SUB-MODULAR OPTIMIZATION

Advanced Topics in Algorithms 236621

Rotem Green 318459484
Hay Elmaliah 315777433

Table of Contents

Question 1

Prove or refute that the following set functions $f : 2^N \to \mathbb{R}_+$ are submodular.

Section 1.a

For every $i = 1, \dots, k$ let $g_i : 2^N \to \mathbb{R}_+$ be k submodular functions over N, and let $a_1, \dots, a_k$ be k non-negative coefficients. For every $S \subseteq N$ define $f(S) = a_1 g_1(S) + \dots + a_k g_k(S)$.

נראה כי f אכן תת-מודולרית.

יהיו $A \subseteq B \subseteq N$ ויהי $u \in N \backslash B$.

נשים לב כי כיוון שלכל $i = 1, \dots, k$,g_i תת-מודולרית ו- $a_i \geq 0$, מתקיים:

$$\forall 1 \leq i \leq k,: \quad a_i\big(g_i(A \cup \{u\}) - g_i(A)\big) \geq a_i\big(g_i(B \cup \{u\}) - g_i(B)\big)$$

לכן מהגדרת f נקבל:

$$\begin{aligned}
&f(A \cup \{u\}) - f(A) \\
&= a_1 g_1(A \cup \{u\}) + \dots + a_k g_k(A \cup \{u\}) - a_1 g_1(A) - \dots - a_k g_k(A) \\
&= a_1\big(g_1(A \cup \{u\}) - g_1(A)\big) + \dots + a_k\big(g_k(A \cup \{u\}) - g_k(A)\big)
\end{aligned}$$

מטענת העזר למעלה, נקבל כי:

$$\begin{aligned}
&f(A \cup \{u\}) - f(A) \\
&\geq a_1\big(g_1(B \cup \{u\}) - g_1(B)\big) + \dots + a_k\big(g_k(B \cup \{u\}) - g_k(B)\big) \\
&= a_1 g_1(B \cup \{u\}) + \dots + a_k g_k(B \cup \{u\}) - a_1 g_1(B) - \dots - a_k g_k(B) \\
&= f(B \cup \{u\}) - f(B)
\end{aligned}$$

כלומר קיבלנו:

$$f(A \cup \{u\}) - f(A) \geq f(B \cup \{u\}) - f(B)$$

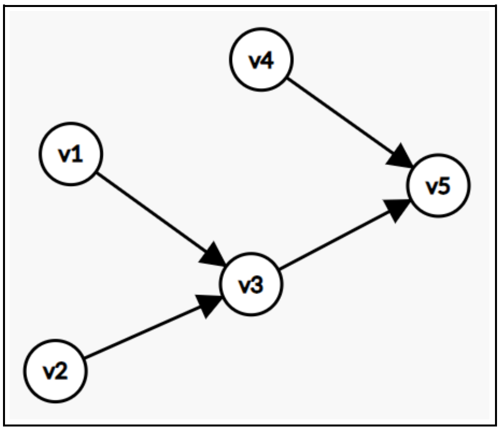
ולכן f תת-מודולרית ■

Section 1.b

Let $G = (V, E)$ be a directed graph, $N = V$, and $g(S)$ is the number of directed edges $(u \to v) \in E$ such that $u \in S$ and $v \notin S$. Additionally, let $b \in \mathbb{N}$ be a given threshold. For every $S \subseteq V$ define $f(S) = \min\{g(S), b\}$.

נראה כי f אינה תת-מודולרית.

נגדיר גרף באופן הבא:



נבחר:

- $A = \{v_1, v_2\}$
- $B = \{v_1, v_2, v_3\}$
- $b = 2$

מתקיים כי $A \subset B$ ונסתכל על הצומת $v_4 \in V \backslash B$:

$$\begin{gathered} f(A \cup \{v_4\}) - f(A) \\ = min\{g(\{v_1, v_2, v_4\}), 2\} - min\{g(\{v_1, v_2\}), 2\} \\ = min\{3, 2\} - min\{2, 2\} \\ = 2 - 2 = 0 \end{gathered}$$

$$\begin{gathered} f(B \cup \{v_4\}) - f(B) \\ = min\{g(\{v_1, v_2, v_3, v_4\}), 2\} - min\{g(\{v_1, v_2, v_3\}), 2\} \\ = min\{2, 2\} - min\{1, 2\} = 2 - 1 = 1 \end{gathered}$$

קיבלנו:

$$\underbrace{f(A \cup \{v_4\}) - f(A)}_{=0} < \underbrace{f(B \cup \{v_4\}) - f(B)}_{=1}$$

עבור הקבוצות $A \subset B$,

לכן עפ"י הגדרה f אינה תת-מודולרית ∎

Section 1.c

> Let $g : 2^U \to \mathbb{R}_+$ be s submodular function over U. Choose the ground set N as follows $N = \{u : u \in U\} \cup \{u' : u \in U\}$, *i.e.*, N is the disjoint union of two independent copies of U. For every $S \subseteq N$ denote by $S_U \subseteq U$ the following set $\{u \in U : u \in S \text{ or } u' \in S\}$. Define $f(S) = g(S_U)$ for every $S \subseteq N$.

נראה כי f אינה תת-מודולרית.

נגדיר:

- $U = \{1,2\}$
- $g(S) = \begin{cases} 1 \; if \; 1 \notin S \\ 0 \; if \; 1 \; \in S \end{cases}$

ראשית נראה כי g תת-מודולית:

יהיו $A \subseteq B \subseteq U$ ויהי $u \in U \backslash B$.

נפריד למקרים עבור u:

- $u = 1$

  נקבל כי $g(A) = g(B) = 1$ וכן $g(A \cup \{1\}) = g(B \cup \{1\}) = 0$

  $g(A \cup \{1\}) - g(A) = g(B \cup \{1\}) - g(B) \Leftarrow$
- $u = 2$

  נקבל כי $g(A \cup \{2\}) = g(A)$ וכן $g(B \cup \{2\}) = g(B)$

  $g(A \cup \{2\}) - g(A) = g(B \cup \{2\}) - g(B) \Leftarrow$

כלומר בכל מקרה נקבל:

$$g(A \cup \{u\}) - g(A) = g(B \cup \{u\}) - g(B)$$

$$g(A \cup \{u\}) - g(A) \geq g(B \cup \{u\}) - g(B) \Leftarrow$$

ולכן g תת-מודולרית.

כעת נראה כי f אינה תת-מודלרית:

נבחר:

- $A = \{2\}$
- $B = \{1', 2\}$
- $1 \in N \backslash B$

נשים לב כי אכן מתקיים $A \subseteq B \subseteq N$.

לבסוף:

$$f(A \cup \{1\}) - f(A) = g((A \cup \{1\})_U) - g(A_U) = g(\{2,1\}) - g(\{2\}) = \; 0 - 1 = -1$$

$$f(B \cup \{1\}) - f(B) = g((B \cup \{1\})_U) - g(B_U) = g(\{1,2\}) - g(\{1,2\}) = 1 - 1 = 0$$

$$\underbrace{f(A \cup \{1\}) - f(A)}_{=(-1)} < \underbrace{f(B \cup \{1\}) - f(B)}_{=0} \Leftarrow$$

כלומר f אינה תת-מודלרית ∎

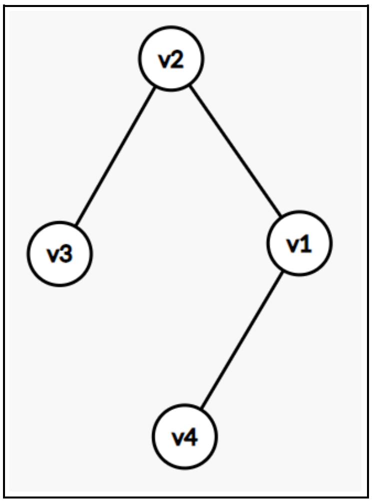
Section 1.d

Let $g : 2^N \to \mathbb{R}_+$ be a submodular function over N and let $T \subseteq N$. Define $f(S) = g((T \setminus S) \cup (S \setminus T))$ for every $S \subseteq N$, *i.e.*, $f(S)$ equals g on the reflection of S with respect to T (all elements of S that are in T are swapped and elements of S outside of T are not swapped).

נראה כי f אינה תת-מודלרית.

נבחר את g להיות פונקצית חתך בגרף לא מכוון (ראינו בכיתה כי פונקציית חתך הינה תת-מודלרית).

נגדיר את הגרף:



נגדיר:

- $A = \{v_1\}$
- $B = \{v_1, v_2\}$
- $T = \{v_1, v_2, v_3$,

מתקיים כי $A \subseteq B$, ומתקיימים הבאים:

1. $(T \setminus A) \cup (A \setminus T) = (\{v_1, v_2, v_3\} \setminus \{v_1\}) \cup (\{v_1\} \setminus \{v_1, v_2, v_3\}) = \{v_2, v_3\}$
2. $(T \setminus (A \cup \{v_4\})) \cup \big((A \cup \{v_4\}) \setminus T\big) = (\{v_1, v_2, v_3\} \setminus \{v_1, v_4\}) \cup (\{v_1, v_4\} \setminus \{v_1, v_2, v_3\}) = \{v_2, v_3, v_4\}$
3. $(T \setminus B) \cup (B \setminus T) = (\{v_1, v_2, v_3\} \setminus \{v_1, v_2\}) \cup (\{v_1, v_2\} \setminus \{v_1, v_2, v_3\}) = \{v_3\}$
4. $(T \setminus (B \cup \{v_4\})) \cup \big((B \cup \{v_4\}) \setminus T\big) = (\{v_1, v_2, v_3\} \setminus \{v_1, v_2, v_4\}) \cup (\{v_1, v_2, v_4\} \setminus \{v_1, v_2, v_3\}) = \{v_3, v_4\}$

לכן:

$$f(A \cup \{v_4\}) - f(A) = g(\{v_1, v_2, v_3\}) - g(\{v_2, v_3\}) = 1 - 1 = 0$$

$$f(B \cup \{v_4\}) - f(B) = g(\{v_3, v_4\}) - g(\{v_3\}) = 2 - 1 = 1$$

$$\underbrace{f(A \cup \{v_4\}) - f(A)}_{=0} < \underbrace{f(B \cup \{v_4\}) - f(B)}_{=1} \Leftarrow$$

לכן f אינה תת-מודולרית ■

Section 1.e

Let $G=(V,E)$ be an undirected graph equipped with non-negative edge weights $b:E\to\mathbb{R}_+$. For every vertex $u\in V$ denote by $B(u)$ the total weight of edges touching u, *i.e.*, $B(u)=\sum_{e=(u,v)\in E}b_e$. Let $\{\theta_u\}_{u\in V}$ be independent identically distributed uniform random variables in $[0,1]$. Given an initial $A_0\subseteq V$ of infected vertices, consider the following dynamics which progresses in discrete steps: in the i^{th} step let A_i be,

$$A_i=A_{i-1}\cup\left\{u\notin A_{i-1}:\sum_{e=(u,v)\in E:v\in A_{i-1}}b_e\geq\theta_uB(u)\right\}.$$

Denote by i_{A_0} the step after which no new infected vertices are added, *i.e.*, $A_{i_{A_0}}=A_{i_{A_0}+1}=A_{i_{A_0}+2}=\dots$ (note that i_{A_0} is a random variable whose distribution depends on the initial infected set A_0). Set the ground set N to be V and for every $S\subseteq V$ define $f(S)=\mathbb{E}[|A_{i_S}|]$.

Hint: prove that f is submodular.

