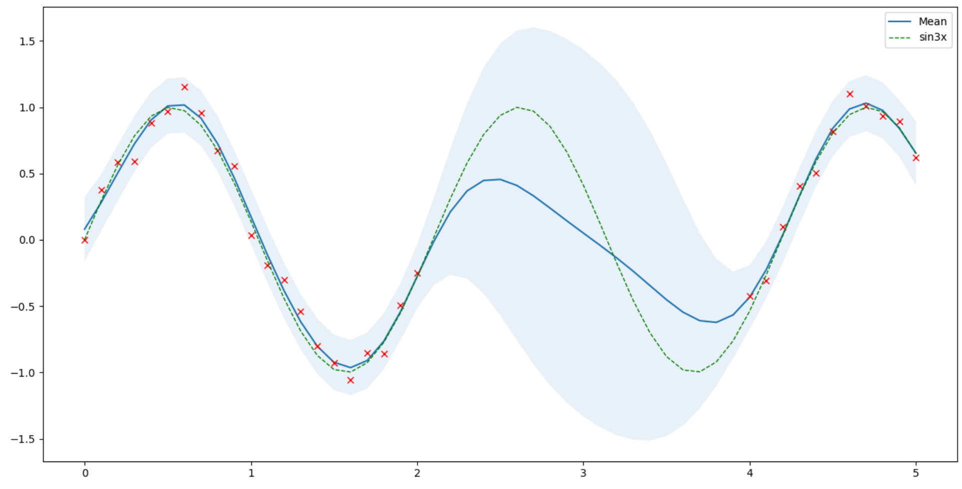

Figure 3: Mean output function predicted vs the ground truth function 'sin3x' with a 100% confidence interval

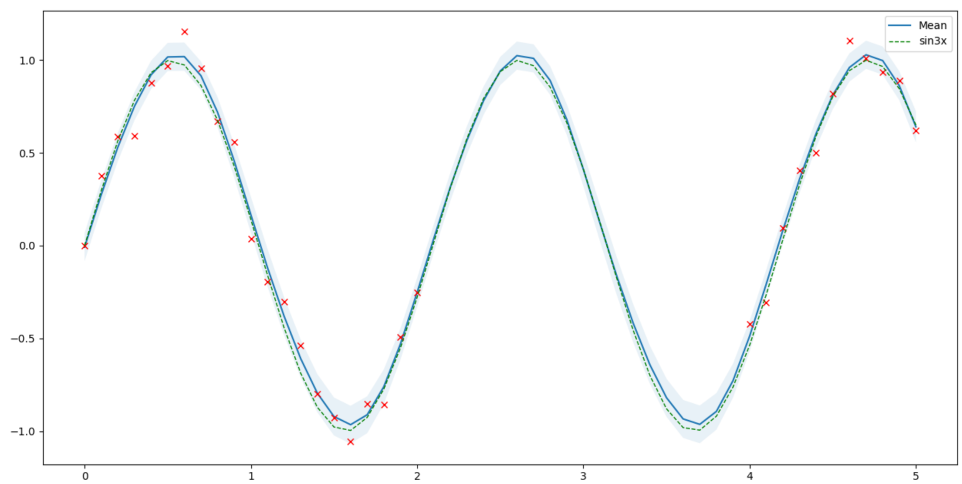

Figure 4: Mean output function predicted vs the ground truth function 'sin3x' with a 100% confidence interval

As a sanity check I also plotted the predicted mean values against the ground truth values to check if the output plot is close to $y = x$ line. The output plot is shown in the figure(5)
The mean-squared error computed for the predicted mean values vs the ground truth is : 4.042

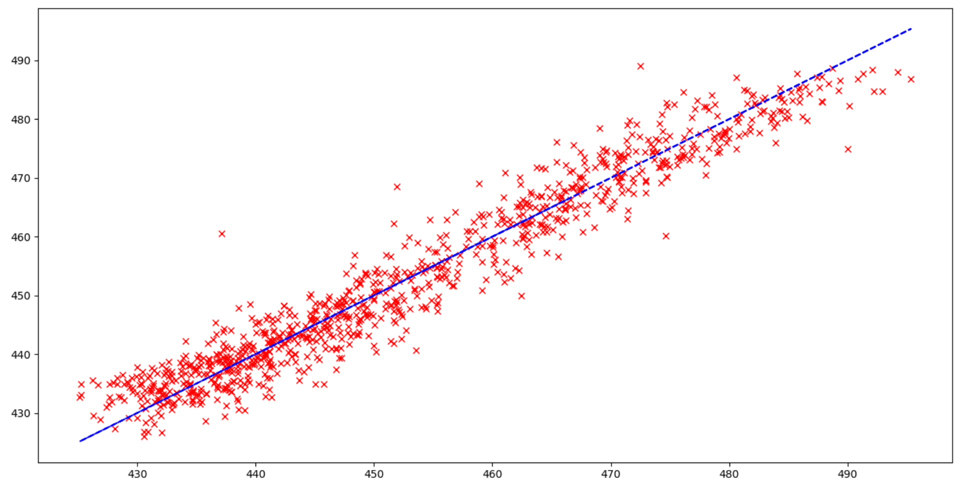

Figure 5: The X-axis is the ground truth values and the blue line is the y=ground_truth and the red(x's) are predicted mean values.