1 אלגוריתמים חמדניים

1.1 הקדמה

1.1.1 דוגמת חימום.

דוגמא נתונים n מספרים שלמים, $a_1,...,a_n$. האם יש שניים מהמספרים האלה שסכומם 2022.

פתרון א' נעבור על כל הזוגות $1 \leq i \leq j \leq n$ ונבדוק האם $a_i + a_j = 2022$. זמן ריצה $O(\binom{n}{2}) = O(n^2)$.

פתרון ב' בשלב הראשון נמיין את המספרים בסדר עולה: $a_1 \leq ... \leq a_n$ בשלב השני לכל $1 \leq 1i \leq n$ נבדוק האם $2022 - a_i$ נמצא ברשימה הממוינת בעזרת חיפוש בינארי. זמן ריצה $O(nlogn)$.

1.1.2 מהו אלגוריתם חמדני?

אלגוריתם חמדני הוא בדר"כ אלגוריתם איטרטיבי המנסה למקסם רווח מקומי.

1.2 בעיית התרמיל השברי $(Fractional\ Knapsack)$

1.2.1 הצגת הבעיה.

בעיית התרמיל השברי

<u>סיפור מסגרת:</u> גנב נכנס לחנות כאשר ברשותו תרמיל המוגבל במשקל המקסימלי שהוא יכול לשאת. המטרה של הגנב היא למקסם את הערך הכולל של הפריטים הנכנסים לתרמיל. הפריטים ניתנים לחלוקה.

<u>קלט:</u> n - מספר הפריטים בחנות. W - המשקל המירבי של התרמיל. n זוגות של מספרים אי שליליים: $(v_1,w_1),...,(v_n,w_n)$ כאשר v_i הוא הערך הכולל של הפריט הפריט ה-i ו-w_i הוא המשקל הכולל של הפריט ה-i.

<u>פלט:</u> רשימה של n מספרים $0 \leq x_1,...,x_n \leq 1$ כאשר x_i מסמל את החלקיות של הפריט ה-i בתוך התרמיל. רשימה כזו היא פתרון חוקי לבעיה אם מתקיים אילוץ המשקל: $\sum_{i=1}^{n} x_i \cdot w_i \leq W$. בהינתן זה נרצה למקסם את הערך הכולל של הפריטים שבתרמיל.

1.2.2 דוגמא

$n = 3$, $W = 50$, הפריטים הם $(60,10),(100,20),(120,30)$.

פתרון 1 (לא נכון) $x_1 = 1, x_2 = 1, x_3 = 0$ הערך הכולל הוא $1 \cdot 120 + 1 \cdot 100 + 0 \cdot 60 = 220$.

פתרון 2 (נכון) לבחור כמה שיותר מהפריטים בעלי ערך סגולי גבוה יותר. $x_1 = \frac{2}{3}, x_2 = 1, x_3 = 1$. הערך הכולל הוא $\frac{2}{3} \cdot 120 + 1 \cdot 100 + 1 \cdot 60 = 240$.

1.2.3 אלגוריתם חמדני גרסא 1.

מקרה קצה פשוט: אם $\sum_{i=1}^{n} w_i \le W$ במקרה זה נחזיר $x_1 = x_2 = ... = x_n = 1$ מעכשיו נניח כי $\sum_{i=1}^{n} w_i > W$

אלגוריתם 1 אלגוריתם לפתרון התרמיל השברי גרסא 1

---

1. נחשב את הערכים הסגוליים $r_i = \frac{v_i}{w_i}$ לכל $1 \le i \le n$.
2. נמיין את הפריטים לפי הערכים הסגוליים שלהם בסדר יורד $r_1 \ge r_2 \ge ... \ge r_n$. (מעתה נניח כי האינדקסים ממוינים כך).
3. נמצא את האינדקס $0 \le t \le n-1$ כך ש-$\sum_{i=1}^{t} w_i \le W$ וגם $\sum_{i=1}^{t+1} w_i > W$ ($t+1$ הוא הפריט הראשון שאי אפשר להכניס בשלמותו).
4. נחזיר $x_1 = x_2 = ... = x_t = 1$, $x_{t+1} = \frac{W - \sum_{i=1}^{t} w_i}{w_{t+1}}$, $x_{t+2} = ... = x_n = 0$.