נראה כי f תת-מודלרית.

יהי $G=(V,E)$ גרף לא מכוון. נגדיר $G'=(V,E')$ גרף מכוון המתקבל מ-G ע"י הפיכת כל קשת ב-E לשתי קשתות אנטי-מקבילות ב-E'.

נגדיר כעת פונקציה $p:V\times V\to[0,1]$ באופן הבא:

$$p(u,v)=\begin{cases}\dfrac{b(u,v)}{B(u)} & if\ (u,v)\in E'\\ 0 & if\ (u,v)\notin E'\end{cases}$$

נגדיר גרף חדש $G''=(V,E')$ עבורו לכל צומת בגרף יש קשת יחידה נכנסת, כאשר הקשת (u,v) נבחרת לגרף בסיכוי $p(u,v)$.

בעת נגדיר תהליך איטרטיבי חדש, בתחילתו מוגרלת קבוצת קשתות S (לפי ההסתברויות שתוארו למעלה) וקבוצת צמתים B_0 ולאחר מכן בכל איטרציה i מתבצע:

$$B_i=B_{i-1}\cup\{v\ \notin B_i|\exists(u,v)\in S,u\in b_{i-1}\}$$

בדומה להגדרת השאלה, נסמן את j_S להיות האיטרציה האחרונה שבה התווסף צומת חדש בתהליך.

נגדיר: $h(S)=\mathbb{E}[|B_{j_S}|]$, $S\subseteq V$.

כעת נראה באינדוקציה כי לכל i, $\mathbb{E}[|B_i|]=\mathbb{E}[|A_i|]$ וכן כי $Pr[u\in B_i]=Pr\ [u\in A_i]$:

<u>בסיס</u>

עבור $i=0$ מתקיים $A_0=B_0$ ולכן $\mathbb{E}[|B_0|]=\mathbb{E}[|A_0|]$.

אם $v\in A_0=B_0$ אזי $Pr[u\in B_i]=Pr[u\in A_i]=1$,

אחרת $Pr[u\in B_i]=Pr[u\in A_i]=0$.

<u>צעד</u>

נניח כי הנחת האינדוקציה מתקיימת לכל $0 \le k \le i$ ונוכיח את הטענה עבור $i+1$:

נשים לב כי:

1. $Pr[u \in A_{i+1} | u \in A_i] = 1$
2. $Pr[u \in A_{i+1} | u \notin A_i]$ ניתן על ידי הפיתוח הנ"ל:

$$Pr[u \in A_{i+1} | u \notin A_i]$$

$$= \frac{Pr\left[\sum_{v \in A_{i-1},(u.v)\in E} b(u,v) < \theta_u B(u) \le \sum_{v \in A_i,(u.v)\in E} b(u,v)\right]}{Pr\left[\sum_{v \in A_{i-1},(u.v)\in E} b(u,v) < \theta_u B(u)\right]}$$

$$= \frac{Pr\left[\sum_{v \in A_{i-1},(u.v)\in E} \frac{b(u,v)}{B(u)} < \theta_u \le \sum_{v \in A_i,(u.v)\in E} \frac{b(u,v)}{B(u)}\right]}{Pr\left[\sum_{v \in A_{i-1},(u.v)\in E} \frac{b(u,v)}{B(u)} < \theta_u\right]}$$

$$= \frac{\sum_{v \in A_i,(u.v)\in E} \frac{b(u,v)}{B(u)} - \sum_{v \in A_{i-1},(u.v)\in E} \frac{b(u,v)}{B(u)}}{1 - \sum_{v \in A_{i-1},(u.v)\in E} \frac{b(u,v)}{B(u)}}$$

$$= \frac{\sum_{v \in A_i \backslash A_{i-1},(u.v)\in E} \frac{b(u,v)}{B(u)}}{1 - \sum_{v \in A_{i-1},(u.v)\in E} \frac{b(u,v)}{B(u)}}$$

לכן, מנוסחת הסתברות השלמה:

$$Pr[u \in A_{i+1}]$$
$$= Pr[u \in A_{i+1} | u \in A_i] \cdot Pr[u \in A_i] + Pr[u \in A_{i+1} | u \notin A_i] \cdot Pr[u \notin A_i]$$
$$= Pr[u \in A_i] + Pr[u \in A_{i+1} | u \notin A_i] \cdot Pr[u \notin A_i]$$

נבצע חישוב דומה עבור B_{i+1}:

1. $Pr[u \in B_{i+1} | u \in B_i] = 1$
2. $Pr[u \in B_{i+1} | u \notin B_i]$ ניתן על ידי הפיתוח הנ"ל:

$$Pr[u \in B_{i+1} | u \notin B_i] = \frac{Pr[u \in B_{i+1} \backslash B_i]}{Pr\ [u \notin B_i]} = \frac{\sum_{v \in B_i \backslash B_{i-1},(u.v)\in E'} \frac{b(u,v)}{B(u)}}{1 - \sum_{v \in B_{i-1},(u.v)\in E'} \frac{b(u,v)}{B(u)}}$$

לכן, מנוסחת הסתברות השלמה:

$$Pr[u \in B_{i+1}] = Pr[u \in B_{i+1} | u \in B_i] \cdot Pr[u \in B_i] + Pr[u \in B_{i+1} | u \notin B_i] \cdot Pr[u \notin B_i]$$
$$= Pr[u \in B_i] + Pr[u \in B_{i+1} | u \notin B_i] \cdot Pr[u \notin B_i]$$

מהגדרת E', לכל קשת לא מכוונת ב-E יש קשת מכוונת ב-E' ההמתאימה לה, לכן:

$$Pr[u \in B_{i+1} | u \notin B_i] = \ Pr[u \in A_{i+1} | u \notin A_i]$$

בנוסף, מהנחת האינדוקציה נקבל:

$$Pr[u \notin B_i] = Pr[u \notin A_i]\,, Pr[u \in B_i] = Pr[u \in A_i]$$

נקבל בסה"כ:

$$Pr[u \in A_{i+1}] = \ Pr[u \in B_{i+1}]$$

מהגדרת התוחלת נקבל $\mathbb{E}[|B_{i+1}|] = \mathbb{E}[|A_{i+1}|]$ ובכך הוכחנו את טענת האינדוקציה.

כעת נראה כי $f(S) = h(S)$.

תהי $S \subseteq V$. נסמן את i_S, j_S להיות האיטרציות האחרונות שבהן נוסף צומת חדש בתהליכים האיטרטיביים.

נגדיר $m = max\{i_S, j_S\}$.

מטענת האינדוקציה מתקיים $\mathbb{E}[|B_m|] = \mathbb{E}[|A_m|]$.

בנוסף, מהגדרת i_S, j_S מתקיים:

$$\mathbb{E}\left[\left|A_{i_S}\right|\right] = \mathbb{E}\left[\left|A_{i_S+1}\right|\right] = \cdots = \mathbb{E}[|A_m|]$$

$$\mathbb{E}\left[\left|B_{j_S}\right|\right] = \mathbb{E}\left[\left|B_{j_S+1}\right|\right] = \cdots = \mathbb{E}[|B_m|]$$

$$\Longrightarrow f(S) = \mathbb{E}\left[\left|A_{i_S}\right|\right] = \mathbb{E}[|A_m|] = \mathbb{E}[|B_m|] = \mathbb{E}\left[\left|B_{j_S}\right|\right] = h(S)$$

כעת נותר להראות כי h תת-מודלרית ובכך נקבל כי f תת-מודלרית:

נשים לב כי ניתן לרשום את h באופן הבא:

$$h(S) = \mathbb{E}\left[\left|B_{j_S}\right|\right] = \sum_{X \subseteq E'} Pr[X = Y] \cdot \mathbb{E}\left[\left|B_{j_S}\right| |\ X = Y\right] = \sum_{X \subseteq E'} Pr[X = Y] \cdot g_X(S)$$

כאשר $g_X(S)$ היא מספר הצמתים הישיגים מצמתי S במסלולים העוברים על קשתות ב-X בלבד, ו-Y הינו משתנה מקרי אשר מסמן את קבוצת הקשתות שהאלגוריתם הגריל.

ראינו בביתה כי $g_X(S)$ הינה תת-מודלרית ולכן מסעיף 1 גם h תת-מודלרית.

סה"כ הראינו כי f תת-מודלרית כנדרש ■

Question 2

Let $f : 2^N \to \mathbb{R}_+$ be a monotone submodular functions over a ground set N, and let $k \in \mathbb{N}$ be a cardinality bound. We are interested in solving: $\max\{f(S) : S \subseteq N, |S| \leq k\}$. The goal of this question is to obtain an algorithm that for every $\varepsilon > 0$ achieves an (almost) tight approximation of $1 - 1/e - \varepsilon$ in running time of $O(n \ln(1/\varepsilon))$. Consider the following sampling based algorithm parameterized by $\varepsilon > 0$:

- $S_0 \leftarrow \emptyset$.
- For $i = 1$ to k do:
  - Let $M_i \subseteq N$ be a uniformly random subset of size $\lceil \frac{n \ln(1/\varepsilon)}{k} \rceil$.
  - $u_i \leftarrow \operatorname{argmax}\{f(S_{i-1} \cup \{u\}) - f(S_{i-1}) : u \in M_i\}$.
  - $S_i \leftarrow S_{i-1} \cup \{u_i\}$.
- Return S_k.

Section 2.a

Show that the running time of the algorithm is $O(n \ln(1/\varepsilon))$.

בכל איטרציה האלגוריתם מבצע את הפעולות הנ"ל:

- הגרלת $M_i \subseteq N$ בגודל $\left\lceil \frac{nln\left(\frac{1}{\epsilon}\right)}{k} \right\rceil$ מתבצעת בזמן $O\left(\frac{nln\left(\frac{1}{\epsilon}\right)}{k}\right)$.
- מציאת u_i – מתבצעות כאן $O(|M_i|)$ שאילות ל-$value\ oracle$, כלומר סה"כ מתבצע בסיבוכיות זמן של $O(|M_i|) = O\left(\frac{nln\left(\frac{1}{\epsilon}\right)}{k}\right)$.
- השמה ל S_i – מתבצעת בזמן $O(1)$.

כלומר כל איטרציה מתבצעת בסיבוכיות זמן של $O\left(\frac{nln\left(\frac{1}{\epsilon}\right)}{k}\right)$, בסה"כ מבוצעות k איטרציות ולכן סיבוכיות זמן האלגוריתם כולו היא $O\left(nln\left(\frac{1}{\epsilon}\right)\right)$ ■

Section 2.b

Let S^* be some optimal solution to the problem. The current goal is to prove that for every iteration $i = 1, \dots, k$ the following holds:

$$\mathbb{E}[f(S_i) - f(S_{i-1})] \geq \frac{1-\varepsilon}{k}(f(S^*) - \mathbb{E}[f(S_{i-1})]) . \quad (*)$$

Fix i and let $v_1, \dots, v_k$ be the k elements of S^* sorted according to the marginal values:

$$f(S_{i-1} \cup \{v_1\}) - f(S_{i-1}) \geq \dots \geq f(S_{i-1} \cup \{v_k\}) - f(S_{i-1}) .$$

Let X_j be the indicator for the event that $M_i \cap \{v_1, \dots, v_j\} \neq \emptyset$. Follow the following steps to prove $(*)$.

