אלגברה לינארית - סיכום תרגול מס' 12

1 פירוק SVD

תהי $A \in M_n(\mathbb{R})$ מטריצה. המטריצה A לכסינה אורתוגונלית אם קיימת מטריצה אורתוגונלית $O \in M_n(\mathbb{R})$ ומטריצה אלכסונית $D \in M_n(\mathbb{R})$ כך ש־$D = O^tAO$ או באופן שקול $A = ODO^t$. העמודות של O מהוות אורתונורמלי של וקטורים עצמיים של A ביחס למכפלה הפנימית הסטנדרטית על $\mathbb{R}^n$. מה המשמעות הגיאומטרית של לכסון אורתוגונלי? אם נסמן ב־$T_A : \mathbb{R}^n_{\text{col}} \to \mathbb{R}^n_{\text{col}}$ את האופרטור שמתאים לכפל ב־A משמאל ($T_A(x) = Ax$) אז אפשר לחשוב על לכסון בשני אופנים:

1. אם נסמן ב־$\mathcal{B} = (v_1, \dots, v_n)$ את הבסיס האורתונורמלי של $\mathbb{R}^n$ שמורכב מוקטורים עצמיים של T_A אז $[T_A]^{\mathcal{B}}_{\mathcal{B}} = D$. כלומר, קיים בסיס אורתונורמלי אשר ביחס אליו T_A מיוצגת באופן הכי פשוט שאפשר - מטריצה אלכסונית. כל וקטור $x \in \mathbb{R}^n$ מתפרק באופן יחיד כסכום

$$x = \sum_{i=1}^{n} \langle x, v_i \rangle v_i$$

ופעולת T_A על הוקטור הנ"ל נתונה על־ידי ניפוח של כל "ציר" v_i בנפרד בפקטור λ_i:

$$T_A(x) = T_A\left(\sum_{i=1}^{n} \langle x, v_i \rangle v_i\right) = \sum_{i=1}^{n} \langle x, v_i \rangle T_A(v_i) = \sum_{i=1}^{n} \lambda_i \langle x, v_i \rangle v_i.$$

2. מבלי לבצע שינוי בסיס, מתקיים $T_A = T_O \circ T_D \circ T_{O^t}$:

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "T_{O^t}"] \arrow[rrr, bend left, "T_A"] & \mathbb{R}^n \arrow[r, "T_D"] & \mathbb{R}^n \arrow[r, "T_O"] & \mathbb{R}^n \end{tikzcd}$$

כלומר, פעולת T_A על $\mathbb{R}^n$ ניתנת להצגה בשלושה שלבים:

(א) קודם כל מפעילים העתקה אורתוגונלית T_{O^t} שמסובבת (ואולי משקפת) את $\mathbb{R}^n$. ההעתקה T_{O^t} שולחת את הוקטורים העצמיים של A לצירים הסטנדרטיים e_i של $\mathbb{R}^n$.

(ב) לאחר מכן מפעילים את T_D שמנפחת את הצירים הסטנדרטיים e_i בפקטור λ_i.

(ג) לבסוף מפעילים את T_O שהיא ההעתקה האורתוגונלית ההופכית של T_{O^t}. ההעתקה T_O שולחת את הצירים הסטנדרטיים e_i בחזרה לוקטורים העצמיים של A.

אנחנו יודעים שלא כל מטריצה $A \in M_n(\mathbb{R})$ ניתנת ללכסון אורתוגונלי. המטריצה A עלולה להיות לא לכסינה כלל וגם אם היא לכסינה, אין סיבה שמרחבים עצמיים שמתאימים לערכים עצמיים שונים יהיו ניצבים זה לזה. ראינו ש־A לכסינה אורתוגונלית אם ורק אם היא סימטרית. האם אנחנו יכולים להגיד משהו אנאלוגי למטריצה כללית שאינה בהכרח סימטרית?

מסתבר שכן - לכל מטריצה A קיימות **זוג** מטריצות אורתוגונליות U,V ומטריצה אלכסונית D כך ש־$A=UDV^t$ (כאן, אנחנו לא בהכרח דורשים ש־$U=V$). הפירוק הנ"ל של A כמכפלה $A=UDV^t$ נקרא **פירוק SVD** ולמעשה הוא נכון למטריצות שאינן בהכרח ריבועיות. הטענה המדוייקת מובאת במשפט הבא:

משפט[1] (פירוק SVD): תהי $A\in M_{m\times n}(\mathbb{R})$ מטריצה. אזי קיימות מטריצות אורתוגונליות $V\in M_n(\mathbb{R})$ ו־$U\in M_m(\mathbb{R})$ ומטריצה $D\in M_{m\times n}(\mathbb{F})$ כך ש־$D_{ij}=0$ לכל $i\neq j$ כך שמתקיים

$$A=UDV^t.$$

איברי האלכסון D_{ii} הם מספרים אי־שליליים המסומנים ב־σ_i ונקראים **הערכים הסינגולריים** של A. אפשר ונהוג לקחת את U ו־V כך שיתקיים

$$\sigma_1\geq\sigma_2\geq\ldots\geq\sigma_m\geq 0.$$

מספר הערכים הסינגולריים **החיוביים** של A שווה לדרגה של A (שהיא כמובן קטנה או שווה ל־$\min(m,n)$).