<u>זמן ריצה:</u> $O(nlogn)$.

1.2.4 הוכחת חוקיות ואופטימליות לגרסא 1.

טענה הפתרון שהחזרנו הוא פתרון חוקי ואופטימלי לבעיה.

הוכחת חוקיות צריך להוכיח כי $0 \le x_i \le 1$ לכל $i \in [n]$, ובנוסף כי אילוץ המשקל מתקיים: $\sum_{i=1}^{n} x_i \cdot w_i \le W$. מתקיים כי $x_1 = ... = x_t = 1$ וגם $x_{t+1} = ... = x_n = 0$.

$$0 = \frac{\sum_{i=1}^{t} w_i - \sum_{i=1}^{t} w_i}{w_{t+1}} \le x_{t+1} = \frac{W - \sum_{i=1}^{t} w_i}{w_{t+1}} \le \frac{\sum_{i=1}^{t+1} w_i - \sum_{i=1}^{t} w_i}{w_{t+1}} = 1$$

אילוץ המשקל:

$$\begin{aligned}
\sum_{i=1}^{n} x_i \cdot w_i &= \sum_{i=1}^{t} x_i w_i + x_{t+1} w_{t+1} + \sum_{i=t+2}^{n} x_i w_i \\
&= \sum_{i=1}^{t} w_i + \frac{W - \sum_{i=1}^{t} w_i}{w_{t+1}} \cdot w_{t+1} + 0 \\
&= \sum_{i=1}^{t} w_i + W - \sum_{i=1}^{t} w_i \\
&= W
\end{aligned}$$

הערה: הפתרון החמדני (שאנו מחזירים) ממלא את התרמיל כולו.

הוכחת אופטימליות יהי $y_1, y_2, ..., y_n$ פתרון חוקי כלשהו לבעיה. נראה כי $\sum_{i=1}^{n} x_i \cdot v_i \geq \sum_{i=1}^{n} y_i \cdot v_i$.

$$\begin{aligned}
\sum_{i=1}^{n} x_i \cdot v_i - \sum_{i=1}^{n} y_i \cdot v_i &= \sum_{i=1}^{n} (x_i - y_i) v_i \\
&= \sum_{i=1}^{n} (x_i - y_i) w_i r_i \\
&= \sum_{i=1}^{t} (x_i - y_i) w_i r_i + (x_{t+1} - y_{t+1}) w_{t+1} r_{t+1} + \sum_{i=t+2}^{n} \underbrace{(x_i - y_i)}_{<0} w_i r_i \\
&\geq \sum_{i=1}^{t} (x_i - y_i) w_i r_{t+1} + (x_{t+1} - y_{t+1}) w_{t+1} r_{t+1} + \sum_{i=t+2}^{n} (x_i - y_i) w_i r_{t+1} \\
&= r_{t+1} \left(\sum_{i=1}^{t} (x_i - y_i) w_i + (x_{t+1} - y_{t+1}) w_{t+1} + \sum_{i=t+2}^{n} (x_i - y_i) w_i \right) \\
&= r_{t+1} \cdot \left(\sum_{i=1}^{n} x_i w_i - \sum_{i=1}^{n} y_i w_i \right) \\
&= r_{t+1} \cdot \left(W - \sum_{i=1}^{n} y_i w_i \right) \\
&\geq 0
\end{aligned}$$

1.2.5 אלגוריתם חמדני גרסא 2.

זה אותו אלגוריתם, שמחזיר את אותו פתרון אבל מנוסח בדרך קצת שונה.

אלגוריתם 2 אלגוריתם לפתרון התרמיל השברי גרסא 2

---

1. עיבוד מידע מקדים: נחשב את הערכים הסגוליים $r_i = \frac{v_i}{w_i}$ של כל הפריטים ונמיין אותם בסדר יורד.
2. אתחול: התרמיל ריק.
3. איטרציה: נעבור על הפריטים לפי הסדר $i = 1, 2, .., n$. בכל שלב "נכניס כמה שאפשר", כלומר נכניס את החוקיות המירבית של הפריט ה-i שאינה פוגעת בחוקיות הפתרון. (הכלל החמדני)
4. סיום: נעצור בסוף n האיטרציות (שקול לעצור כשהתרמיל מלא).