Section 2.b.i

Prove that the following inequality always holds (note that S_{i-1}, X_j and u_i are random variables):

$$f(S_{i-1} \cup \{u_i\}) - f(S_{i-1}) \geq \sum_{j=1}^{k-1} X_j \left[(f(S_{i-1} \cup \{v_j\}) - f(S_{i-1})) - (f(S_{i-1} \cup \{v_{j+1}\}) - f(S_{i-1})) \right] + X_k (f(S_{i-1} \cup \{v_k\}) - f(S_{i-1}))$$

יהי $1 \leq t \leq k$ האינדקס המינימלי עבורו $X_t = 1$, אם לא קיים כזה נגדיר $t = \infty$.

נפריד למקרים עבור t:

- $t = \infty$:

  במקרה זה לכל i, $X_i = 0$. לכן:

$$\sum_{j=1}^{k-1} X_j \left[\left(f(S_{i-1} \cup \{v_j\}) - f(S_{i-1})\right) - \left(f(S_{i-1} \cup \{v_{j+1}\}) - f(S_{i-1})\right) \right] + X_k\left(f(S_{i-1} \cup \{v_k\}) - f(S_{i-1})\right) = 0$$

  בנוסף, ממונוטוניות f מתקיים כי $f\left(S_{i-1} \cup \{u_j\}\right) - f(S_{i-1}) \geq 0$.

  כלומר הטענה מתקיימת.

- $t = k$:

  נסתכל על איטרציה כלשהי בה האלגוריתם ביצע $u_i \leftarrow v_j$.

  נשים לב כי בהכרח $j = k$ כיוון ש $v_1, \dots, v_j$ הינם האיברים של הפיתרון האופטימלי, ממויינים לפי גודל הערכים השוליים. לכן בכל איטרציה של האלגוריתם מתבצע $u_i \leftarrow v_k$ ונקבל:

$$\sum_{j=1}^{k-1} X_j \left[\left(f(S_{i-1} \cup \{v_j\}) - f(S_{i-1})\right) - \left(f(S_{i-1} \cup \{v_{j+1}\}) - f(S_{i-1})\right) \right] + X_k\left(f(S_{i-1} \cup \{v_k\}) - f(S_{i-1})\right) = f(S_{i-1} \cup \{u_i\}) - f(S_{i-1})$$

  כלומר הטענה מתקיימת בשוויון ממש.

- $t < k$:

נשים לב כי מהגדרת X_j מתקיים $X_j = \begin{cases} 0, & j < t \\ 1, & j \geq t \end{cases}$.

מכאן ש:

$$\sum_{j=1}^{k-1} X_j \left[\left(f\left(S_{i-1} \cup \{v_j\}\right) - f(S_{i-1}) \right) - \left(f\left(S_{i-1} \cup \{v_{j+1}\}\right) - f(S_{i-1}) \right) \right] + X_k \left(f(S_{i-1} \cup \{v_k\}) - f(S_{i-1}) \right)$$

$$= \sum_{j=t}^{k-1} \left(f\left(S_{i-1} \cup \{v_j\}\right) - f(S_{i-1}) \right) - \left(f\left(S_{i-1} \cup \{v_{j+1}\}\right) - f(S_{i-1}) \right) + \left(f(S_{i-1} \cup \{v_k\}) - f(S_{i-1}) \right)$$

$$= \sum_{j=t}^{k-1} \left(f\left(S_{i-1} \cup \{v_j\}\right) - f\left(S_{i-1} \cup \{v_{j+1}\}\right) \right) + \left(f(S_{i-1} \cup \{v_k\}) - f(S_{i-1}) \right)$$

$$\underbrace{=}_{\text{טור טלסקופי}} \left(f(S_{i-1} \cup \{v_t\}) - f(S_{i-1} \cup \{v_k\}) \right) + \left(f(S_{i-1} \cup \{v_k\}) - f(S_{i-1}) \right)$$

$$= f(S_{i-1} \cup \{v_t\}) - f(S_{i-1})$$

נסתכל על איטרציה כלשהי בה האלגוריתם ביצע $u_i \longleftarrow v_j$.

כאמור כיוון ש- v_i ממויינים לפי הערכים השוליים בפיתרון האופטימלי ו- v_t הינו האיבר בעל הערך השולי הגבוה ביותר ב-M_i, באירטציה ה- i יתבצע $u_i \longleftarrow v_t$.

לכן,

$$f(S_{i-1} \cup \{v_t\}) - f(S_{i-1}) = f(S_{i-1} \cup \{u_i\}) - f(S_{i-1})$$

כלומר הטענה מתקיימת בשיוויון ממש.

בסה"כ הראינו כי הטענה מתקיימת בשלמותה ■

Section 2.b.ii

Prove that: $\mathbb{E}[X_j] \ge 1 - \left(1 - \frac{\ln(1/\varepsilon)}{k}\right)^j$ for every $j = 1, \ldots, k$.

יהי $j \in [k]$.

נסמן $R = |M_i| = \left\lceil \frac{n ln\left(\frac{1}{\epsilon}\right)}{k} \right\rceil$.

נשים לב שיש $\binom{n}{R}$ אפשרויות לM_i כלשהו ו-$\binom{n-j}{R}$ אפשרויות ל-M_i אשר לא מכיל אף אחד מ $v_1, \dots, v_j$.

לכן,

$$Pr(X_j = 0) = Pr(M_i \cap \{v_1, \dots, v_j\} = \emptyset) = \frac{\binom{n-j}{R}}{\binom{n}{R}}$$

מכאן נקבל:

$$\mathbb{E}(X_j) = Pr(X_j = 1) = 1 - Pr(X_j = 0)$$

$$= 1 - \frac{\binom{n-j}{R}}{\binom{n}{R}} = 1 - \frac{\frac{(n-j)!}{R!\,(n-j-R)!}}{\frac{n!}{R!\,(n-R)!}} = 1 - \left(1 - \frac{R}{n}\right)\left(1 - \frac{R}{n-1}\right) \dots \left(1 - \frac{R}{n-j+1}\right)$$

$$\ge 1 - \left(1 - \frac{R}{n}\right)^j = 1 - \left(1 - \frac{\left\lceil \frac{n ln\left(\frac{1}{\epsilon}\right)}{k} \right\rceil}{n}\right)^j \ge 1 - \left(1 - \frac{\frac{n ln\left(\frac{1}{\epsilon}\right)}{k}}{n}\right)^j$$

$$= 1 - \left(1 - \frac{ln\left(\frac{1}{\epsilon}\right)}{k}\right)^j$$

בסה"כ הוכחנו כי לכל $j \in [k]$ מתקיים $\mathbb{E}(X_j) \ge 1 - \left(1 - \frac{ln\left(\frac{1}{\epsilon}\right)}{k}\right)^j$ כנדרש ■

Section 2.b.iii

Prove $(*)$.

Hint: use Chebyshev's sum inequality, which states that for $a_1 \geq \ldots \geq a_n$ and $b_1 \geq \ldots \geq b_n$ the following holds:

$$\frac{1}{n}\sum_{\ell=1}^{n} a_\ell b_\ell \geq \left(\frac{1}{n}\sum_{\ell=1}^{n} a_\ell\right)\left(\frac{1}{n}\sum_{\ell=1}^{n} b_\ell\right).$$

ראשית, על מנת להשתמש ברמז, נוכיח טענת עזר:

הסדרות $\{\mathbb{E}[f(S_{i-1} \cup \{v_j\}) - f(S_{i-1})]\}_{j=1}^{k}$, $\{\mathbb{E}[X_j - X_{j-1}]\}_{j=1}^{k}$ הינן מונוטוניות יורדות.

נראה כי לכל j, $\mathbb{E}[X_j - X_{j-1}] - \mathbb{E}[X_{j+1} - X_j] \geq 0$:

$$\begin{aligned}
&\mathbb{E}[X_j - X_{j-1}] - \mathbb{E}[X_{j+1} - X_j] \\
&= \mathbb{E}[X_j] - \mathbb{E}[X_{j-1}] - (\mathbb{E}[X_{j+1}] - \mathbb{E}[X_j]) \\
&= Pr(X_j = 1) - Pr(X_{j-1} = 1) - (Pr(X_{j+1} = 1) - Pr(X_j = 1)) \\
&= Pr(M_i \cap \{v_1, \ldots, v_j\} \neq \emptyset) - Pr(M_i \cap \{v_1, \ldots, v_{j-1}\} \neq \emptyset) \\
&\qquad - (Pr(M_i \cap \{v_1, \ldots, v_{j+1}\} \neq \emptyset) - Pr(M_i \cap \{v_1, \ldots, v_j\} \neq \emptyset)) \\
&\underbrace{=}_{\text{חיסור אירועים מוכלים}} Pr(M_i \cap \{v_1, \ldots, v_j\} \neq \emptyset \text{ and } M_i \cap \{v_1, \ldots, v_{j-1}\} = \emptyset) \\
&\qquad - Pr(M_i \cap \{v_1, \ldots, v_{j+1}\} \neq \emptyset \text{ and } M_i \cap \{v_1, \ldots, v_j\} = \emptyset) \\
&\underbrace{=}_{\text{מהגדרת האלגוריתם}} Pr(u_i = v_j) - Pr(u_i = v_{j+1}) \underbrace{\geq}_{(*)} 0
\end{aligned}$$

$(*)$ אם $v_j \in M_i$, מאופן פעולת האלגוריתם נקבל כי $u_i \neq v_{j+1}$ ולכן $Pr(u_i = v_j) \geq Pr(u_i = v_{j+1})$.

בנוסף, מהגדרת ערכי v_i (ממויינים לפי הערכים השוליים) מתקבל מיידית כי הסדרה:

$$\{\mathbb{E}[f(S_{i-1} \cup \{v_j\}) - f(S_{i-1})]\}_{j=1}^{k}$$

הינה מונוטונית יורדת.

עד כה הוכחנו את טענת העזר.

נוכיח כעת את הטענה המרכזית:

$$\mathbb{E}[f(S_i)-f(S_{i-1})]=\mathbb{E}[f(S_{i-1}\cup\{u_i\})-f(S_{i-1})]$$

$$\underbrace{\geq}_{\text{תת סעיף } i \text{ ומונוטוניות התוחלת}} \mathbb{E}\left[\begin{array}{c}\sum_{j=1}^{k-1}X_j\left[\left(f\left(S_{i-1}\cup\{v_j\}\right)-f(S_{i-1})\right)-\left(f\left(S_{i-1}\cup\{v_{j+1}\}\right)-f(S_{i-1})\right)\right]\\ +X_k\left(f(S_{i-1}\cup\{v_k\})-f(S_{i-1})\right)\end{array}\right]$$

$$=\mathbb{E}\left[\sum_{j=1}^{k}X_j\left(f\left(S_{i-1}\cup\{v_j\}\right)-f(S_{i-1})\right)-\sum_{j=1}^{k-1}X_j\left(f\left(S_{i-1}\cup\{v_{j+1}\}\right)-f(S_{i-1})\right)\right]$$

$$=\left[\sum_{j=1}^{k}X_j\left(f\left(S_{i-1}\cup\{v_j\}\right)-f(S_{i-1})\right)-\sum_{j=1}^{k-1}X_{j-1}\left(f\left(S_{i-1}\cup\{v_j\}\right)-f(S_{i-1})\right)\right]$$

$$\underbrace{=}_{\text{המשתנים בסכום בלתי תלויים}}\sum_{j=1}^{k}\mathbb{E}\left[X_j-X_{j-1}\right]\cdot\mathbb{E}\left[f\left(S_{i-1}\cup\{v_j\}\right)-f(S_{i-1})\right]$$

$$=k\frac{1}{k}\sum_{j=1}^{k}\mathbb{E}\left[X_j-X_{j-1}\right]\cdot\mathbb{E}\left[f\left(S_{i-1}\cup\{v_j\}\right)-f(S_{i-1})\right]$$

$$\underbrace{\geq}_{\text{טענת העזר ואי שיוויון צבישב}} k\left(\frac{1}{k}\sum_{j=1}^{k}\mathbb{E}\left[X_j-X_{j-1}\right]\right)\left(\frac{1}{k}\sum_{j=1}^{k}\mathbb{E}\left[f\left(S_{i-1}\cup\{v_j\}\right)-f(S_{i-1})\right]\right)$$