בדומה לקודם, אם נסמן ב־$T_A\colon\mathbb{R}^n_{\text{col}}\to\mathbb{R}^m_{\text{col}}$ את ההעתקה הלינארית שמתאימה לכפל ב־A משמאל נקבל כי אפשר לחשוב על פירוק SVD בשני אופנים:

1. אם נסמן ב־$\mathcal{B}_1=(v_1,\ldots,v_n)$ את העמודות של V (בסיס אורתונורמלי ל־$\mathbb{R}^n$) וב־$\mathcal{B}_2=(u_1,\ldots,u_m)$ את העמודות של U (בסיס אורתונורמלי ל־$\mathbb{R}^m$) אז $[T_A]^{\mathcal{B}_1}_{\mathcal{B}_2}=D$. כלומר, קיימים **זוג בסיסים אורתונורמליים** (לתחום ולטווח) אשר ביחס אליהם T מיוצגת באופן הכי פשוט שאפשר כ־$[T_A]^{\mathcal{B}_1}_{\mathcal{B}_2}=D$.
כל וקטור $x\in\mathbb{R}^n$ מתפרק באופן יחיד בסכום

$$x=\sum_{i=1}^{n}\langle x,v_i\rangle v_i$$

ופעולת T_A על הוקטור הנ"ל נתונה על־ידי העברת הצירים v_i ל־u_i וניפוח בפקטור σ_i:[2]

$$T_A(x)=T_A\left(\sum_{i=1}^{n}\langle x,v_i\rangle v_i\right)=\sum_{i=1}^{n}\langle x,v_i\rangle T_A(v_i)=\sum_{i=1}^{m}\langle x,v_i\rangle\sigma_i u_i.$$

2. מבלי לבצע שינויי בסיסים, מתקיים $T_A=T_U\circ T_D\circ T_{V^t}$:

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "T_{V^t}"] \arrow[rrr, bend left, "T_A"] & \mathbb{R}^n \arrow[r, "T_D"] & \mathbb{R}^m \arrow[r, "T_U"] & \mathbb{R}^m
\end{tikzcd}$$

כלומר, פעולת T_A על $\mathbb{R}^n$ ניתנת להצגה בשלושה שלבים:

(א) קודם כל מפעילים העתקה אורתוגונלית T_{V^t} שמסובבת (ואולי משקפת) את $\mathbb{R}^n$. ההעתקה T_{V^t} שולחת את הבסיס האורתונורמלי $(v_1,\ldots,v_n)$ לצירים הסטנדרטיים e_i של $\mathbb{R}^n$.

(ב) לאחר מכן מפעילים את T_D (ששולחת את הצירים e_i ב־$\mathbb{R}^n$ לצירים $\sigma_i e_i$ ב־$\mathbb{R}^m$)

(ג) לאחר מכן מפעילים העתקה אורתוגונלית שלישית T_U שמסובבת (ואולי משקפת) את $\mathbb{R}^m$. ההעתקה T_U שולחת את הצירים הסטנדרטיים של $\mathbb{R}^m$ לבסיס האורתונורמלי $(u_1,\ldots,u_m)$.

[1] הגרסא השקולה להעתקות אומרת שאם V מרחב מכפלה פנימית ממשי ממימד n ו־W מרחב מכפלה פנימית ממשי ממימד m ו־$T\colon V\to W$ העתקה לינארית אזי קיימים בסיסים אורתונורמליים $\mathcal{B}_1$ ל־V ו־$\mathcal{B}_2$ ל־W כך ש־$[T]^{\mathcal{B}_1}_{\mathcal{B}_2}=D$ כאשר $D\in M_{m\times n}(\mathbb{R})$ עם $D_{ij}=0$ לכל $i\neq j$.

[2] אם $n\geq m$ אז $\sigma_i=0$ לכל $i>m$ ולכן השיוויון מתקיים.