אנחנו דנים בבעיות אופטימיזציה. כל בעיה מגדירה מרחב פתרונות חוקיים משלה. במקרה הזה המרחב הזה הוא:

$$\mathcal{S} = \left\{ (x_1, ..., x_n) \in \mathbb{R}^n : 0 \leq x_i \leq 1 \forall i \in [n] \wedge \sum_{i=1}^{n} x_i \cdot w_i \leq W \right\}$$

ועל המרחב הזה אנו מגדירים פונקציית ערך $f : \mathcal{S} \to \mathbb{R}_+$ שאנו רוצים למצוא נקודת מקסימום שלה. במקרה שלנו $f(x) = \sum_{i=1}^{n} x_i v_i$.

אנו צריכים לנמק מדוע בחירה על פי העיקרון החמדני בכל צעד תביא אותנו לפתרון אופטימלי. נעשה זאת באמצעות למת החלפה - נראה שאם בשלב מסויים האלגוריתם לא בוחר בחירה חמדנית, ניתן להחליף את בחירתו בבחירה חמדנית ולהרוויח (לפחות לא להפסיד).

1.2.6 הוכחת חוקיות ואופטמליות גרסא 2.

(הנחה: $r_1 > r_2 > ... > r_n > 0$ - ניתן להתעלם מהפריטים שערכם הסגולי קטן שווה 0, וניתן לאחד פריטים בעלי אותו ערך סגולי).

למת החלפה יהי $y = (y_1, y_2, ..., y_n)$ פתרון חוקי לבעיה, כך שקיים $1 \leq j \leq n$ עבורו $y_j < 1$ ו-$\sum_{i=1}^{j} y_i w_i < W$ אזי קיים פתרון חוקי $z = (z_1, z_2, ..., z_n)$ כך ש-$z_1 = y_1, ..., z_{j-1} = y_{j-1}$, $z_j > y_j$ וכך ש-$\sum_{i=1}^{n} z_i v_i > \sum_{i=1}^{n} y_i v_i$.

הוכחת למת ההחלפה נבדיל בין שתי אפשרויות: (1) $\sum_{i=1}^{n} y_i w_i < W$. (2) $\sum_{i=1}^{n} y_i w_i = W$. נטפל באפשרות השנייה (הראשונה דומה אך קלה יותר). מעכשיו נניח כי $\sum_{i=1}^{n} y_i w_i = W$. על פי הנחת הלמה, $\sum_{i=1}^{j} y_i w_i < W$. לכן, קיים $j \leq k \leq n$ כך ש- $y_k > 0$. נסמן $d_j = (1 - y_j)w_j$. נסמן $d_k = y_k w_k$. נסמן $0 < d = min\{d_j, d_k\}$ (d זה המשקל שנזיז). נגדיר פתרון z באופן הבא:

$z_i = y_i$ לכל $i \neq j, k$. $z_j = y_j + \frac{d}{w_j}$, $z_k = y_k - \frac{d}{w_k}$. (זו ההחלפה).

נשים לב כי $z_1 = y_1, ..., z_{j-1} = y_{j-1}$. נבדוק כי z הוא פתרון חוקי. מתקיים כי $0 \leq z_i = y_i \leq 1$ לכל $i \neq j, k$ מתקיים:

$$0 \leq y_j < z_j = y_j + \frac{d}{w_k} \leq y_j + \frac{d_j}{w_k} = y_j + (1 - y_j) = 1$$

וכנ"ל לגבי z_k.

נראה כי z מקיים את אילוץ המשקל.

$$\begin{aligned}
\sum_{i=1}^{n} z_i w_i - \sum_{i=1}^{n} y_i w_i &= \sum_{i=1}^{n} (z_i - y_i) w_i \\
&= (z_j - y_j) w_j + (z_k - y_k) w_k \\
&= \frac{d}{w_j} \cdot w_j - \frac{d}{w_k} \cdot w_k \\
&= 0
\end{aligned}$$

לכן $\sum_{i=1}^{n} y_i w_i = \sum_{i=1}^{n} z_i w_i = W$ ומכאן ש-z פתרון חוקי.