לצורך נוחות נגדיר $X_0=0$,

$$=\frac{1}{k}\ \mathbb{E}[X_k-X_0]\sum_{j=1}^{k}\mathbb{E}\left[f\left(S_{i-1}\cup\{v_j\}\right)-f(S_{i-1})\right]$$

$$\underbrace{=}_{\mathbb{E}[X_0]=0}\frac{1}{k}\ \mathbb{E}[X_k]\sum_{j=1}^{k}\mathbb{E}\left[f\left(S_{i-1}\cup\{v_j\}\right)-f(S_{i-1})\right]$$

$$\underbrace{\geq}_{\text{טענה מתת סעיף } ii}\frac{1}{k}\left(1-\left(1-\frac{ln\left(\frac{1}{\epsilon}\right)}{k}\right)^k\right)\sum_{j=1}^{k}\mathbb{E}\left[f\left(S_{i-1}\cup\{v_j\}\right)-f(S_{i-1})\right]$$

כיוון ש $\lim_{n\to\infty}\left(1-\frac{a}{n}\right)^n=e^{-a}$ מלמטה, נובע כי $\left(1-\frac{a}{n}\right)^n\leq e^{-a}$ ולכן:

$$\geq\frac{1}{k}\left(1-e^{-\ln\left(\frac{1}{\epsilon}\right)}\right)\sum_{j=1}^{k}\mathbb{E}\left[f\left(S_{i-1}\cup\{v_j\}\right)-f(S_{i-1})\right]$$

$$= \frac{1-\epsilon}{k}\sum_{j=1}^{k}\mathbb{E}\left[f\left(S_{i-1}\cup\{v_j\}\right)-f(S_{i-1})\right]$$

$$\underbrace{\geq}_{\text{תת מודולריות של } f} \frac{1-\epsilon}{k}\mathbb{E}[f(S_{i-1}\cup\{v_1,\dots,v_k\})-f(S_{i-1})]$$

$$= \frac{1-\epsilon}{k}\mathbb{E}[f(S_{i-1}\cup S^*)-f(S_{i-1})]$$

$$\underbrace{\geq}_{\text{מונוטוניות } f} \frac{1-\epsilon}{k}\mathbb{E}[f(S^*)-f(S_{i-1})]$$

$$= \frac{1-\epsilon}{k}\left(f(S^*)-\mathbb{E}[f(S_{i-1})]\right)$$

סה"כ הוכחנו את הטענה כנדרש ■

Section 2.c

Assuming $\varepsilon \leq 1 - 1/e$ solve $(*)$ and prove that $\mathbb{E}[f(S_k)] \geq (1 - 1/e - \varepsilon) f(S^*)$.

מהטענה בסעיף הקודם מתקבל:

$$\mathbb{E}[f(S_i) - f(S_{i-1})] \geq \frac{1-\epsilon}{k}\left(f(S^*) - \mathbb{E}[f(S_{i-1})]\right)$$
$$\underbrace{\Longrightarrow}_{\text{ליניאריות התוחלת}} \mathbb{E}[f(S_i)] \geq \frac{1-\epsilon}{k} f(S^*) + \left(1 - \frac{1-\epsilon}{k}\right)\mathbb{E}[f(S_{i-1})]$$

קיבלנו משוואה רקורסיבית מהצורה הכללית $X_n \geq aX_{n-1} + b$ כאשר תנאי הבסיס הינו $X_0 = \mathbb{E}[f(S_0)]$.

ידוע כי פיתרון למשוואה מצורה זו הינו $X_n \geq X_0 a^{n-1} + b\frac{a^n - 1}{a-1}$.

בהצבת הפיתרון במשוואה שלנו נקבל:

$$\mathbb{E}[f(S_k)]$$
$$\geq \mathbb{E}[f(S_0)]\left(1 - \frac{1-\epsilon}{k}\right)^{k-1} + \frac{1-\epsilon}{k} f(S^*) \frac{\left(1 - \frac{1-\epsilon}{k}\right)^k - 1}{\left(1 - \frac{1-\epsilon}{k}\right) - 1}$$
$$= \mathbb{E}[f(S_0)]\left(1 - \frac{1-\epsilon}{k}\right)^{k-1} + \left(1 - \left(1 - \frac{1-\epsilon}{k}\right)^k\right) f(S^*)$$
$$\underbrace{\geq}_{(1)} \left(1 - \left(1 - \frac{1-\epsilon}{k}\right)^k\right) f(S^*)$$
$$\underbrace{\geq}_{(2)} \left(1 - \left(\frac{1}{e}\right)^{1-\epsilon}\right) f(S^*)$$

(1) מתקיים כי $\mathbb{E}[f(S_0)]\left(1 - \frac{1-\epsilon}{k}\right)^{k-1} \geq 0$

(2) כאמור $\left(1 - \frac{a}{n}\right)^n \leq e^{-a}$

עפ"י הזהות: $\forall 0 < a < b < 1,\ a^b < a - b + 1$

עבור: $a = \frac{1}{e},\ b = 1 - \epsilon$

נקבל: $\left(1 - \left(\frac{1}{e}\right)^{1-\epsilon}\right) f(S^*) \geq \left(1 - \frac{1}{e} - \epsilon\right) f(S^*)$

לבסוף, קיבלנו:

$$\mathbb{E}[f(S_k)] \geq \left(1 - \frac{1}{e} - \epsilon\right) f(S^*) \blacksquare$$

Question 3

Let $f : 2^N \to \mathbb{R}_+$ be a submodular function, which is not necessarily monotone, over ground set N. Let $M = (N, I)$ be a partition matroid of rank k where $N = N_1 \cup N_2 \cup \ldots \cup N_k$ is the partition of the ground set N that is associated with the matroid. We are interested in solving $\max \{f(S) : S \subseteq N, S \in I\}$. The goal of this question is obtaining a fast combinatorial algorithm that achieves an approximation of $1/4$ for the problem. Consider the following algorithm that is parameterized by $t \in \mathbb{N}$, which indicates how many iterations it performs:

- $S_0 \leftarrow \emptyset$ and $J_0 \leftarrow \emptyset$.
- for $i = 1$ to t do:
  - $\forall j = 1, \ldots, k$: $u_j^i \leftarrow \text{argmax}\left\{f(S_{i-1} \cup \{u\}) - f(S_{i-1}) : u \in N_j\right\}$.
  - Let j_i be a uniformly random number from $\{1, \ldots, k\}$.
  - $S_i \leftarrow S_{i-1}$ and $J_i \leftarrow J_{i-1}$.
  - if $j_i \notin J_{i-1}$ then:
    - $J_i \leftarrow J_i \cup \{j_i\}$.
    - if $f(S_{i-1} \cup \{u_{j_i}^i\}) - f(S_{i-1}) \geq 0$ then: $S_i \leftarrow S_i \cup \{u_{j_i}^i\}$.
- return S_t.

Section 3.a

Show that $S_i \in I$ for every $i = 1, \ldots, t$.

We show the above property using induction:

Base case $(\mathbf{i = 1})$

As defined in the question:

1. $S_0 = \emptyset$.
2. $J_0 = \emptyset$ then the first condition holds for any choice of random $1 \leq j_i \leq k$:
   Either $S_1 = S_0 \cup \{u_{j_i}^1\}$ or $S_1 = S_0$ (where $u_{j_i}^1$ is in one of N's partition sets N_{j_i}).

Therefore, by the definition of set union we get that $S_1 = \{u_{j_i}^1\}$ or $S_1 = \emptyset$.

As a set of at most one element, it is necessarily in I (the set of all independent sets):

1. S_1 has at most one element from every partition set of N.
2. For all $1 \leq i \leq k$ the dependency bound k_i holds that $k_i \geq 1$.

$\Longrightarrow S_1 \in I$ by definition $(|S_1 \cap N_y| \leq k_y, \forall y = 1, \ldots, k)$.

Inductive Hypothesis

Say $S_i \in I$ for all $i < t_0$ for some $1 < t_0 \leq t$.

Inductive Step ($\boldsymbol{i = t_0}$)

- If $j_i \in J_{i-1}$:
  In that case there was no update in the algorithm beside that of $S_i \longleftarrow S_{i-1}$ and $J_i \longleftarrow J_{i-1}$, then $S_i = S_{i-1}$. The inductive hypothesis yields $S_{i-1} \in I$, then finally $S_i \in I$.
- Else ($j_i \notin J_{i-1}$):
  Since $j_i \notin J_{i-1}$, we now check the following property for $u_{j_i}^i$:
  If $f\left(S_{i-1} \cup \{u_{j_i}^i\}\right) - f(S_{i-1}) < 0$ then $S_i = S_{i-1} \in I$.
  Else:
  1. $S_i = S_{i-1} \cup \{u_{j_i}^i\}$
  2. The inductive hypothesis holds that $S_{i-1} \in I$, therefore, by I's definition S_{i-1} has at most k_y elements from every N_y – the y_{th} partition set of N:
     $|S_{i-1} \cap N_y| \leq k_y, \forall y = 1, \dots, k$.

  The only partition set N_y whose bound k_y might potentially not hold (in I's definition) is N_{j_i}, then we would like to show that $|S_i \cap N_{j_i}| \leq k_{j_i}$.
  Indeed, $j_i \notin J_{i-1}$ so this index was not sampled before so there was never any addition of an element from N_{j_i} to S_{i-1}.
  It means that after the addition of $u_{j_i}^i \in N_{j_i}$ to S_{i-1} we get: $|S_i \cap N_{j_i}| = 1 \leq k_{j_i}$.
  $\Longrightarrow S_i \in I$.

We saw that $S_i \in I$ for all $i \in \{1, \dots, k\}$ ■

Section 3.b

For every $S \in I$ denote by O_S the best feasible completion of S, *i.e.*,

$$O_S = \text{argmax}\{f(S \cup T) : T \subseteq N \setminus S, S \cup T \in I\}.$$

Prove that for every $i = 1, \dots, t$:

$$\mathbb{E}[f(S_i)] \geq \left(1 - \frac{1}{k}\right) \mathbb{E}[f(S_{i-1})] + \frac{1}{k}\mathbb{E}\left[f(S_{i-1} \cup O_{S_{i-1}})\right].$$

Throughout this proof, until just before the end, we fix any choice that the algorithm had made up to the current i (or $i-1$ if relevant) iteration.