איור שממחיש את נקודת מבט מס' 2 כאשר $n = m = 2$ (כאשר בציור $D = \Sigma$) הוא האיור הבא:

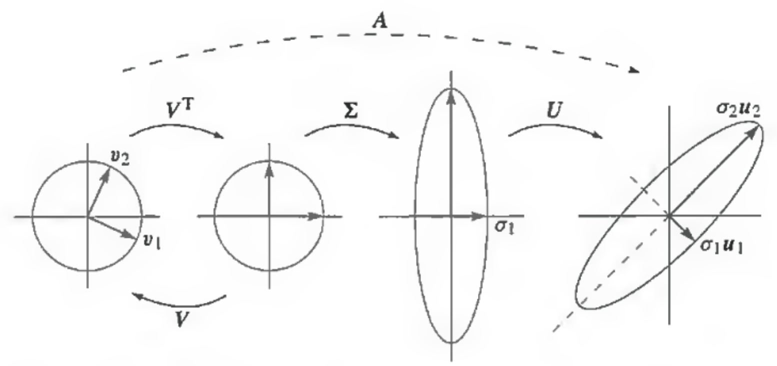


Figure 6.5 U and V are rotations and reflections. Σ is a stretching matrix.

הפירוק הנ"ל שימושי מאוד גם באלגברה לינארית וגם ביישומים של אלגברה לינארית לתחומים אחרים (כמו מדעי המחשב). היתרון שלו טמון במשמעויות הגיאומטריות ובכך שהוא תקף **לכל מטריצה** $A \in M_{m\times n}(\mathbb{F})$. לפני שנעיר בקצרה למה קיים פירוק כזה ואיך מוצאים אותו, ניתן מספר דוגמאות ושימושים.

דוגמא: תהי $A \in M_2(\mathbb{R})$ נתונה על־ידי

$$A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}.$$

הפולינום האופייני של A הוא

$$\chi_A(t) = t^2 - \mathrm{tr}(A)t + \det(A) = t^2 - 3t + 4 = \left(t - \frac{3}{2}\right)^2 + \frac{7}{4}$$

ולכן ל־A אין וקטורים עצמיים והיא לא לכסינה (מעל $\mathbb{R}$). פירוק ה־SVD שלה נתון על־ידי

$$A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{V^t}.$$

המטריצה

$$V = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \cos\left(\frac{\pi}{4}\right) & -\sin\left(\frac{\pi}{4}\right) \\ \sin\left(\frac{\pi}{4}\right) & \cos\left(\frac{\pi}{4}\right) \end{bmatrix}$$

היא מטריצת סיבוב ב־$45°$ נגד כיוון השעון. מכאן, אנחנו מקבלים תיאור גיאומטרי של פעולת T_A על $\mathbb{R}^2$: ההעתקה $T_A = T_D \circ T_{V^t}$ מסובבת תחילה וקטור ב־$45°$ **עם כיוון השעון** ולאחר מכן מפעילה ניפוח בכיוון ציר ה־x בפקטור $\sigma_1 = 2\sqrt{2}$ וניפוח בציר ה־y בפקטור $\sigma_2 = \sqrt{2}$ (בדוגמא הזאת, לא מופעל לבסוף סיבוב נוסף). מכאן, אפשר לראות מיד ש־T_A מעתיקה את מעגל היחידה לאליפסה שהצירים שלה מתלכדים עם ציר ה־x וציר ה־y ואורכי הצירים של האליפסה הם $2\sqrt{2}$ ו־$\sqrt{2}$ בהתאמה.

מה המשמעות הגיאומטרית של פעולת T_{A^t} על $\mathbb{R}^2$? מתקיים $A^t = (DV^t)^t = VD$ ולכן T_{A^t} פועלת **קודם כל** על־ידי ניפוח של הצירים ולאחר מכן על־ידי סיבוב ב־$45°$ נגד כיוון השעון.

אבחנות:

1. באופן כללי, אם $A \in M_n(\mathbb{R})$ ריבועית והפיכה אז הפירוק $A = UDV^t$ מראה ש־T_A מעתיקה ספרות סביב הראשית לאליפסואידים סביב הראשית "מאותו המימד". אם A ריבועית אך לא הפיכה, ההעתקה T_A מעתיקה ספרות סביב הראשית לאליפסואידים סביב הראשית נמצאים על תת־מרחב קטן יותר של $\mathbb{R}^n$ (ממימד $\mathrm{rank}(A)$).