כעת נראה כי $\sum_{i=1}^{n} z_i v_i > \sum_{i=1}^{n} y_i v_i$.

$$\begin{aligned}
\sum_{i=1}^{n} z_i v_i - \sum_{i=1}^{n} y_i v_i &= \sum_{i=1}^{n} (z_i - y_i) v_i \\
&= (z_j - y_j) v_j + (z_k - y_k) v_k \\
&= \frac{d}{w_j} v_j - \frac{d}{w_k} v_k \\
&= d \cdot r_j - d \cdot r_k \\
&= d(r_j - r_k) \\
&> 0
\end{aligned}$$

וסיימנו את הוכחת הלמה.

הוכחת חוקיות מכיוון שעל פי דרך פעולתו של האלגוריתם החמדן הוא שומר על החוקיות של הפתרון בכל איטרציה, גם בסיום הריצה של האלגוריתם הפתרון המוחזר הוא פתרון חוקי.

הוכחת אופטימליות נסמן ב-$x = (x_1, ..., x_n)$ את הפתרון החמדן שהאלגוריתם מחזיר. נראה כי אם פתרון חוקי אחר y שונה מ-x, אז y אינו אופטימלי. מכיוון שקיים פתרון אופטימלי לבעיה (קומפקטיות) נסיק ש-x הוא הפתרון האופטימלי. יהי y פתרון חוקי ששונה מ-x. יהי $1 \leq j \leq n$ האינדקס הראשון עבורו $y_j \neq x_j$. על פי דרך פעולתו של האלגוריתם החמדן, x_j היא החלקיות המירבית של הפריט ה-j שניתן להכניס לתרמיל בלי לפגוע בחוקיות הפתרון. לכן $y_j \neq x_j$ גורר ש-$y_j < x_j$. נראה כי הפתרון y מקיים את התנאים של למת ההחלפה ולכן אינו אופטימלי. אכן $y_j < x_j \leq 1$ וגם $\sum_{i=1}^{j} y_i w_i < \sum_{i=1}^{j} x_i w_i \leq W$. מש"ל.

1.3 בעיית שיבוץ משימות

1.3.1 הצגת הבעיה

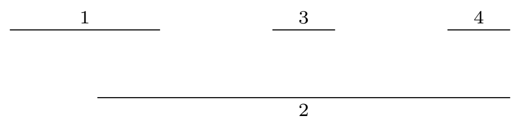
בעיית שיבוץ משימות

<u>סיפור מסגרת</u>: נתונות n משימות, כך שלא ניתן לבצע שתי משימות בו זמנית. אנו רוצים למצוא תת קבוצה גדולה כמה שיותר של משימות שכן ניתן לבצען.

<u>קלט</u>: n - מספר המשימות. ו-n קטעים סגורים הנתונים על ידי נקודות ההתחלה והסיות שלהם: $[s_1, f_1], ..., [s_n, f_n]$.

<u>פלט</u>: הפתרונות החוקיים הם תתי קבוצה $S \subseteq [n]$ כך שהקטעים עם אינדקסים ב-S זרים זה לזה. אנו רוצים למצוא את הקבוצה בעלת הגודל המקסימלי במרחב.

1.3.2 דוגמא

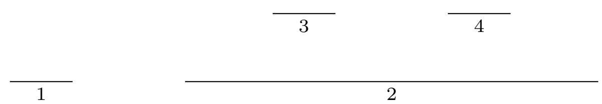


$n=5$. הפתרון האופטימלי הוא $S^*=\{1,3,4\}$. יכולים להיות קלטים עם יותר מפתרון אופטימלי אחד.

1.3.3 רעיונות לאלגוריתמים חמדניים

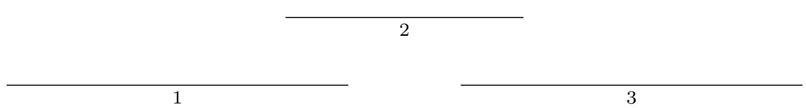
נחפש אלגוריתם איטרטיבי שבכל איטרציה מוסיף קטע חדש לפתרון באמצעות כלל חמדני.