Let us work on the right-term of the inequality:

$$\left(1 - \frac{1}{k}\right) \mathbb{E}[f(S_{i-1})] + \frac{1}{k}\mathbb{E}[f(S_{i-1} \cup O_{S_{i-1}})] = \mathbb{E}[f(S_{i-1})] - \frac{1}{k}\mathbb{E}[f(S_{i-1})] + \frac{1}{k}\mathbb{E}[f(S_{i-1} \cup O_{S_{i-1}})]$$

$$= \mathbb{E}[f(S_{i-1})] + \frac{1}{k}\left(\mathbb{E}[f(S_{i-1} \cup O_{S_{i-1}})] - \mathbb{E}[f(S_{i-1})]\right)$$

$$\underbrace{=}_{\substack{expectation's \\ linearity}} \mathbb{E}[f(S_{i-1})] + \mathbb{E}\left[\frac{1}{k}\left(f(S_{i-1} \cup O_{S_{i-1}}) - f(S_{i-1})\right)\right]$$

Now we focus on the second expectancy argument: $\frac{1}{k}\left(f(S_{i-1} \cup O_{S_{i-1}}) - f(S_{i-1})\right)$

By the sub-modularity of f we get:

$$\frac{1}{k}\left(f(S_{i-1} \cup O_{S_{i-1}}) - f(S_{i-1})\right) \leq \frac{1}{k} \sum_{u \in O_{S_{i-1}}} \{f(S_{i-1} \cup \{u\}) - f(S_{i-1})\}$$

Considering each element u in $O_{S_{i-1}}$, the marginal gain of adding u to S_{i-1} will be less than or equal to the gain of adding $u_{j_i}^i$ (because $u_{j_i}^i$ is chosen to maximize the gain).

Summing over all u in $O_{S_{i-1}}$, the total gain will be less than or equal to the gain obtained by summing over all $u_{j_i}^i$ for $j_i \notin J_{i-1}$, then:

$$\frac{1}{k} \sum_{u \in O_{S_{i-1}}} \{f(S_{i-1} \cup \{u\}) - f(S_{i-1})\} \leq \frac{1}{k} \sum_{j_i \notin J_{i-1}} \{f(S_{i-1} \cup \{u_{j_i}^i\}) - f(S_{i-1})\}$$

$$\underbrace{\leq}_{\substack{bigger \\ minuend}} \frac{1}{k} \sum_{j_i \notin J_{i-1}} \{\max\{f(S_{i-1} \cup \{u_{j_i}^i\}), f(S_{i-1})\} - f(S_{i-1})\}$$

We notice that according to the algorithm, for any $j_i \notin J_{i-1}$:

- If $f\left(S_{i-1} \cup \{u_{j_i}^i\}\right) - f(S_{i-1}) \geq 0$:
  1. $S_i = S_{i-1} \cup \{u_{j_i}^i\}$
  2. $f\left(S_{i-1} \cup \{u_{j_i}^i\}\right) \geq f(S_{i-1})$

  $\Longrightarrow f(S_i) = f(S_{i-1} \cup \{u_{j_i}^i\}) = \max\{f\left(S_{i-1} \cup \{u_{j_i}^i\}\right), f(S_{i-1})\}$

- If $f\left(S_{i-1} \cup \{u_{j_i}^i\}\right) - f(S_{i-1}) < 0$:
  1. $S_i = S_{i-1}$
  2. $f\left(S_{i-1} \cup \{u_{j_i}^i\}\right) < f(S_{i-1})$

  $\Longrightarrow f(S_i) = f(S_{i-1}) = \max\{f\left(S_{i-1} \cup \{u_{j_i}^i\}\right), f(S_{i-1})\}$

$\Longrightarrow$ Overall, we see that in case that $j_i \notin J_{i-1}$: $f(S_i) = \max\{f\left(S_{i-1} \cup \{u_{j_i}^i\}\right), f(S_{i-1})\}$.

$$\Longrightarrow \frac{1}{k} \sum_{j_i \notin J_{i-1}} \{\max\{f\left(S_{i-1} \cup \{u_{j_i}^i\}\right), f(S_{i-1})\} - f(S_{i-1})\} \underbrace{=}_{\substack{\text{the}\\ \text{property}\\ \text{above}}} \frac{1}{k} \sum_{j_i \notin J_{i-1}} \{f(S_i) - f(S_{i-1})\}$$

Back to the term at the top of the page, we get:

$$\Longrightarrow \frac{1}{k}\left(f\left(S_{i-1} \cup O_{S_{i-1}}\right) - f(S_{i-1})\right) \leq \frac{1}{k} \sum_{j_i \notin J_{i-1}} \{f(S_i) - f(S_{i-1})\}$$

We notice the following property:

$$\mathbb{E}[f(S_i) - f(S_{i-1}) \mid S_{i-1}] = \sum_{j_i \notin J_{i-1}} \{f(S_i) - f(S_{i-1})\} \cdot P(j_i | S_{i-1}) \underbrace{=}_{(1)} \frac{1}{k} \sum_{j_i \notin J_{i-1}} \{f(S_i) - f(S_{i-1})\}$$

Combining the previous result with the last one:

$$\Longrightarrow \frac{1}{k}\left(f\left(S_{i-1} \cup O_{S_{i-1}}\right) - f(S_{i-1})\right) \leq \mathbb{E}[f(S_i) - f(S_{i-1}) \mid S_{i-1}]$$

$$\mathbb{E}[f(S_i) - f(S_{i-1}) \mid S_{i-1}] \underbrace{=}_{(1)} \mathbb{E}[f(S_i) \mid S_{i-1}] - \mathbb{E}[f(S_{i-1}) \mid S_{i-1}] \underbrace{=}_{(2)} \mathbb{E}[f(S_i) \mid S_{i-1}] - f(S_{i-1})$$

Consequently, we get:

$$\frac{1}{k}\left(f\left(S_{i-1} \cup O_{S_{i-1}}\right) - f(S_{i-1})\right) \leq \mathbb{E}[f(S_i) \mid S_{i-1}] - f(S_{i-1})$$

$$\Longrightarrow \mathbb{E}[f(S_i) \mid S_{i-1}] \geq f(S_{i-1}) + \frac{1}{k}\left(f\left(S_{i-1} \cup O_{S_{i-1}}\right) - f(S_{i-1})\right)$$

Now we unfix the conditions over previous choices by the algorithm.

Since the inequalities above hold almost surely,

$$\mathbb{E}\left[\mathbb{E}[f(S_i)|\, S_{i-1}]\right] \geq \mathbb{E}\left[f(S_{i-1}) + \frac{1}{k}\left(f\left(S_{i-1} \cup O_{S_{i-1}}\right) - f(S_{i-1})\right)\right]$$

$$\underbrace{=}_{(1)} \mathbb{E}[f(S_{i-1})] + \mathbb{E}\left[\frac{1}{k}\left(f\left(S_{i-1} \cup O_{S_{i-1}}\right) - f(S_{i-1})\right)\right]$$

$$\underbrace{=}_{(3)} \left(1 - \frac{1}{k}\right)\mathbb{E}[f(S_{i-1})] + \frac{1}{k}\mathbb{E}\left[f\left(S_{i-1} \cup O_{S_{i-1}}\right)\right]$$

The law of total expectation yields $\mathbb{E}\left[\mathbb{E}[f(S_i)|S_{i-1}]\right] = \mathbb{E}[f(S_i)]$

$$\Longrightarrow \mathbb{E}[f(S_i)] \geq \left(1 - \frac{1}{k}\right)\mathbb{E}[f(S_{i-1})] + \frac{1}{k}\mathbb{E}\left[f\left(S_{i-1} \cup O_{S_{i-1}}\right)\right] \blacksquare$$

(1) j_i *is a uniformly random number from* $\{1, \dots, k\}$.

(2) *given* S_{i-1} *the value* $f(S_{i-1})$ *is constant and the expectation of a constant equals the constant.*

(3) *the derivations are shown at the beginning of the section.*

Section 3.c

Prove that for every $i-1,\dots,t$:

$$\mathbb{E}\left[f(S_i \cup O_{S_i})\right] \geq \left(1-\frac{2}{k}\right)\mathbb{E}\left[f(S_{i-1} \cup O_{S_{i-1}})\right].$$

Like section b, throughout this proof, until just before the end, we fix any choice that the algorithm had made up to the current i (or $i-1$ if relevant) iteration.

Recall $O_{S_{i-1}}$ the best feasible completion of S_{i-1}, and define $M_i = \{u_1^i, u_2^i, \dots, u_k^i\}$ for every iteration i, where each u_j^i is the element chosen in the algorithm as argmax for all j.

We first define a function $g: M_i \to O_{S_{i-1}}$ in a way that for every $u_{j_i}^i \in M_i$: $g(u_{j_i}^i)$ and $u_{j_i}^i$ are both elements of the same partition set N_{j_i} of N.

(It can be played by setting each partition set to have a "dummy" element that yields no gain, in that case for any partition set N_j that contributed no elements to the current S, the completion O_S has no elements from N_j as well (and vice versa). This way g as a bijection is well defined).

We start by exploring the expression:

$$\mathbb{E}[f(S_i \cup O_{S_i})] - f(S_{i-1} \cup O_{S_{i-1}})$$

As we have shown in section b, given the environment based by S_{i-1}:

$$\mathbb{E}[f(S_i \cup O_{S_i})] = \frac{1}{k}\sum_{u^i \in M_i} f\left((S_{i-1} \cup \{u^i\}) \cup O_{S_{i-1}\cup\{u^i\}}\right)$$

$$\Rightarrow \mathbb{E}[f(S_i \cup O_{S_i})] - f(S_{i-1} \cup O_{S_{i-1}}) = \frac{1}{k}\sum_{u^i \in M_i} f\left((S_{i-1} \cup \{u^i\}) \cup O_{S_{i-1}\cup\{u^i\}}\right) - f(S_{i-1} \cup O_{S_{i-1}})$$

$$\underbrace{=}_{\substack{|M_i|=k \\ \text{and the value} \\ \text{is constant in } u^i}} \frac{1}{k}\sum_{u^i \in M_i} f\left((S_{i-1} \cup \{u^i\}) \cup O_{S_{i-1}\cup\{u^i\}}\right) - \frac{1}{k}\sum_{u^i \in M_i} f(S_{i-1} \cup O_{S_{i-1}})$$

$$\underbrace{=}_{linearity} \frac{1}{k}\sum_{u^i \in M_i}\left(f\left((S_{i-1} \cup \{u^i\}) \cup O_{S_{i-1}\cup\{u^i\}}\right) - f(S_{i-1} \cup O_{S_{i-1}})\right)$$

Since $O_{S_{i-1}\cup\{u^i\}}$ the best completion of $S_{i-1} \cup \{u^i\}$, we get for any set $O_{S_{i-1}\cup\{u^i\}} \neq A \subseteq N$:

$$f\left((S_{i-1} \cup \{u^i\}) \cup O_{S_{i-1}\cup\{u^i\}}\right) \geq f\left((S_{i-1} \cup \{u^i\}) \cup A\right)$$

Specifically, $A = O_{S_{i-1}} \backslash \{g(u^i)\}$ for any $u^i \in M_i$.

$$\Longrightarrow \frac{1}{k}\sum_{u^i\in M_i}\left(f\left(\left(S_{i-1}\cup\{u^i\}\right)\cup O_{S_{i-1}\cup\{u^i\}}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)\right)$$
$$\geq \frac{1}{k}\sum_{u^i\in M_i}\left(f\left(S_{i-1}\cup\{u^i\}\cup O_{S_{i-1}}\backslash\{g(u^i)\}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)\right)$$

Notice that if $u^i = g(u^i)$, then $\{u^i\}\cup O_{S_{i-1}}\backslash\{g(u^i)\} = O_{S_{i-1}}$ and the summation argument will be zero.

