אלגוריתמים מתקדמים – סיכום שיעור 12.1.2005

הגדרה: שיכון

שיכון של מרחב מטרי (X,d_x) במרחב מטרי (Y,d_y) היא פונקציה $f:X\rightarrow Y$.

הגדרה: שיכון איזומטרי

שיכון המשמר מרחקים בצורה מלאה – לכל $u,v\in X$ $d_y(f(u),f(v))=d_x(u,v)$.

נשים לב כי מרחב נורמי הינו מרחב מטרי בו $d(x,y)=\|x-y\|$ (כי מתקיים אי שוויון משולש).

נגדיר $(R^n,l_p)=l_p^n$ כמרחב מטרי בו הנקודות שיכות ל R^n ונורמת l_p מוגדרת כ

$$\|X\|_p=\left(\sum_{i=1}^{n}|X_i|\right)^{\frac{1}{p}}.$$

$p=1$ - נורמת מנהטן.

$p=2$ - נורמה אוקלידית.

$p=\infty$ - נורמת אינסוף: $\|X\|_\infty=\max_i|X_i|$

קונבנציה

נסמן ב L_p מרחב באינסוף קוארדינטות ובעל פונקציית מרחק l_p. (נוח להגיד שמתייחסים לנורמה הנ"ל באיזשהו מימד).

משפט (Dvoretsky):

L_2 משתכן איזומטרית ב L_p עבור $p\geq 1$.

משפט (Frechet):

כל מרחב מטרי X בגודל n ניתן לשיכון איזומטרי ב l_∞^n

<u>הוכחה</u>: כל קוארדינטה במרחב ה-n מימדי נייצג באמצעות נקודה במרחב המקורי שלנו (כלומר הקוארדינטה הראשונה תסומן כ Z_1 השנייה כ Z_2 וכך הלאה כאשר $Z_i\in X$).

נגדיר את פונקציית השיכון f להיות $y=(f_{z1}(x),f_{z2}(x)...,f_{zn}(x))$ - כלומר השיכון יתבצע על ידי הפעלת פונקציה f_{zi} עבור כל קוארדינטה i של מרחב Y בצורה הבא: $f_{zi}(x)=d(x,x_i)$ כלומר עבור x מסויים, הערך של הקוארדינטה i בשיכון של x ב Y יהיה $d(x,x_i)$.

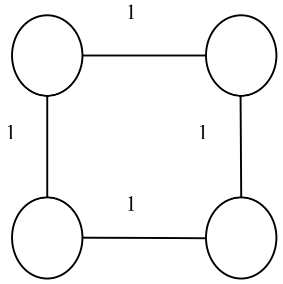
<u>טענה</u>: $\|f(x)-f(y)\|_\infty=d(x,y)$

<u>הוכחה</u>: $\|f(x)-f(y)\|_\infty=\max_i|f_{zi}(x)-f_{zi}(y)|=\max_z|d(x,z)-d(y,z)|$.

לפי אי שוויון משולש $d(x,z)-d(y,z)\leq d(x,y)$. כלומר $\|f(x)-f(y)\|_\infty\leq d(x,y$.

מבניית המרחב עבור z=x (וגם עבור z=y) $|d(x,z)-d(y,z)|=d(x,y)$ ולכן

$\|f(x)-f(y)\|=d(x,y)$.



הגדרה: מרחב מטרי על גרף (לא מכוון)

מרחב מטרי בו הקדקדים הם נקודות ו (A,B) הינו אורך המסלול הקצר ביותר בין A ל B.

טענה:

C_4 לא ניתן לשיכון איזומטרי במרחב אוקלידי:

<u>הוכחה</u>:

במרחב הנ"ל אי שוויון משולש מתקיים כשוויון $d(A,C)=d(A,B)+d(A,C)$. במרחב האוקלידי משמעו שכל שלוש נקודות נמצאות על ישר אחד – לא ניתן לשכן כך 4 נקודות ולכן סתירה.

הגדרה: עיוות (distortion)

לשיכון של X ל Y יש עיוות α אם קיים קבוע c>0 כך שלכל $u,v\in X$ מתקיים $d_y(f(u),f(v))\le c\cdot\alpha\cdot d(x,y)$. הגורם c מבטל scaling (אם אפשר לשכן מרחב X ב Y אזי אם נכפיל את כל הקוארדינטות בקבוע עדיין השיכון אפשרי).