1. הכנסת הקטע שזמן התחלתו מינימלי שלא מתנגש עם הקטעים שכבר בחרנו. דוגמא נגדית:



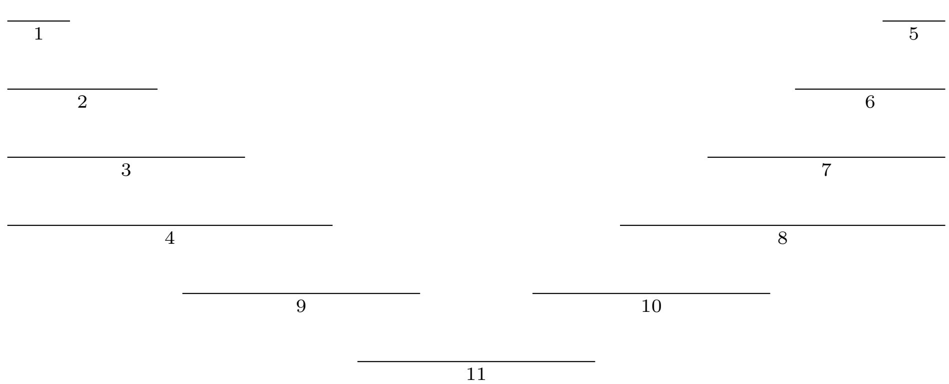
הפתרון החמדני $\{1,2\}$, הפתרון האופטימלי $S^*=\{1,3,4\}$.

2. הכנסת הקטע בעל האורך המינימלי שלא מתנגש עם הקטעים שכבר בחרנו. דוגמא נגדית:



הפתרון החמדני $\{2\}$, הפתרון האופטימלי $S^*=\{1,3\}$.

3. הכנסת הקטע שמתנגש עם כמה שפחות מהקטעים שנשארו ולא מתנגש עם הקטעים שכבר בחרנו. דוגמא נגדית:



הפתרון החמדני $\{11,1,5\}$ אופטימלי $S^*=\{1,5,9,10\}$.

4. קטע עם זמן סיום מינימלי (אין דוגמא נגדית כי זה עובד).

1.3.4 האלגוריתם החמדן

אלגוריתם 3 אלגוריתם חמדן לפתרון בעיית שיבוץ משימות

1. עיבוד מידע מקדים: נמיין את הקטעים בסדר עולה לפי זמני הסיום שלהם. (מעתה נניח כי $f_1 \leq ... \leq f_n$).
2. אתחול: $G = \emptyset$, $A = [n]$. (G זה המשתנה המכיל את הפתרון החמדן, A זה מה שעדיין ניתן לבחור).
3. איטרציה: בכל שלב נוסיף ל-G את האינדקס המינימלי מ-A. נמחק מ-A את כל הקטעים החותכים את הקטע שכרגע הוספנו ל-G.
4. סיום: כאשר $A = \emptyset$ נעצור ונחזיר את G.

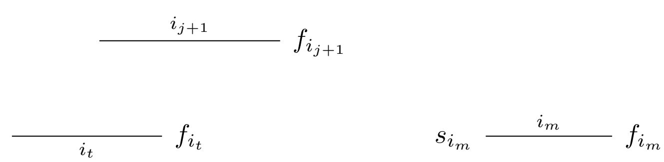
זמן ריצה: מיון: $O(nlogn)$. אתחול וסיום: $O(1)$. איטרציה: יש לכל היותר n איטרציות (A מתחיל עם n קטעים ובכל איטרציה מוציאים לכל הפחות קטע אחד מ-A וכל איטרציה עולה לכל היותר $O(n)$) אז בסה"כ $O(n^2)$ - אבל ניתן לעשות עם עוד קצת מחשבה ב-$O(nlogn)$.

1.3.5 הוכחת חוקיות ואופטימליות

נרצה להוכיח שהאלגוריתם מחזיר פתרון אופטימלי באמצעות למת החלפה.

למת החלפה יהי $S = \{i_1, ..., i_k\}$ כך ש-$i_1 < i_2 < ... < i_k$ פתרון חוקי לבעיה. יהי $0 \leq j \leq k-1$. יהי t האינדקס המינימלי של קטע שלא חותך את הקטעים $i_1, ..., i_j$. אזי גם הפתרון $S_1 = \{i_1, ..., i_j, t, i_{j+2}, ..., i_k\}$ גם הוא פתרון חוקי ($|S_1| = |S|$).