Therefore, we can focus on $u^i \in M_i$ s.t $u^i \neq g(u^i)$:

$$= \frac{1}{k}\sum_{g(u^i)\neq u^i\in M_i}\left(f\left(S_{i-1}\cup\{u^i\}\cup O_{S_{i-1}}\backslash\{g(u^i)\}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)\right)$$

Consider the following sets:

- $A = S_{i-1}\cup O_{S_{i-1}}$
- $B = S_{i-1}\cup O_{S_{i-1}}\backslash\{g(u^i)\}$

We see that $B \subseteq A$, and since f is sub-modular, we get for any $u^i \in M_i$:

$$f\left(B\cup\{u^i\}\right)-f(B)\geq f\left(A\cup\{u^i\}\right)-f(A)$$

$$\Longrightarrow f\left(S_{i-1}\cup O_{S_{i-1}}\backslash\{g(u^i)\}\cup\{u^i\}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\backslash\{g(u^i)\}\right)$$
$$\geq f\left(S_{i-1}\cup O_{S_{i-1}}\cup\{u^i\}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)$$

$$\Longrightarrow f\left(S_{i-1}\cup O_{S_{i-1}}\backslash\{g(u^i)\}\cup\{u^i\}\right)\geq f\left(S_{i-1}\cup O_{S_{i-1}}\cup\{u^i\}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)$$
$$+f\left(S_{i-1}\cup O_{S_{i-1}}\backslash\{g(u^i)\}\right)$$

Subtracting $f\left(S_{i-1}\cup O_{S_{i-1}}\right)$ from both sides, we get:

$$f\left(S_{i-1}\cup O_{S_{i-1}}\backslash\{g(u^i)\}\cup\{u^i\}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)$$
$$\geq f\left(S_{i-1}\cup O_{S_{i-1}}\cup\{u^i\}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)$$
$$+f\left(S_{i-1}\cup O_{S_{i-1}}\backslash\{g(u^i)\}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)$$

Summing over all $u^i \in M_i$ s.t $u^i \neq g(u^i)$, and multiplying by $\frac{1}{k}$:

$$\frac{1}{k}\sum_{g(u^i)\neq u^i\in M_i}\left(f\left(S_{i-1}\cup\{u^i\}\cup O_{S_{i-1}}\setminus\{g(u^i)\}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)\right)$$
$$\geq \frac{1}{k}\sum_{g(u^i)\neq u^i\in M_i}\left(f\left(S_{i-1}\cup\{u^i\}\cup O_{S_{i-1}}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)\right)$$
$$+\frac{1}{k}\sum_{g(u^i)\neq u^i\in M_i}\left(f\left(S_{i-1}\cup O_{S_{i-1}}\setminus\{g(u^i)\}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)\right)$$

Overall, by now we got:

$$\mathbb{E}\left[f\left(S_i\cup O_{S_i}\right)\right]-f\left(S_{i-1}\cup O_{S_{i-1}}\right)$$
$$\geq \frac{1}{k}\sum_{g(u^i)\neq u^i\in M_i}\left(f\left(S_{i-1}\cup\{u^i\}\cup O_{S_{i-1}}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)\right)$$
$$+\frac{1}{k}\sum_{g(u^i)\neq u^i\in M_i}\left(f\left(S_{i-1}\cup O_{S_{i-1}}\setminus\{g(u^i)\}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)\right)$$

We could reduce the expression even more if we iterate over $u^i \in M_i$ whether $u^i = g(u^i)$ or not:

- If $u^i = g(u^i)$:
  Since $g: M_i \to O_{S_{i-1}}$, we get that $g(u^i) \in O_{S_{i-1}}$, and with this equality $u^i \in O_{S_{i-1}}$.
  Therefore, $\{u^i\}\cup O_{S_{i-1}} = O_{S_{i-1}}$ so $f\left(S_{i-1}\cup\{u^i\}\cup O_{S_{i-1}}\right) = f\left(S_{i-1}\cup O_{S_{i-1}}\right)$ and we will be adding more zeros to the first summation so it will not change.
- If $u^i \neq g(u^i)$:
  Since $O_{S_{i-1}}$ is the best completion to S_{i-1}, $f\left(S_{i-1}\cup O_{S_{i-1}}\right) \geq f\left(S_{i-1}\cup O_{S_{i-1}}\setminus\{g(u^i)\}\right)$ so we only add non-positive values to the sum, thus it helps us reduce the expression.

Then,

$$\geq \frac{1}{k}\sum_{u^i\in M_i}\left(f\left(S_{i-1}\cup\{u^i\}\cup O_{S_{i-1}}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)\right)$$
$$+\frac{1}{k}\sum_{u^i\in M_i}\left(f\left(S_{i-1}\cup O_{S_{i-1}}\setminus\{g(u^i)\}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)\right)$$

By property of sub-modularity, we get:

$$\geq \frac{1}{k}\left[f\left(S_{i-1}\cup M_i\cup O_{S_{i-1}}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)\right]+\frac{1}{k}\left[f\left(S_{i-1}\cup O_{S_{i-1}}\setminus O_{S_{i-1}}\right)-f\left(S_{i-1}\cup O_{S_{i-1}}\right)\right]$$
$$=\frac{1}{k}f\left(S_{i-1}\cup M_i\cup O_{S_{i-1}}\right)-\frac{1}{k}f\left(S_{i-1}\cup O_{S_{i-1}}\right)+\frac{1}{k}f\left(S_{i-1}\right)-\frac{1}{k}f\left(S_{i-1}\cup O_{S_{i-1}}\right)$$
$$\underbrace{\geq}_{f\geq 0} -\frac{1}{k}f\left(S_{i-1}\cup O_{S_{i-1}}\right)-\frac{1}{k}f\left(S_{i-1}\cup O_{S_{i-1}}\right) = -\frac{2}{k}f\left(S_{i-1}\cup O_{S_{i-1}}\right)$$

Finally, we have:

$$\mathbb{E}\left[f\left(S_i \cup O_{S_i}\right)\right] - f\left(S_{i-1} \cup O_{S_{i-1}}\right) \geq -\frac{2}{k} f\left(S_{i-1} \cup O_{S_{i-1}}\right)$$

$$\Longrightarrow \mathbb{E}\left[f\left(S_i \cup O_{S_i}\right)\right] \geq \left(1 - \frac{2}{k}\right) f\left(S_{i-1} \cup O_{S_{i-1}}\right)$$

Now we unfix the conditions over previous choices by the algorithm, hence:

$$\mathbb{E}\left[f\left(S_i \cup O_{S_i}\right) \mid S_{i-1}\right] \geq \left(1 - \frac{2}{k}\right) f\left(S_{i-1} \cup O_{S_{i-1}}\right)$$

Since the inequality holds almost surely,

$$\mathbb{E}\left[\mathbb{E}\left[f\left(S_i \cup O_{S_i}\right) \mid S_{i-1}\right]\right] \geq \mathbb{E}\left[\left(1 - \frac{2}{k}\right) f\left(S_{i-1} \cup O_{S_{i-1}}\right)\right]$$

By the law of total expectation, and linearity of expectation:

$$\mathbb{E}\left[f\left(S_i \cup O_{S_i}\right)\right] \geq \left(1 - \frac{2}{k}\right) \mathbb{E}\left[f\left(S_{i-1} \cup O_{S_{i-1}}\right)\right] \blacksquare$$

Section 3.d

> Solve the above recursive formula and prove that for every $i = 1, \ldots, t$:
>
> $$\mathbb{E}[f(S_i)] \geq \left(\left(1 - \frac{1}{k}\right)^i - \left(1 - \frac{2}{k}\right)^i \right) f(S^*).$$
>
> Here S^* is some optimal solution to the problem.
>
> Hint: note that without loss of generality $O_\emptyset = S^*$.

Using the results from both sections b and c, we will expand the following recursive formula:

$$\mathbb{E}[f(S_i)] \geq \left(1 - \frac{1}{k}\right) \mathbb{E}[f(S_{i-1})] + \frac{1}{k} \mathbb{E}\left[f\left(S_{i-1} \cup O_{S_{i-1}}\right)\right]$$

First derivation:

$$\begin{cases} \mathbb{E}[f(S_{i-1})] \geq \left(1 - \frac{1}{k}\right) \mathbb{E}[f(S_{i-2})] + \frac{1}{k} \mathbb{E}\left[f\left(S_{i-2} \cup O_{S_{i-2}}\right)\right] \\ \quad \mathbb{E}\left[f\left(S_{i-1} \cup O_{S_{i-1}}\right)\right] \geq \left(1 - \frac{2}{k}\right) \mathbb{E}\left[f\left(S_{i-2} \cup O_{S_{i-2}}\right)\right] \end{cases}$$

$$\geq \left(1 - \frac{1}{k}\right)\left(\left(1 - \frac{1}{k}\right) \mathbb{E}[f(S_{i-2})] + \frac{1}{k} \mathbb{E}\left[f\left(S_{i-2} \cup O_{S_{i-2}}\right)\right]\right) + \frac{1}{k}\left(1 - \frac{2}{k}\right) \mathbb{E}\left[f\left(S_{i-2} \cup O_{S_{i-2}}\right)\right]$$

$$= \left(1 - \frac{1}{k}\right)\left(1 - \frac{1}{k}\right) \mathbb{E}[f(S_{i-2})] + \underbrace{\frac{1}{k}\left(1 - \frac{1}{k}\right) \mathbb{E}\left[f\left(S_{i-2} \cup O_{S_{i-2}}\right)\right]}_{\geq 0} + \frac{1}{k}\left(1 - \frac{2}{k}\right) \mathbb{E}\left[f\left(S_{i-2} \cup O_{S_{i-2}}\right)\right]$$

$$\geq \left(1 - \frac{1}{k}\right)^2 \mathbb{E}[f(S_{i-2})] + \frac{1}{k}\left(1 - \frac{2}{k}\right) \mathbb{E}\left[f\left(S_{i-2} \cup O_{S_{i-2}}\right)\right]$$

Second derivation:

$$\begin{cases} \mathbb{E}[f(S_{i-2})] \geq \left(1 - \frac{1}{k}\right) \mathbb{E}[f(S_{i-3})] + \frac{1}{k} \mathbb{E}\left[f\left(S_{i-3} \cup O_{S_{i-3}}\right)\right] \\ \quad \mathbb{E}\left[f\left(S_{i-2} \cup O_{S_{i-2}}\right)\right] \geq \left(1 - \frac{2}{k}\right) \mathbb{E}\left[f\left(S_{i-3} \cup O_{S_{i-3}}\right)\right] \end{cases}$$

$$\geq \left(1 - \frac{1}{k}\right)^2 \left(\left(1 - \frac{1}{k}\right) \mathbb{E}[f(S_{i-3})] + \frac{1}{k} \mathbb{E}\left[f\left(S_{i-3} \cup O_{S_{i-3}}\right)\right]\right) + \frac{1}{k}\left(1 - \frac{2}{k}\right)\left(1 - \frac{2}{k}\right) \mathbb{E}\left[f\left(S_{i-3} \cup O_{S_{i-3}}\right)\right]$$

$$= \left(1 - \frac{1}{k}\right)^3 \mathbb{E}[f(S_{i-3})] + \underbrace{\frac{1}{k}\left(1 - \frac{1}{k}\right)^2 \mathbb{E}\left[f\left(S_{i-3} \cup O_{S_{i-3}}\right)\right]}_{\geq 0} + \frac{1}{k}\left(1 - \frac{2}{k}\right)\left(1 - \frac{2}{k}\right) \mathbb{E}\left[f\left(S_{i-3} \cup O_{S_{i-3}}\right)\right]$$

$$\geq \left(1 - \frac{1}{k}\right)^3 \mathbb{E}[f(S_{i-3})] + \frac{1}{k}\left(1 - \frac{2}{k}\right)^2 \mathbb{E}\left[f\left(S_{i-3} \cup O_{S_{i-3}}\right)\right]$$

