פעולות במטריצות

28 בינואר 2004

אחת הפעולות הבסיסיות שלמדתם באלגברה ליניארית היא כפל מטריצות. אם $C = AB$ כש-A, B מטריצות ממשיות מוגדר, כזכור

$$c_{ij} = \sum_k a_{ik} b_{kj}$$

אם מחשבים את c_{ij} באופן ישיר ע"פ ההגדרה, נדרשות $\Theta(n^3)$ פעולות כפל ופעולות חיבור. בשנת 1969 הפתיע Volker Strassen רבים והציג אלגוריתם יעיל יותר לכפל מטריצות. האלגוריתם שלו דורש רק $O(n^\alpha)$ פעולות אלמנטריות לחישוב כפל של מטריצות $n \times n$ כש $\alpha = \log_2 7 = 2.81\ldots$. בהמשך נמצאו גם אלגוריתמים יעילים יותר. השיטה היעילה ביותר הידועה היום דורשת $O(n^\beta)$ פעולות $\beta = 2.376\ldots$. הרעיון הבסיסי הוא שניתן לכפול שתי מטריצות 2×2 תוך שימוש בשבע (ולא בשמונה) פעולות כפל. היבט חשוב של האלגוריתם הזה הוא שהוא פועל גם אם מדובר במטריצות שאיבריהן שייכים לחוג לא קומוטטיבי (כלומר, xy לא בהכרח שווה ל-yx). נניח לרגע שידוע לנו האלגוריתם האמור לכפל מטריצות. כמו כן נניח לרגע ש-n זוגי (מייד נטפל בעניין זה). נפצל את A, B לבלוקים מגודל $n/2$ ונכפול אותן כך:

$$A = \begin{matrix} A_1 & A_2 \\ A_3 & A_4 \end{matrix} \qquad B = \begin{matrix} B_1 & B_2 \\ B_3 & B_4 \end{matrix}$$

קל לבדוק שאם $A = \begin{matrix} A_1 & A_2 \\ A_3 & A_4 \end{matrix}$ אז $C_1 = A_1 B_1 + A_2 B_3$ וכן הלאה. שימו לב שהכפלים הם כפלי מטריצות (מגודל $n/2$) ואינם קומוטטיביים. הנקודה העיקרית היא שהאלגוריתם לכפל מטריצות שאת קיומו הנחנו יכול לבצע את הכפל תוך שימוש ב-7 פעולות כפל בלבד. פעולות הכפל הן מהצורה $A_i B_j$. מעבר לפעולות הכפל אנו נדרשים לעוד $O(n^2)$ פעולות של חיבור וחיסור. מתקבלת ההערכה הרקורסיבית הבאה למספר הפעולות האריתמטיות $T(n)$ שדורש האלגוריתם המלא לכפל שתי מטריצות דהיינו,

$$T(n) \leq 7T(n/2) + O(n^2)$$

שפתרונו כאמור $T(n) \leq n^\alpha$. בספר תוכלו למצוא ניתוח מטיפוס Reverse engineering (הסדנה?) המתאר דרך שיטתית אפשרית לבניית האלגוריתם הנ"ל ולהלן נאמר עוד משהו על הרעיונות שמאחוריו.

איך נטפל ב-n אי זוגי? נשלים את המטריצה ע"י אפסים למטריצה $2^k \times 2^k$ באשר $k = \lceil \log n \rceil$. זה מגדיל את n בגורם כפלי קטן מ-2 ולכן אינו משפיע על הסיבוכיות האסימפטוטית.

את האלגוריתם לפרטיו תוכלו לקרוא בספר. נאמר רק שבתחילה מחשבים 7 מטר-
יצות

$$P_1 = A_1(B_2 - B_4); P_2 = (A_1 + A_2)B_4, \dots P_7 = (A_1 - A_3)(B_1 + B_2)$$

ואח"כ מחשבים את ה-C_i, למשל $C_2 = P_4 + P_5 + P_6 - P_2$.