2. באופן כללי, עבור $A \in M_{n\times m}(\mathbb{R})$ הפירוק $A = UDV^t$ נותן משמעות גיאומטרית לפעולת A^t (ודרך כך, להעתקה הצמודה מאחר ו־$(T_A)^* = T_{A^t}$). מתקיים

$$A^t = \left(UDV^t\right)^t = \left(V^t\right)^t D^t U^t = VD^tU^t$$

ולכן T_{A^t} "מפעילה את הסיבובים בסדר הפוך" לפעולת T_A. הדבר ברור יותר כאשר $n = m$ ו־A מטריצה ריבועית. במקרה זה, $A^t = VD^tU^t = VDU^t$ ולכן T_{A^t} פועלת קודם כל בעזרת ההעתקה האורתוגונלית T_{U^t}, לאחר מכן מפעילה את אותו הניפוח T_D ולבסוף פועלת בעזרת ההעתקה האורתוגונלית T_V. כלומר, סדר הסיבובים מתחלף.

אחד השימושים החזקים של SVD הוא בתחום של כיווץ מידע. תהי $A \in M_{m\times n}(\mathbb{R})$ מטריצה (אפשר לחשוב למשל ש־A מתארת תמונת greyscale של m פיקסלים על n פיקסלים). באופן כללי, על־מנת לתאר את המטריצה עלינו לתת mn מספרים ממשיים. אם, לעומת זאת, אנחנו יודעים שהמטריצה A היא מדרגה 1, אז ניתן לרשום אותה כ־

$$A = uv = \begin{bmatrix} u^1 \\ \vdots \\ u^m \end{bmatrix} \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix} = \begin{bmatrix} u^1 v \\ \vdots \\ u^m v \end{bmatrix} = \begin{bmatrix} u^1v_1 & u^1v_2 & \cdots & u^1v_{n-1} & u^1v_n \\ u^2v_1 & u^2v_2 & \cdots & u^2v_{n-1} & u^2v_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ u^{m-1}v_1 & u^{m-1}v_2 & \cdots & u^{m-1}v_{n-1} & u^{m-1}v_n \\ u^mv_1 & u^mv_2 & \cdots & u^mv_{n-1} & u^mv_n \end{bmatrix}$$

כאשר $u \in \mathbb{R}^m_{\mathrm{col}}$ ו־$v \in \mathbb{R}^n_{\mathrm{row}}$. כלומר, בשביל לתאר מטריצה A מדרגה 1, מספיק לתת $m + n$ מספרים ממשיים. הרעיון הבסיסי מאחורי כיווץ A הוא לנסות ולמצוא מטריצה מדרגה נמוכה שתהיה "קרובה ככל האפשר" ל־A. איך SVD עוזר לנו? אנחנו יכולים לרשום[3]

$$A = UDV^t = \begin{bmatrix} u_1^1 & \cdots & u_m^1 \\ \vdots & \ddots & \vdots \\ u_1^m & \cdots & u_m^m \end{bmatrix} \begin{bmatrix} \sigma_1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \dots & \vdots \\ 0 & \cdots & \sigma_m & \cdots & 0 \end{bmatrix} \begin{bmatrix} v_1^1 & \cdots & v_1^n \\ \vdots & \ddots & \vdots \\ v_n^1 & \cdots & v_n^n \end{bmatrix}.$$

קל לבדוק שאם נחליף את D במטריצה D' בה אנחנו שומרים רק את k הערכים הסינגולריים הגדולים ביותר של A ($\sigma_1, \dots, \sigma_k$) ואת שאר הערכים מחליפים באפסים, אנחנו מקבלים מטריצה A' מדרגה לכל היותר k שניתנת לכתיבה

$$A' = UD'V^t = \sigma_1 u_1 v_1^t + \sigma_2 u_2 v_2^t + \ldots + \sigma_k u_k v_k^t$$