And so on, until:

$(\boldsymbol{i-1})_{\boldsymbol{th}}$ derivation:

$$\begin{cases} \mathbb{E}[f(S_1)] \geq \left(1-\frac{1}{k}\right)\mathbb{E}[f(S_0)] + \frac{1}{k}\mathbb{E}\left[f\left(S_0 \cup O_{S_0}\right)\right] = \left(1-\frac{1}{k}\right)\mathbb{E}[f(\emptyset)] + \frac{1}{k}\mathbb{E}[f(\emptyset \cup O_\emptyset)] \\ \mathbb{E}\left[f\left(S_1 \cup O_{S_1}\right)\right] \geq \left(1-\frac{2}{k}\right)\mathbb{E}\left[f\left(S_0 \cup O_{S_0}\right)\right] = \left(1-\frac{2}{k}\right)\mathbb{E}[f(\emptyset \cup O_\emptyset)] \end{cases}$$

$$\geq \left(1-\frac{1}{k}\right)^{i-1}\mathbb{E}[f(S_0)] + \frac{1}{k}\left(1-\frac{2}{k}\right)^{i-2}\mathbb{E}\left[f\left(S_0 \cup O_{S_0}\right)\right]$$

$$\geq \left(1-\frac{1}{k}\right)^{i-1}\left(\left(1-\frac{1}{k}\right)\mathbb{E}[f(\emptyset)] + \frac{1}{k}\mathbb{E}[f(\emptyset \cup O_\emptyset)]\right) + \frac{1}{k}\left(1-\frac{2}{k}\right)^{i-2}\left(1-\frac{2}{k}\right)\mathbb{E}[f(\emptyset \cup O_\emptyset)]$$

$$= \underbrace{\left(1-\frac{1}{k}\right)^{i}\mathbb{E}[f(\emptyset)]}_{\geq 0} + \frac{1}{k}\left(1-\frac{1}{k}\right)^{i-1}\mathbb{E}[f(\emptyset \cup O_\emptyset)] + \frac{1}{k}\left(1-\frac{2}{k}\right)^{i-1}\mathbb{E}[f(\emptyset \cup O_\emptyset)]$$

$$\underbrace{\geq}_{\left(1-\frac{1}{k}\right)\geq\frac{1}{k}} \left(1-\frac{1}{k}\right)^{i}\mathbb{E}[f(\emptyset \cup O_\emptyset)] + \left(1-\frac{2}{k}\right)^{i}\mathbb{E}[f(\emptyset \cup O_\emptyset)]$$

$$\underbrace{=}_{\substack{Union\ with \\ empty\ set}} \left(1-\frac{1}{k}\right)^{i}\mathbb{E}[f(O_\emptyset)] + \left(1-\frac{2}{k}\right)^{i}\mathbb{E}[f(O_\emptyset)]$$

$$\underbrace{=}_{\substack{Given\ in \\ the\ question}} \left(1-\frac{1}{k}\right)^{i}\mathbb{E}[f(S^*)] + \left(1-\frac{2}{k}\right)^{i}\mathbb{E}[f(S^*)]$$

$$= \left(\left(1-\frac{1}{k}\right)^{i} + \left(1-\frac{2}{k}\right)^{i}\right)\mathbb{E}[f(S^*)]$$

$$\underbrace{\geq}_{\substack{C(A+B)\geq C(A-B) \\ where\ A,B,C\geq 0}} \left(\left(1-\frac{1}{k}\right)^{i} - \left(1-\frac{2}{k}\right)^{i}\right)\mathbb{E}[f(S^*)]$$

$\Longrightarrow$ For every $i = 1,..,t$: $\mathbb{E}[f(S_i)] \geq \left(\left(1-\frac{1}{k}\right)^{i} - \left(1-\frac{2}{k}\right)^{i}\right)\mathbb{E}[f(S^*)]$ ■

Section 3.e

Show that there is a choice of $t \in \mathbb{N}$ such that $\mathbb{E}[f(S_t)] \geq 1/4 \cdot f(S^*)$.

The choice for t is a function of k since we would like to use the property from section d:

$$\text{For every } i = 1,..,t\text{: } \mathbb{E}[f(S_i)] \geq \left(\left(1-\tfrac{1}{k}\right)^i - \left(1-\tfrac{2}{k}\right)^i\right) \mathbb{E}[f(S^*)]$$

We choose $\boldsymbol{t = \lceil k \cdot \ln(2) \rceil}$

- If $k = 1$: $t = \lceil \ln(2) \rceil = 1$

$$\Rightarrow \mathbb{E}[f(S_t)] = \mathbb{E}[f(S_1)] \underbrace{\geq}_{d} \left(\left(1-\frac{1}{1}\right)^1 - \left(1-\frac{2}{1}\right)^1\right) \mathbb{E}[f(S^*)] = \mathbb{E}[f(S^*)] \underbrace{\geq}_{\substack{f \text{ is non} \\ \text{negative}}} \frac{1}{4}\mathbb{E}[f(S^*)]$$

- *else* $(k \geq 2)$:

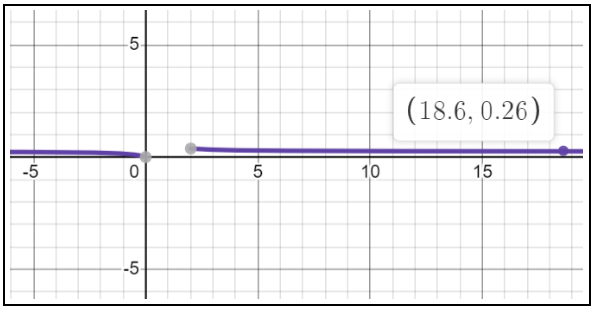
$$\Rightarrow \mathbb{E}[f(S_t)] \geq \left(\left(1-\frac{1}{k}\right)^{\lceil k \cdot \ln(2) \rceil} - \left(1-\frac{2}{k}\right)^{\lceil k \cdot \ln(2) \rceil}\right) \mathbb{E}[f(S^*)]$$

We observe the function $g(k) = \left(\left(1-\frac{1}{k}\right)^{\lceil k \cdot \ln(2) \rceil} - \left(1-\frac{2}{k}\right)^{\lceil k \cdot \ln(2) \rceil}\right)$ as k increases:

Using Wolfram-Alpha website for limits calculations:

$$\lim_{k \to \infty} \left(\left(1-\frac{1}{k}\right)^{\lceil k \log(2) \rceil} - \left(1-\frac{2}{k}\right)^{\lceil k \log(2) \rceil}\right) = \frac{1}{4}$$

We also see that (without the ceiling for $kln(2)$) the function converges from above:



(Below we show these properties mathematically, we now use this method for simplicity)

Therefore, in case $k \geq 2$, the choice of $t = \lceil k \cdot \ln(2) \rceil \in \mathbb{N}$ guarantees that:

$$\mathbb{E}[f(S_t)] \geq \left(\left(1-\frac{1}{k}\right)^t - \left(1-\frac{2}{k}\right)^t\right) \mathbb{E}[f(S^*)] \underbrace{\geq}_{\substack{\text{converge from} \\ \text{above to } \frac{1}{4}}} \frac{1}{4}\mathbb{E}[f(S^*)]$$

$$\Rightarrow \textit{There is a choice of } t \in \mathbb{N} \textit{ such that } \mathbb{E}[f(S_t)] \geq \frac{1}{4}\mathbb{E}[f(S^*)] \ \blacksquare$$

Appendix for section 3.e

We will show that:

$$\lim_{k\to\infty}\left(\left(1-\frac{1}{k}\right)^{k\cdot\ln(2)}-\left(1-\frac{2}{k}\right)^{k\cdot\ln(2)}\right)=\frac{1}{4}\ \text{(From above)}$$

$\Longrightarrow$ We deduce the same for $g(k)=\left(\left(1-\frac{1}{k}\right)^{\lceil k\cdot\ln(2)\rceil}-\left(1-\frac{2}{k}\right)^{\lceil k\cdot\ln(2)\rceil}\right)$

Indeed,

$$\lim_{k\to\infty}\left(\left(1-\frac{1}{k}\right)^{k\cdot\ln(2)}-\left(1-\frac{2}{k}\right)^{k\cdot\ln(2)}\right)=\lim_{k\to\infty}\left(\left(\left(1-\frac{1}{k}\right)^{k}\right)^{\ln(2)}-\left(\left(1-\frac{2}{k}\right)^{k}\right)^{\ln(2)}\right)$$

$$\underbrace{=}_{\substack{Euler\\ limit\,form}}\ e^{-\ln(2)}-e^{-2\ln(2)}=\frac{1}{2}-\frac{1}{4}=\frac{1}{4}$$

- Since $\left(1-\frac{a}{x}\right)^{g(x)}$ for a positive and monotone $g(x)$ and a positive a, is a form of Euler limit that converges from above, we get that the calculated limit is 0.25 from above.
- The same holds for the original expression with the ceiling:

$$\left(\left(1-\frac{1}{k}\right)^{\lceil k\cdot\ln(2)\rceil}-\left(1-\frac{2}{k}\right)^{\lceil k\cdot\ln(2)\rceil}\right)$$

  Due to the definition of limit in infinity, which holds for any series of arguments that converges to infinity, inserted in the function that satisfies the limit (it is with a slight approximation but it’s the closer we get to the point in a readable and straight-forward approach).

Question 4

Prove the following properties of continuous extensions of submodular functions.

Section 4.a

(a) Recall that the concave closure $f^+ : [0,1]^N \to \mathbb{R}_+$ of a set function $f : 2^N \to \mathbb{R}_+$ is defined as follows:

$$f^+(\mathbf{x}) \triangleq \max\{\sum_{S\subseteq N} \alpha_S f(S) : \sum_{S\subseteq N} \alpha_S = 1, \alpha_S \geq 0\ \forall S \subseteq N, \sum_{S\subseteq N} \alpha_S \mathbf{1}_S = \mathbf{x}\} \quad \forall \mathbf{x} \in [0,1]^N$$

Prove that there exists $\mathbf{x} \in [0,1]^N$ for which it is NP-hard to compute $f^+(\mathbf{x})$ assuming f is submodular.

Hint: reduce from the Max-Cut problem.

Reminder- the max cut problem:

Let $G = (V, E)$ be an unweighted graph. The problem is to find a set $S \subset V$ such that the number of edges between S and $\bar{S} = V/S$ is maximized.

Given a set of vertices $\emptyset \neq S \subset V$, let $f(S)$ be the number of edges crossing the cut defined by S (i.e., the number of edges with one endpoint in S and the other endpoint in $\bar{S}$).

Notice that f is sub-modular, as seen in class.

We wish to find the maximal non-trivial cut, reducing to the problem of finding the maximum of a concave function, more specifically the vector $\bar{x}$ that maximizes f^+ that is defined here below.

Define the following:

1. f^+- the concave closure of f we defined above.
2. S_{OPT} – the set of vertices from V for which the number of vertices crossing the cut it defines is maximal, meaning that $\forall \emptyset \neq S \subset V$: $f(S_{OPT}) \geq f(S)$.
3. $\{\alpha_s\}_{\emptyset \neq S \subset V}$ – an arbitrary distribution.