הערה נוספת על דרגות של מטריצות ואלגוריתמים לכפל מטריצות

באלגברה ליניארית למדתם את המושג של דרגה של מטריצה. כרגיל מגדירים תחילה דרגת שורות ודרגת עמודות. מאפיינים אותם ע"י כך שרואים את המטריצה כאופרטור ליניארי ומחשבים את המימדים של הגרעין ושל הטווח. יש גם דרך חילופית להגדרת המימד והיא זו: מגדירים תחילה מהי מטריצה מדרגה 1. אומרים ש-A מטריצה מדרגה 1, אם יש שני וקטורים u, v כך ש-$a_{ij} = u_i v_j$ לכל i, j. באופן כללי, אומרים שהדרגה של B היא r, אם ניתן להציג את B כסכום של r מטריצות מדרגה 1 (כמוגדר לעיל) אך לא פחות.

שימו לב כי בפיתוח האלגוריתם של Strassen, לכל אחת מהמטריצות P_i יש הצורה XY כש-X צירוף ליניארי של ה-A_i ו-Y צירוף ליניארי של ה-B_j את ה-C_i ביטאנו כצירופים ליניאריים של ה-P_i . ניתן להראות (אך לא נעשה זאת כאן) שבהצטמצמות לשיטה כזו לא איבדנו מן הכלליות.

לביטוי מן הצורה $P = (\sum \alpha_i A_i)(\sum \beta_i B_i)$ אנו מתאימים את המטריצה שבמקום ה-(i, j) שלה רשום $\alpha_i \beta_j$. זוהי כמובן מטריצה מדרגה 1. גם לכל אחד מה-C_i אנו מתאימים מטריצה. למשל, היות ש-$C_2 = A_1 B_2 + A_2 B_4$ תותאם לה המטריצה

$$\begin{matrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix}$$

(השורה ה-i כאן מתאימה ל-A_i והעמודה ה-j ל-B_j). השאלה נהיית אם כן: מצא 7 מטריצות מדרגה 1 כך שכל 4 המטריצות המתאימות ל-C_i (כמו זו של C_2 המופיעה מקודם) ניתנות לביטוי כצירופים ליניאריים בהן. זו כבר אינה בעייה בנוגע לדרגה של מטריצה אלא של טנסור מסויים, אולם כאן כבר גלשנו הרבה מעבר לתחומנו.

פירוק LUP: רוצים להציג מטריצה ריבועית נתונה A בצורה $PA = LU$ באשר P מטריצת תמורה, L מטריצה משולשית תחתונה ו-U עליונה. להצגה כזו יש היתרון שהיא מקלה על פתרון מערכת משוואות ליניאריות $Ax = b$ מערכת זו מיתרגמת ל-$LUx = Pb$. מערכת מהצורה $Ly = v$ קל לפתור.

$$y_1 = \frac{v_1}{l_{11}}$$

מציבים זאת במשוואה השנייה, מוצאים את y_2 וכו'. באופן דומה פותרים גם מערכת מהצורה $Uz = w$.

איך מוצאים פירוק LUP? אלימינציית גאוס שאותה הכרתם בקורס באלגברה ליניארית היא אלגוריתם המוצא פירוק כזה. נסביר כיצד, ללא פירוט רב. מומלץ לחזור ולרענן את זכרונכם בחומר מהקורס באלגברה ליניארית. תהיה A מטריצה לא סינגולרית ונניח למען הנוחיות כי $a_{11} \neq 0$. נרשום $A = \begin{matrix} a_{11} & w^T \\ v & A' \end{matrix}$ כש-v, w ווקטורי עמודה ו-A' מטריצה $(n-1) \times (n-1)$

פעולות על מטריצות-חלק 2.

מבצעים אלימינציית גאוס עם a_{11} על העמודה שלו בכך מושגת ההצגה

$$A = \begin{pmatrix} 1 & \vec{0} \\ v/a_{11} & I_{n-1} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & v^T \\ 0 & A' \end{pmatrix}$$

וממשיכים ברקורסיה עם A'. אם $a_{11} = 0$ נדרש שינוי בסדר השורות כזה שיביא למקור ה- $(1,1)$ איבר שונה מ- 0 וכך נכנסת מטריצת התמורה P למשחק.

נוכיח עתה: שכפל מטריצות אינו קשה יותר מחישוב ההופכי של מטריצה. זו דוגמה חשובה לעיקרון הצמצום או הרדוקציה בין בעיות שהיא אבן יסוד הן בתורת האלגור-יתמים והן בסיבוכיות החישוב (את הבדיחה על הקומקום כבר שמעתם, אני מקווה).