הוכחת למת ההחלפה צריכים להוכיח שכל הקטעים ב-S_1 זרים זה לזה. מכיוון שכל הקטעים ב-S_1 פרט ל-t שייכים לפתרון חוקי S, נתון כי הם זרים זה לזה. נשאר לבדוק כי t זר לכל הקטעים האלה. על פי בחירת t, הוא זר לקטעים $i_1, ..., i_j$. נשאר לבדוק כי t זר לקטעים $i_{j+2}, ..., i_k$. יהי $j+2 \leq m \leq k$ נוודא כי t זר לקטע i_m. אבחנה מרכזית: יהיו $[s, f]$, $[s', f']$ שני קטעים כלשהם כך ש-$f' \geq f$, אזי הם זרים אמ"ם $s' > f$. מכיוון שהקטעים i_{j+1} ו-i_m שייכים לפתרון חוקי S, הם קטעים זרים. מכיוון ש-$i_m > i_{j+1}$ מתקיים $f_{i_m} \geq f_{i_{j+1}}$ ולכן לפי האבחנה $s_{i_m} > f_{i_{j+1}}$. נזכור כי t הוא האינדקס המינימלי של קטע שלא חותך את הקטעים $i_1, ..., i_j$ ואילו i_{j+1} הוא קטע שלא חותך את הקטעים האלה. לכן $t \leq i_{j+1}$ ולכן $f_t \leq f_{i_{j+1}} < s_{i_m}$. (ראה ציור למטה) ולכן מהאבחנה בכיוון השני t ו-i_m הם קטעים זרים.



טענה הפתרון החמדן הוא פתרון חוקי ואופטימלי.

הוכחת חוקיות על פי דרך פעולתו של האלגוריתם החמדן, הקבוצה G היא קבוצה של קטעים זרים בסוף כל איטרציה של האלגוריתם. ולכן גם בסיום הריצה שלו.

הוכחת אופטימליות יהי $G = \{g_1, ..., g_r\}$ הפתרון של האלגוריתם החמדן. בשלב הראשון נראה כי לכל $1 \leq l \leq r$ קיים פתרון אופטימלי O_l המתחיל ב-$g_1, ..., g_l$. בפרט נקבל, עבור $l = r$ כי קיים פתרון אופטימלי O_r המתחיל בכל הקטעים $g_1, ...g_r$ של הפתרון החמדני. נעשה

זאת באינדוקציה על l.

<u>בסיס</u> $(l=1)$: יהי $O=\{i_1,...,i_m\}$ פתרון אופטימלי כלשהו (קיים כזה כי יש קבוצה סופית של פתרונות). נפעיל את למת ההחלפה. על הפתרון O עבור $j=0$. לפי הלמה, ניתן להחליף את i_1 בקטע t המסתיים הכי מוקדם, כלומר ב-$t=1$. נשים לב כי גם $g_1=1$. אז קיבלנו פתרון חוקי $O_1=g_1,i_2,...,i_m$. נשים לב כי $|O|=|O_1|$ ולכן גם O_1 אופטימלי.

<u>צעד</u> (נניח שהטענה נכונה עבור $1\leq l<r$ ונוכיח עבור $l+1$): יהי $O_l=\{g_1,...,g_l,i_{l+1},...,i_m\}$ פתרון אופטימלי המתחיל ב-l הקטעים הראשונים של הפתרון החמדן. נפעיל את למת ההחלפה על הפתרון O_l עם $j=l$. לפי הלמה ניתן להחליף את i_{l+1} בקטע t המסתיים הכי מוקדם מכל הקטעים שלא חותכים את $g_1,...,g_l$. לפי דרך פעולתו של האלגוריתם החמדן, $t=g_{l+1}$ ולכן קיבלנו פתרון חוקי ואופטימלי $O_{l+1}=\{g_1,...,g_{l+1},i_{l+1},...,i_m\}$ וזה מסיים את צעד האינדוקציה.

מטענת האינדוקציה, קיים פתרון אופטימלי O_r המתחיל בכל r הקטעים של הפתרון החמדן G. $O_r=\{g_1,...,g_r,i_{r+1},...,i_m\}$. נראה כי O_r מכיל <u>רק</u> את הקטעים של G ולכן $G=O_r$ ובכך נוודא כי G אופטימלי. נניח בשלילה כי קיים קטע $i\in O_r\setminus G$ כלומר $i\notin G$ וגם i זר לכל הקטעים ב-G. נשים לב כי על פי דרך פעולתו של האלגוריתם החמדן, הקטע i לא יוצא מקבוצה A לאורך כך האיטרציות של האלגוריתם. לכן, בסיום הריצה של האלגוריתם, $i\in A$ כלומר $A\neq\emptyset$ בסתירה לכלל העצירה של האלגוריתם.