Let $\bar{x} \in [0,1]^N$,

$$f^+(\bar{x}) \underbrace{=}_{definition} \sum_{\emptyset \neq S \subset V} \alpha_s f(S) \underbrace{\leq}_{S_{OPT}\ is\ optimal} \sum_{\emptyset \neq S \subset V} \alpha_s f(S_{OPT}) \underbrace{=}_{f(S_{OPT})\ is\ constant} f(S_{OPT}) \sum_{\emptyset \neq S \subset V} \alpha_s \underbrace{=}_{\sum_{\emptyset \neq S \subset V}(\alpha_s)=1} f(S_{OPT})$$

$\Longrightarrow$ Thus $f^+(\bar{x})$ is bounded by $f(S_{OPT})$, which sets up the stage for us to show that computing $f^+(\bar{x})$ is also NP-Hard.

In order to establish the NP-Hardness of computing $f^+(\bar{x})$, we will focus on a specific vector $\bar{x}$ where each component is set to 0.5: $\bar{x} = 0.5 \cdot \bar{1}$

Our goal is to show that for this particular $\bar{x}$, the value of $f^+(\bar{x})$ corresponds to the maximum cut in the graph, denoted as $f(S_{OPT})$.

We finally define a distribution:

$$\beta_s = \begin{cases} 0.5, & S = S_{OPT}, V/S_{OPT} \\ 0, & else \end{cases}$$

β_s answers the requirements:

1. $\sum_{S\subset V} \alpha_s = \alpha_{S_{opt}} + 0 + \cdots + 0 + \alpha_{V/S_{opt}} = 0.5 + 0.5 = 1$
2. $\forall S \subset N, \alpha_s \geq 0$
3. $\sum_{S\subset N} \alpha_s 1_s \underbrace{=}_{S_{opt}\ vs.\ V/S_{opt}} (0.5, \dots, 0.5) = \bar{x}$

Then,

$$f^+(\bar{x}) = \sum_{\emptyset \neq S \subset V} \beta_s f(S) = \beta_{S_{OPT}} f(S_{OPT}) + \beta_{V/S_{OPT}} f(V/S_{OPT})$$

$$\underbrace{=}_{\substack{symmetry\ of \\ S\ vs.\ V/S\ cut}} \beta_{S_{OPT}} f(S_{OPT}) + \beta_{S_{OPT}} f(S_{OPT}) = 0.5 f(S_{OPT}) + 0.5 f(S_{OPT}) = f(S_{OPT})$$

By demonstrating this equivalence, we determine that given a graph, if $f^+(\bar{x})$ (which is based on the graph) could be computed in polynomial time, then the maximum cut size in the graph could also be found in polynomial time, which contradicts the known NP-Hardness of the Max-Cut problem.

$\Longrightarrow$ given the $\bar{x}$ vector we chose: $\bar{x} = 0.5 \cdot \bar{1}$, computing $f^+(\bar{x})$ is NP-Hard ■

Section 4.b

> Let $F : [0,1]^N \to \mathbb{R}_+$ be the multilinear extensions of a submodular set function $f : 2^N \to \mathbb{R}_+$. Prove that for every $\mathbf{x} \in [0,1]^N$ and non-negative direction vector $\mathbf{d} \in \mathbb{R}_+^N$, *i.e.*, $d_i \geq 0$ for every $i \in N$, the function $g(s) \triangleq F(\mathbf{x} + s\mathbf{d})$ is concave in s (here $s \in \mathbb{R}$).

Let $\bar{x} \in [0,1]^N$ of the form $\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$ and an arbitrary number $j \in [N]$.

Let $\bar{x}_{[j\to 1]}$ be a modified version of vector $\bar{x}$ where the j_{th} entry is set to 1, while all the rest remain unchanged.

Similarly, $\bar{x}_{[j\to 0]}$ is a modified version of vector $\bar{x}$ where the j_{th} entry is set to 0.

Formally:

1. $\bar{x}_{[j\to 0]} = \begin{bmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_N \end{bmatrix} \wedge \begin{bmatrix} 1 \\ \vdots \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \bar{x} \wedge 1_{\{1,\dots,N\}/\{j\}}$
2. $\bar{x}_{[j\to 1]} = \begin{bmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_N \end{bmatrix} \vee \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \bar{x} \vee 1_{\{j\}}$

As seen in class:

$$\frac{\partial F}{\partial x_j} = \frac{F(\bar{x}_{[j\to 1]}) - F(\bar{x}_{[j\to 0]})}{1-0} = F(\bar{x}_{[j\to 1]}) - F(\bar{x}_{[j\to 0]})$$

Let $k \in [N]$.

We define $\bar{x}_{[j\to a][k\to b]}$ be a modified version of vector $\bar{x}$ where the j_{th} entry is set to a and the k_{th} entry is set to b, while all the rest remain unchanged, for $a, b \in \{0,1\}$.

If $k = j$, we clearly get $\frac{\partial^2 F}{\partial x_k \partial x_j} = 0$.

Else,

$$\frac{\partial^2 F}{\partial x_k \partial x_j} \underbrace{=}_{\substack{we\ saw \\ above}} \frac{\partial\left(F(\bar{x}_{[j\to 1]}) - F(\bar{x}_{[j\to 0]})\right)}{\partial x_k} = \frac{\partial\left(F(\bar{x}_{[j\to 1]})\right) - \partial\left(F(\bar{x}_{[j\to 0]})\right)}{\partial x_k}$$

$$= \frac{\partial\left(F(\bar{x}_{[j\to 1]})\right)}{\partial x_k} - \frac{\partial\left(F(\bar{x}_{[j\to 0]})\right)}{\partial x_k}$$

Just as before,

$$= \frac{F(\bar{x}_{[j\to 1][k\to 1]}) - F(\bar{x}_{[j\to 1][k\to 0]})}{1-0} - \frac{F(\bar{x}_{[j\to 0][k\to 1]}) - F(\bar{x}_{[j\to 0][k\to 0]})}{1-0}$$

$$= \left(F(\bar{x}_{[j\to 1][k\to 1]}) - F(\bar{x}_{[j\to 1][k\to 0]})\right) - \left(F(\bar{x}_{[j\to 0][k\to 1]}) - F(\bar{x}_{[j\to 0][k\to 0]})\right)$$

Let $R \subseteq N$ be a set of with elements were sampled independently with probabilities x_i for $i \in [N]$, then as we saw in class: $F(\bar{x}) = \mathbb{E}[f(R)]$. Therefore, we get the following:

1. $F(\bar{x}_{[j\to 1][k\to 1]}) \underbrace{=}_{adding\ j\ and\ k} \mathbb{E}[f(R \cup \{j,k\})]$
2. $F(\bar{x}_{[j\to 1][k\to 0]}) \underbrace{=}_{\substack{adding\ j \\ subtracting\ k}} \mathbb{E}[f((R \cup \{j\})/\{k\})]$
3. $F(\bar{x}_{[j\to 0][k\to 1]}) \underbrace{=}_{\substack{adding\ k \\ subtracting\ j}} \mathbb{E}[f((R \cup \{k\})/\{j\})]$
4. $F(\bar{x}_{[j\to 0][k\to 0]}) \underbrace{=}_{subtracting\ j\ and\ k} \mathbb{E}[f(R/\{j,k\})]$

$$\Longrightarrow \frac{\partial^2 F}{\partial x_k \partial x_j} = \left(\mathbb{E}[f(R \cup \{j,k\})] - \mathbb{E}[f((R \cup \{j\})/\{k\})]\right) - \left(\mathbb{E}[f((R \cup \{k\})/\{j\})] - \mathbb{E}[f(R/\{j,k\})]\right)$$

Let $B \subseteq A \subseteq R$ two sets defined as follows:

- $A = (R \cup \{j\})/\{k\}$
- $B = R/\{j,k\}$

$$\Longrightarrow \frac{\partial^2 F}{\partial x_k \partial x_j} = (\mathbb{E}[f(A \cup \{k\})] - \mathbb{E}[f(A)]) - (\mathbb{E}[f(B \cup \{k\})] - \mathbb{E}[f(B)])$$

$$= \mathbb{E}[f(A \cup \{k\}) - f(A)] - \mathbb{E}[f(B \cup \{k\}) - f(B)]$$

$$= \mathbb{E}\left[\left(f(A \cup \{k\}) - f(A)\right) - \left(f(B \cup \{k\}) - f(B)\right)\right]$$

Since f is sub-modular, we get by definition:

$$f(A \cup \{k\}) - f(A) \underbrace{\leq}_{B \subseteq A} f(B \cup \{k\}) - f(B)$$

$$\Longrightarrow \left(f(A \cup \{k\}) - f(A)\right) - \left(f(B \cup \{k\}) - f(B)\right) \leq 0$$

$$\Longrightarrow \mathbb{E}\left[\left(f(A \cup \{k\}) - f(A)\right) - \left(f(B \cup \{k\}) - f(B)\right)\right] \leq 0$$

$$\Longrightarrow \frac{\partial^2 F}{\partial x_k \partial x_j} \leq 0$$

$$\Longrightarrow \begin{cases} \dfrac{\partial^2 F}{\partial x_k \partial x_j} = 0, & k = j \\ \dfrac{\partial^2 F}{\partial x_k \partial x_j} \leq 0, & k \neq j \end{cases}$$

$$\Longrightarrow \forall j,k \in [N]: \frac{\partial^2 F}{\partial x_k \partial x_j} \leq 0$$

Consider the given function $g(s)$:

$$g(s) = F(\bar{x} + s\bar{d})$$

The first derivative of $g(s)$:

$$\frac{d(g(s))}{ds} \underbrace{=}_{\substack{chain \\ rule}} \sum_{j=1}^{N} \left(\frac{d(x_j + sd_j)}{ds}\right)\left(\frac{\partial F}{\partial(x_j + sd_j)}\right) = \sum_{j=1}^{N} d_j \left(\frac{\partial F}{\partial(x_j + sd_j)}\right)$$

The second derivative of $g(s)$:

$$\frac{d^2(g(s))}{ds^2} \underbrace{=}_{\substack{chain \\ rule}} \sum_{k=1}^{N} \left[\left(\frac{d(x_k + sd_k)}{ds}\right)\left(\frac{\partial F}{\partial(x_k + sd_k)}\right)\left(\sum_{j=1}^{N} d_j \left(\frac{\partial F}{\partial(x_j + sd_j)}\right)\right)\right]$$

$$= \sum_{k=1}^{N} \left[d_k \left(\frac{\partial F}{\partial(x_k + sd_k)}\right)\left(\sum_{j=1}^{N} d_j \left(\frac{\partial F}{\partial(x_j + sd_j)}\right)\right)\right]$$

$$\underbrace{=}_{\substack{k\ is\ constant \\ iterating\ over\ j}} \sum_{k=1}^{N} \left(\sum_{j=1}^{N} \left(d_j d_k \left(\frac{\partial F}{\partial(x_k + sd_k)}\right)\left(\frac{\partial F}{\partial(x_j + sd_j)}\right)\right)\right)$$

$$= \sum_{k=1}^{N} \sum_{j=1}^{N} \left(d_j d_k \frac{\partial^2 F}{\partial(x_k + sd_k)\partial(x_j + sd_j)}\right)$$

Since $d_i \geq 0$ for every $i \in [N]$, both d_j and d_k are non-negative.

Also, we saw that $\frac{\partial^2 F}{\partial x_k \partial x_j} \leq 0$.

Overall, the expression we got for the second derivative of $g(s)$ is non-positive:

$$\frac{d^2(g(s))}{ds^2} = \sum_{k=1}^{N} \sum_{j=1}^{N} \left(d_j d_k \frac{\partial^2 F}{\partial(x_k + sd_k)\partial(x_j + sd_j)}\right) \leq 0$$

$\Longrightarrow$ The function $g(s)$ is concave ■