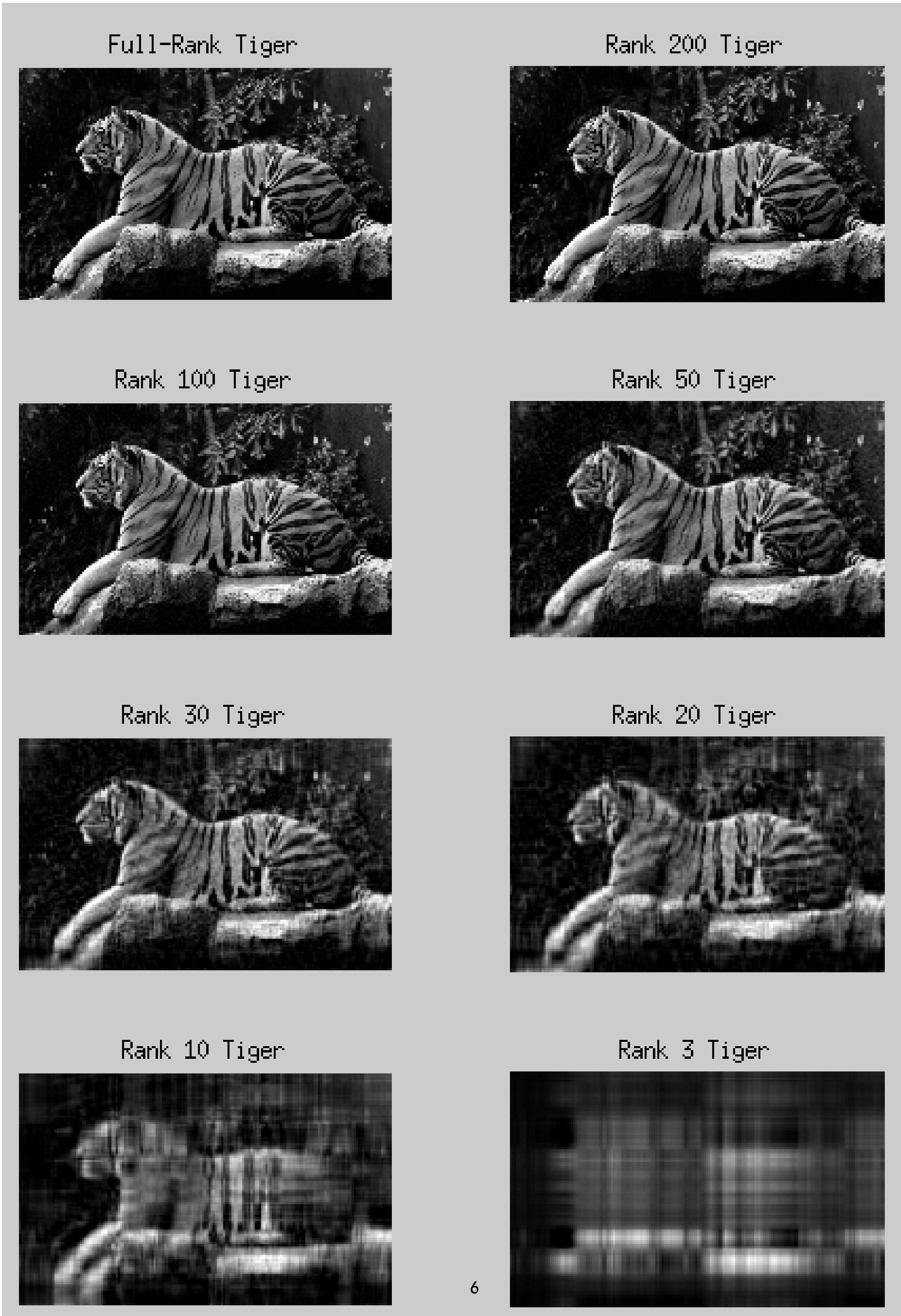
כאשר הוקטורים u_i, v_i הם העמודות של המטריצות U, V בהתאמה. מסתבר שהמטריצה A' נותנת את הקירוב הטוב ביותר למטריצה A מבין המטריצות מדרגה לכל היותר k ויתרה מזאת, הגודל $\frac{\sum_{i=1}^k \sigma_i}{\sum_{i=1}^m \sigma_i}$ נותן מדד "לכמה מידע" אנחנו שומרים מתוך A (כמה טוב הקירוב). בשביל לתאר את המטריצה A' מספיק לתת $k(n+m)$ מספרים. אם אנחנו רוצים לשמור על 90% "מהמידע" ב־A, אנחנו צריכים לקחת k כך ש־$\frac{\sum_{i=1}^k \sigma_i}{\sum_{i=1}^m \sigma_i} \geq 0.9$.

[3] אנחנו מניחים בכתיב זה כי $m > n$. אם $m = n$ או $m < n$, הפיתוח דומה.

בתור דוגמא[4] נתבונן בתמונה הבאה של נמר:

לאחר המרה של התמונה ל־grayscale (הפיכתה לשחור־לבן) והורדת הרזולוציה שלה מתקבלת מטריצה $A \in M_{500\times 800}(\mathbb{R})$ כאשר בכל תא יש מספר ממשי בין 0 ל־1 (שמתאר את העוצמה של הפיקסל). לאחר הפעלת האלגוריתם הנ"ל, אנחנו מקבלים את התוצאות הבאות:

[4] הדוגמא לקוחה מתוך
`http://andrew.gibiansky.com/blog/mathematics/cool-linear-algebra-singular-value-decomposition/`,
שם ניתן למצוא שימושים נוספים לפירוק SVD וקוד MATLAB שמממש את האלגוריתם שתיארנו על תמונות.