1.4 בעיית מציאת קבוצת ווקטורים בת"ל בעלת משקל מקסימלי

1.4.1 הצגת הבעיה

קבוצה בת"ל של ווקטורים בעלת משקל מקסימלי

<u>קלט:</u> קבוצה של n ווקטורים $F=\{v_1,...,v_n\}$ במרחב ווקטורי V, ופונקציית משקל $\mu:F\to\mathbb{R}_+$.

<u>פלט:</u> תת-קבוצה $S\subseteq[n]$ כך שהווקטורים עם אינדקסים ב-S מהווים קבוצה בת"ל וכך ש-$\mu(S)$ מירבי בתנאי זה (כאשר $\mu(S)=\sum_{i\in S}\mu(v_i)$).

1.4.2 דוגמא

$V=\mathbb{R}^4$, $n=5$.

$$v_1=\begin{bmatrix}1\\2\\3\\4\end{bmatrix},v_2=\begin{bmatrix}\sqrt{2}\\e\\\pi\\\sqrt{3}\end{bmatrix},v_3=\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix},v_4=\begin{bmatrix}3\\0\\5\\2\end{bmatrix},v_5=\begin{bmatrix}3\\5\\7\\11\end{bmatrix}$$
$$\mu(v_1)=7,\mu(v_2)=6,\mu(v_3)=5,\mu(v_4)=4,\mu(v_5)=3,$$

מתקיים $v_4=v_1+2v_3$. אם בכל שלב נוסיף ווקטור בעל משקל מקסימלי שלא מייצר תלות עם מה שבחרנו נקבל $G=\{1,2,3,5\}$.

1.4.3 האלגוריתם החמדן

אלגוריתם 4 אלגוריתם חמדן לפתרון בעיית מציאת קבוצת ווקטורים בת"ל בעלת משקל מקסימלי

---

1. עיבוד מידע מקדים: נמיין את הווקטורים בסדר יורד לפי משקלם. (מעכשיו נניח כי $\mu(v_1) \geq ... \geq \mu(v_n)$).
2. אתחול: $G = \emptyset$, $A = [n]$.
3. איטרציה: בכל שלב נעביר מ-A ל-G את האינדקס המינימלי מ-A. נמחק מ-A את כל הווקטורים התלויים ליניארית בווקטורים ב-G.
4. עצירה: כאשר $A = \emptyset$ נעצור ונחזיר את G.

1.4.4 זמן ריצה

מיון: $O(nlogn)$. אתחול ועצירה $O(1)$. באשר לשלב האיטרציה:

שאלה: איך פותרים את הבעיה הבאה: נתינים k ווקטורים $g_1, ..., g_k \in V$ ווקטור $a \in V$. האם הווקטורים $a, g_1, .., g_k$ תלויים ליניארית?

תשובה (העשרה): פותרים את המערכת:

$$\begin{bmatrix} | & & | \\ g_1 & \cdots & g_k \\ | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{dim(V)} \end{bmatrix} = \begin{bmatrix} | \\ a \\ | \end{bmatrix}$$

זמן הריצה של זה הוא $O(dim(V) \cdot n^2)$. בכל איטרציה עוברים על כל הווקטורים ב-A ופותרים מערכת כזאת. סה"כ זמן הריצה לאיטרציה הוא $O(dim(V) \cdot n^3)$. יש n איטרציות ולכן זמן הריצה הכולל הוא $O(dim(V) \cdot n^4)$. (אבל אפשר בקלות גם להגיע ל-$O(dim(V) \cdot n^3)$).

1.4.5 הוכחת חוקיות ואופטימליות

למת החלפה החלפה לא נכונה יהי $S = \{i_1, ..., i_k\}$ פתרון חוקי, ויהי $0 \leq j \leq k-1$ ויהי t האינדקס המינימלי של ווקטור שלא תלוי ליניארית ב-$i_1, ..., i_l$ אזי $S_1 = \{i_1, ..., i_j, t, i_{j+2}, ..., i_k\}$ פתרון חוקי.