הגדרה: הרחבה (expansion) של שיכון

$$\text{expansion}(f)=\max\frac{d_y(f(x),f(y))}{d_x(x,y)}\le\alpha\cdot c$$

הגדרה: כיווץ (contruction) של שיכון

$$contraction(f)=\max\frac{d_x(x,y)}{d_y(f(x),f(y)}=\frac{1}{c}$$

העיוות הינו המכפלה של הכיווץ בהרחבה: α=expansion(f)*contraction(f).

משפט (Bourgain):

כל מרחב מטרי על n נקודות ניתן לשיכון ב L_2 בעיוות $O(\log n)$.

בהוכחה של המשפט מראים שניתן לבצע שיכון במימד $d=O(\log^2 n)$.

הרחבה למשפט מראה שניתן לבצע שיכון ל L_p בעיוות של $O(\frac{\lceil\log n\rceil}{p})$.

מאמר של Linial,London, Rabinovich מראה שלא ניתן לבצע שיכון ל L_p בעיוות קטן מ $\Omega(\frac{\lceil\log n\rceil}{p})$.

למה (Johnson - Lindenstrauss):

כל קבוצה של n נקודות ב L_2 ניתנת לשיכון ב L_2^d לכל $\varepsilon>0$ כאשר $d=O(\frac{\log^2 n}{\varepsilon^2})$ בעיוות $1+\varepsilon$.

הערה: זה אינו אפשרי ב L_1.

משפט (LLR - Linial,London, Rabinovich):

נגדיר $c_2(x)$ העיוות המינימלי של X ב L_2. אזי $c_2(x)$ ניתן לחישוב בזמן פולינומיאלי בגודל המרחב X. (כלומר קיים אלגוריתם לחישוב השיכון בעל העיוות המינימלי).

הוכחה: נתון X בגודל n. נניח $f : X \to L_2$. ניתן להניח בה"ע שהשיכון ל L_2^{n-1}.

נסמן $\alpha = c_2(x)$ ונניח ש f שיכון מרחיב, אזי מתקיים:

$d(x,y) \le \|f(x) - f(y)\| \le \alpha \cdot d(x,y)$. נעלה בריבוע ונקבל

$d(x,y)^2 \le \|f(x) - f(y)\|^2 \le \alpha^2 \cdot d(x,y)^2$.

מכאן ש:

$$\|f(x) - f(y)\|^2 =< f(x) - f(y), f(x) - f(y) >=< f(x), f(x) > + < f(y), f(y) > -2 < f(x), f(y) >$$

לכל $x \in X$ נגדיר ווקטור u_x n-1 מימדי וקיבלנו את בעיית ה SDP הבאה:

$$\begin{aligned}
&\min \alpha \\
&s.t \\
&u_x u_x + u_y u_y - 2 \cdot u_y u_x \ge d(x,y)^2 \\
&u_x u_x + u_y u_y - 2 \cdot u_y u_x \le \alpha^2 d(x,y)^2
\end{aligned}$$

בעיה זו ניתן לפתור בזמן פולינומיאלי.

בעיה:

נתונה קבוצה X של n נקודות ב L_2^d כאשר d קבוע קטן. רוצים לחשב את הקוטר של X (מציאת נקודות $u, v \in X$ כך ש $\max_{x,y\in X} \|x - y\| = d(u,v)$.

אלגוריתם טריוויאלי הוא ב $O(n^2 d)$ - ריצה על כל הנקודות.

אם נסתכל על הבעיה ב $L_\infty^{d'}$ אזי

$$\max_{x,y\in X} \|x - y\|_\infty = \max_{x,y} \max_i |x_i - y_i| = \max_i \max_{x\neq y} |x_i - y_i| = \max \left|\max_x x_i - \min_y y_i\right|$$

חישוב זה ניתן לביצוע ב $O(nd')$.

למה (תוכח בשיעור הבא):

L_1^d ניתן לשיכון ב $L_\infty^{d'}$ כאשר $d' = 2^d$.