Full-Rank Tiger
Rank 200 Tiger
Rank 100 Tiger
Rank 50 Tiger
Rank 30 Tiger
Rank 20 Tiger
Rank 10 Tiger
Rank 3 Tiger

כבר כאשר לוקחים קירוב מדרגה (קטנה או שווה ל־) 50, קשה לראות הבדלים בין התמונה המקורית לתמונה המקורבת. בדוגמא הזאת, זה מאפשר לנו לתאר את התמונה באמצעות $50 \cdot (500 + 800) = 65000$ מספרים במקום 400000 ־ יחס כיווץ של $0.1625 = \frac{65000}{400000}$.

1.1 (העשרה) מדוע פירוק SVD קיים?

נניח לרגע כי קיים פירוק SVD ל־$A \in M_{m\times n}(\mathbb{R})$. כלומר, נרשום $A = UDV^t$. אזי, נקבל כי

$$A^tA = (UDV^t)^t(UDV^t) = VD^tU^tUDV = VD^tDV^t = VD'V^t$$

כאשר $D' \in M_n(\mathbb{R})$ אלכסונית. כלומר, פירוק SVD של A מוביל ללכסון אורתוגונלי של המטריצה הריבועית $A^tA \in M_n(\mathbb{R})$ שהיא מטריצה חיובית (סימטרית עם ערכים עצמיים אי־שליליים). הערכים הסינגולריים של A הם השורשים של הערכים העצמיים של המטריצה A^tA ועמודותיה של המטריצה האורתוגונלית V הן בסיס אורתונורמלי של וקטורים עצמיים של A^tA. בדומה,

$$AA^t = (UDV^t)(UDV^t)^t = UDV^tVD^tU^t = UDD^tU^t = UD''U^t$$

כאשר $D'' \in M_m(\mathbb{R})$ אלכסונית. כלומר, פירוק SVD של A מוביל גם ללכסון אורתוגונלי של המטריצה הריבועית $AA^t \in M_m(\mathbb{R})$ שהיא מטריצה חיובית (סימטרית עם ערכים עצמיים אי־שליליים). הערכים הסינגולריים של A הם גם השורשים של הערכים העצמיים של המטריצה AA^t ועמודותיה של המטריצה האורתוגונלית U הן בסיס אורתונורמלי של וקטורים עצמיים של AA^t.

עתה, אפשר ללכת הפוך. תהי $A \in M_{m\times n}(\mathbb{R})$ ונסמן ב־$r = \text{rank}(A)$. המטריצה A^tA היא חיובית ולכן מהמשפט הספקטרלי היא לכסינה אורתונורמלית וקיים בסיס אורתונורמלי $(v_1, \dots, v_n)$ (ביחס למכפלה הפנימית הסטנדרטית) כך ש־$A^tAv_i = \lambda_i v_i$ ו־$\lambda_i \geq 0$ לכל $1 \leq i \leq n$. נניח בה"כ כי

$$\lambda_1 \geq \ldots \geq \lambda_r \geq \ldots \geq \lambda_n \geq 0$$

(כאשר $\lambda_i > 0$ אם $1 \leq i \leq r$ ו־$\lambda_i = 0$ אם $r < i \leq n$). אם נחשב, נגלה כי $||Av_i||^2 = \lambda_i$ ולכן אם נגדיר $\sigma_i = \sqrt{\lambda_i}$ ו־$u_i = \frac{Av_i}{\sigma_i}$ עבור $1 \leq i \leq r$ ונקבל מיד מההגדרה כי $Av_i = \sigma_i u_i$. בדיקה תראה ש־$(u_1, \dots, u_r)$ היא סדרה אורתונורמלית ואם נשלים אותה לבסיס אורתונורמלי $(u_1, \dots, u_m)$ נקבל כי

$$Av_i = \begin{cases} \sigma_i u_i & 1 \leq i \leq r \\ 0 & r < i \leq n \end{cases}$$

כנדרש.

2 תבניות ריבועיות

2.1 מוטיבציה

במישור האוקלידי $\mathbb{R}^2$, אוסף הנקודות שהקואורדינטות שלהן מקיימת משוואה לינארית נתונה $ax + by = 0$ מהווה ישר דרך הראשית ולהפך (כל ישר כנ"ל מוגדר על־ידי משוואה לינארית). אם מאפשרים משוואות לינאריות לא הומוגניות, כלומר מהצורה $ax + by = c$, מקבלים בדיוק את כל הישרים (האפיניים) במישור. שאלה טבעית בגיאומטריה, היא לסווג את כל העקומים שמוגדרים על־ידי משוואה **ריבועית** (לאו דווקא הומוגנית). כלומר, אוסף הפתרונות למשוואות מהצורה

$$ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

עקומים כנ"ל נקראים גם "שניוניות". אנחנו נתמקד במקרה הפרטי בו למשוואה יש את הצורה

$$ax^2 + bxy + cy^2 = d.$$

כלומר, כאשר הפולינום שמגדיר את המשוואה הוא פולינום **הומוגני** ממעלה שניה.