למת החלפה נכונה אבל מסובכת מדי יהי $S = \{i_1, ..., i_k\}$ פתרון חוקי, ויהי $0 \leq j \leq k-1$ ויהי t האינדקס המינימלי של ווקטור שלא תלוי ליניארית ב-$i_1, ..., i_l$ אזי קיים פתרון חוקי $S_1 = \{i_1, ..., i_j, t, r_{j+2}, ..., r_k\}$.

למת החלפה (מטרואידלית) תהיינה X, Y שתי קבוצות סופיות בת"ל (בנפרד) של ווקטורים במרחב ווקטורי V, כך ש-$|X| < |Y|$. אזי קיים $y \in Y \setminus X$ כך ש- $X \cup \{y\}$ בת"ל.

תזכורת לפני הוכחת הלמה בסיס של מרחב ווקטורי היא קבוצה בת"ל של ווקטורים הפורשים את המרחב. בסיס הוא קבוצה בת"ל בגדול מקסימלי בתוך המרחב. כל הבסיסים במרחב הם באותו גודל שנקרא המימד של המרחב.

הוכחת למת ההחלפה יהי U המרחב הווקטורי הנפרש על ידי X. X הוא בסיס של U ולכן $dim(U) = |X|$. מכיוון ש- $|Y| > |X|$, ו-Y בת"ל, קיים $y \in Y$ כך ש- $y \notin U$. באופן שקול, y אינו צירוף ליניארי של ווקטורים מ-X, ובפרט $y \notin X$ ו-$X \cup \{y\}$ בת"ל.

טענה האלגוריתם החמדן מחזיר פתרון חוקי ואופטימלי לבעיה.

הוכחת חוקיות חוקיות הפתרון ברורה מהעדכון של קבוצה A המתבצע בכל איטרציה של האלגוריתם החמדן.

הוכחת אופטימליות יהי G הפתרון של האלגוריתם החמדן, ויהי S^* פתרון אופטימלי. נראה כי $\mu(G) = \mu(S^*)$ ומכך נסיק כי G פתרון אופטימלי. נעשה זאת בשני שלבים.

<u>השלב הראשון</u> נוכיח כי $|G| = |S^*|$. אם בשלילה $|G| > |S^*|$ אז לפי למת ההחלפה קיים $s \in S^* \setminus G$ כך ש- $G \cup \{s\}$ קבוצה בת"ל. במצב כזה הווקטור s לא יוצא מהקבוצה A באף איטרציה של האלגוריתם החמדן. והאלגוריתם החמדן עוצר כאשר $s \in A$, כלומר $A \neq \emptyset$ בסתירה. אם בשלילה $|G| < |S^*|$ אז לפי למת ההחלפה קיים $g \in G \setminus S^*$ כך ש-$S^* \cup \{g\}$ הוא פתרון חוקי שמשקלו $\mu(S^*) + \mu(g) > \mu(S^*)$, בסתירה לאופטימליות של $|S^*|$.

<u>השלב השני</u> נוכיח כי $\mu(G) = \mu(S^*)$. נסמן $k = |S^*| = |G|$. נכתוב $G = \{g_1, ..., g_k\}$ ו-$S^* = \{s_1, ..., s_k\}$ (האינדקסים של שתי הקבוצות הוא בסדר יורד. נניח בשלילה כי $\mu(S^*) > \mu(G)$ ונגיע לסתירה. אז $\sum_{i=1}^k \mu(s_i) > \sum_{i=1}^k \mu(g_i)$, ולכן קיים $1 \leq j \leq k$ כך ש-$\mu(s_j) > \mu(g_j)$. נגדיר $X = \{g_1, ..., g_{j-1}\}$, $Y = \{s_1, ..., s_j\}$. לפי למת ההחלפה קיים $y \in Y \setminus X$ כך ש- $X \cup \{y\}$ בת"ל. כלומר קיים $1 \leq t \leq j$ כך ש-s_t אינו תלוי ליניארית ב-$\underbrace{g_1, ..., g_{j-1}}_{X}$. לכן, s_t ניתן לבחירה באיטרציה ה-j של האלגוריתם החמדן. בנוסף מתקיים כי $\mu(s_t) \geq \mu(s_j) \geq \mu(g_j)$ כי הם מסודרים בסדר עולה של המשקלים. לכן האלגוריתם החמדן לא בחר בחירה חמדנית באיטרציה ה-j שלו, <u>בסתירה</u> לדרך פעולתו.