2.2 ניסוח בשפה של אלגברה לינארית ותבניות

תזכורת: יהי V מרחב וקטורי מעל $\mathbb{F}$. תבנית בילינארית על V זאת העתקה $g\colon V\times V\to\mathbb{F}$ כך שלכל $a\in\mathbb{F}$ ולכל $u,u',v,v'\in V$ מתקיים:

1. (לינאריות במשתנה הראשון) $g(u+u',v)=g(u,v)+g(u',v)$ ו־$g(au,v)=ag(u,v)$.
2. (לינאריות במשתנה השני) $g(u,v+v')=g(u,v)+g(u,v')$ ו־$g(u,av)=ag(u,v)$.

נאמר ש־g **סימטרית** אם $g(u,v)=g(v,u)$ לכל $u,v\in V$. בהינתן תבנית בילינארית סימטרית[5] g אפשר להגדיר פונקציה $q\colon V\to\mathbb{F}$ על־ידי הנוסחא

$$Q(v)=g(v,v).$$

הפונקציה q נקראת **תבנית ריבועית**. יהי $d\in\mathbb{F}$. הקבוצה

$$Q^{-1}(d)=\{v\in V\,|\,q(v)=d\}$$

נקראת **קו גובה** של q (שמתאים לגובה d).

דוגמא: ניקח $\mathbb{F}=\mathbb{R}$ ו־$V=\mathbb{R}^2$. אתם תראו בכיתה שכל תבנית בילינארית סימטרית על V נתונה על־ידי

$$g\left(\begin{bmatrix}x^1\\x^2\end{bmatrix},\begin{bmatrix}y^1\\y^2\end{bmatrix}\right)=ax^1y^1+b(x^1y^2+x^2y^1)+cx^2y^2$$

עבור איזשהם $a,b,c\in\mathbb{R}$. התבנית הריבועית המתאימה ל־g תהיה

$$Q\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=g\left(\begin{bmatrix}x\\y\end{bmatrix},\begin{bmatrix}x\\y\end{bmatrix}\right)=ax^2+2bxy+cy^2.$$

כלומר התבנית הריבועית המתאימה Q היא בדיוק פולינום הומוגני ממעלה 2 במשתנים x,y. אם אנחנו חושבים על הגרף של Q כעל מפה טופוגרפית ($Q(x,y)$ קובעת את הגובה מעל פני הים בנקודה (x,y)) אז הקבוצה $Q^{-1}(d)$ היא באמת קו גובה במפה ־ אוסף הנקודות בגובה d. מקרים פרטיים חשובים:

1. אם

$$g\left(\begin{bmatrix}x^1\\x^2\end{bmatrix},\begin{bmatrix}y^1\\y^2\end{bmatrix}\right)=x^1y^1+x^2y^2$$

אז g היא המכפלה הפנימית הסטנדרטית על $\mathbb{R}^2$ והתבנית הריבועית שמתאימה לה היא

$$Q_1\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=x^2+y^2.$$

קו הגובה $Q_1^{-1}(d)$ הוא מעגל ברדיוס $\sqrt{d}$ סביב הראשית (אם $d>0$), נקודה (אם $d=0$) או קבוצה ריקה (אם $d<0$).

2. אם

$$g\left(\begin{bmatrix}x^1\\x^2\end{bmatrix},\begin{bmatrix}y^1\\y^2\end{bmatrix}\right)=x^1y^1$$

[5] אפשר כמובן לדבר על הפונקציה הזאת גם אם g לא סימטרית, אבל היא משמעותית פחות מעניינת. למשל, אם $\mathrm{char}\mathbb{F}\neq 2$ אז אפשר לשחזר את התבנית הבילינארית מהתבנית הריבועית אבל זה לא יהיה נכון אם g לא סימטרית.

אז

$$Q_2\left(\left[\begin{array}{c} x \\ y \end{array}\right]\right) = x^2.$$

קו הגובה $Q_2^{-1}(d)$ מורכב מזוג ישרים מקבילים (אם $d>0$), קו ישר "כפול" (אם $d=0$) או קבוצה ריקה (אם $d<0$).

שאלה: האם g מגדירה מכפלה פנימית?
תשובה: לא, יש וקטורים ב־$\mathbb{R}^2$ "שהאורך" שלהם הוא אפס.

3. אם

$$g\left(\left[\begin{array}{c} x^1 \\ x^2 \end{array}\right], \left[\begin{array}{c} y^1 \\ y^2 \end{array}\right]\right) = x^1y^1 - x^2y^2$$

אז

$$Q_3\left(\left[\begin{array}{c} x \\ y \end{array}\right]\right) = x^2 - y^2.$$

קוי הגובה $Q_3^{-1}(d)$ הן היפרבולות (אם $d \neq 0$) או זוג ישרים נחתכים (אם $d=0$). בפרט, g איננה מכפלה פנימית מאחר ויש ביחס ל־g וקטורים "שהאורך שלהם הוא 0".

4. אם

$$g\left(\left[\begin{array}{c} x^1 \\ x^2 \end{array}\right], \left[\begin{array}{c} y^1 \\ y^2 \end{array}\right]\right) = x^1y^2 + x^2y^1$$

אז

$$Q_4\left(\left[\begin{array}{c} x \\ y \end{array}\right]\right) = 2xy.$$

קוי הגובה $Q_4^{-1}(d)$ הן היפרבולות (אם $d \neq 0$) או זוג ישרים נחתכים (אם $d=0$).

במקרים 3 ו־4 ראינו שמתקבלים אותן צורות גיאומטריות כקוי גובה. האם נוכל לפרמל את הקשר בין המקרים? נסמן

$$O = \left[\begin{array}{cc} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{array}\right]$$

ונבצע חילוף משתנה

$$\left[\begin{array}{c} x \\ y \end{array}\right] = O\left[\begin{array}{c} u \\ v \end{array}\right] = \left[\begin{array}{c} \frac{1}{\sqrt{2}}(u-v) \\ \frac{1}{\sqrt{2}}(u+v) \end{array}\right].$$

אם נציב את חילוף המשתנה בתבנית Q_4 נקבל כי

$$Q_4\left(\left[\begin{array}{c} x \\ y \end{array}\right]\right) = Q_4\left(\left[\begin{array}{c} \frac{1}{\sqrt{2}}(u-v) \\ \frac{1}{\sqrt{2}}(u+v) \end{array}\right]\right) = 2\frac{1}{\sqrt{2}}(u-v)\frac{1}{\sqrt{2}}(u+v) = u^2 - v^2 = Q_3\left(\left[\begin{array}{c} u \\ v \end{array}\right]\right).$$

כלומר, מתקיים

$$Q_4\left(T_O\left(\left[\begin{array}{c} u \\ v \end{array}\right]\right)\right) = Q_3\left(\left[\begin{array}{c} u \\ v \end{array}\right]\right).$$

זה גורר ש־

$$Q_3^{-1}(d) = (Q_4 \circ T_O)^{-1}(d) = T_O^{-1}(Q_4^{-1}(d))$$

או

$$T_O(Q_3^{-1}(d)) = Q_4^{-1}(d).$$

כלומר, ההעתקה T_O (סיבוב ב־$45°$ נגד כיוון השעון) מעתיקה את קוי הגובה של התבנית הריבועית Q_3 לקוי הגובה של התבנית הריבועית Q_4 (ציירו!).

באופן כללי יותר, ניתן להגדיר יחס שקילות בין תבניות ריבועיות שאומר שתבנית ריבועית Q שקולה לתבנית ריבועית Q' אם קיימת העתקה הפיכה $T\colon V \to V$ כך ש־$Q = Q' \circ T$. בשפה הזאת, מיון קוי גובה של Q (פתרונות למשוואות ריבועיות) הופך להיות למיון מחלקות השקילות של התבניות הריבועיות תחת יחס השקילות הנ"ל. אם מתרגמים זאת בחזרה לתבניות בילינאריות, יחס השקילות הנ"ל הופך לחפיפה והתשובה המלאה תתקבל ממשפט המיון לתבניות בילינאריות מעל $\mathbb{R}$ שתוכיחו בכיתה (משפט סילבסטר)[6].

[6]בדוגמא שלנו, ההעתקה T הייתה אורתוגונלית. העתקה כזאת שומרת על מרחקים ולכן מעתיקה למשל מעגלים למעגלים. אם לעומת זאת דורשים רק ש־T תהיה הפיכה, אז אנחנו מאפשרים גם "מתיחה" של הצירים. תחת מושג השקילות הנ"ל, התבנית הריבועית $x^2 + y^2$ שקוי הגובה שלה הם מעגלים תהיה שקולה לתבנית הריבועית $2x^2 + y^2$ שקוי הגובה שלה הם אליפסות. הדרישה ש־T תהיה הפיכה מול T אורתוגונלית מובילה לשני יחסי שקילות שונים ־ חפיפה רגילה וחפיפה באמצעות מטריצות אורתוגונליות. למשפט סילבסטר יש שתי גרסאות ־ אחת לחפיפה רגילה ואחת לחפיפה אורתוגונלית.