משפט: כפל מטריצות אינו קשה יותר מהיפוך.

הוכחה:

$$\begin{pmatrix} I & A & 0 \\ 0 & I & B \\ ) &) & I \end{pmatrix}^{-1} = \begin{pmatrix} I & -A & AB \\ 0 & I & -B \\ 0 & 0 & I \end{pmatrix}$$

והתוצאה של הכפל נמצאת בפינה הימנית העליונה.

ניתן להראות גם שהיפוך מטריצות אינו קשה יותר מכפל מטריצות, ולכן שתי הבעיות שקולות זל"ז מבחינת סיבוכיות החישוב שלהן.

עוד שני שימושים חשובים של מטריצות קירוב ע"י ריבועים פחותים: למדתם בוודאי נוסחאות אינטרפולציה. בהנתן n נק' במישור $(x_1, y_1), \ldots, (x_n, y_n)$ יש פולינום אחד ויחיד P ממעלה $\leq n-1$ כך שלכל i מתקיים $P(x_i) = y_i$. יש גם נוסחה מפורשת פשוטה (נוסחת האינטרפולציה של לגרנז') המחשב את הפולינום הזה. מה ניתן לעשות אם ברצוננו לקרב היטב ככל הניתן אולם בעזרת פולינום ממעלה קטנה יותר?

ראשית יש להבהיר מהו מדד הקירבה שבו משתמשים. המדד החשוב ביותר (אך בהחלט לא היחידי) הוא קירוב במובן l_2 שגם נקרא קירוב בריבועים פחותים least squares approximation.

בעצם נטפל בשאלה כללית עוד יותר: נתונה לנו מטריצה $A_{m \times n}$ (בדוגמה של הפולינומים A היא מטריצה ונדרמונדה) ואנו מניחים ש $m >> n$. יש לנו גם ווקטור מטרה y ואנו מחפשים את הקירוב הטוב ביותר כלומר את c שמשיג

$$\min_c ||Ac - y||_2^2.$$

דרך אחרת לחשוב על הבעייה היא גיאומטרית: ב R^m נבנה לנו נקודה y וגם מרחב ליניארי n-מימדי L (המרחב הנפרש ע"י עמודות המטריצה A). עלינו למצוא את הנקודה הקרובה ביותר ל-y במרחב L. נגזור את $||Ac - y||_2^2$ לפי c ונקבל $(Ac - y)^T A = 0$ כלומר, $A^T Ac = A^T y$ ולכן $c = ((A^T A)^{-1} A^T)y$. למטריצה $(A^T A)^{-1} A^T$ קוראים המטריצה הפסאודו-הופכית של A.

פירוק SVD (סקירה בלבד) כזכור לכם מאלגברה ליניארית, למטריצה ממשית סימ-טרית שי ערכים עצמיים ממשיים. הרבה בעיות הכרוכות בעבודה עם מטריצות נעשות פשוטות יותר ע"י שימוש בווקטורים עצמיים וערכים עצמיים ("אנליזה ספקטרלית"). האם יש פעולות אנלוגיות שניתן לעשות גם עם מטריצות שהן לאו דווקא סימטריות או אף מלבניות?

מסתבר שהתשובה היא במידה רבה חיובית. התחליף הנאות הוא הערכים הסינ-גולריים של המטריצה. תהיה $A_{m\times n}$ מטריצה לאו דווקא ריבועית. המטריצה AA^T היא מטריצה ממשית סימטרית. זו מטריצה שהאיבר ה-(i,j) שלה שווה למכפלה הפ-נימית של השורה ה- i בשורה ה- j של A. יתר על כן, קל לראות ש- AA^T היא מטריצה מוגדרת אי-שלילית ולכן יש לה m ערכים עצמיים ממשיים ואי שליליים $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m \geq 0$. למספרים אלה קוראים הערכים הסינגולריים של A . ניתן לבטא כל $A_{m\times n}$ (כאשר $m \leq n$) ממשית בצורה $A = X\Lambda Y$. Λ אלכסונית $\lambda_i = \sqrt{\sigma_i}$. העמודות של X הם הו"ע של XX^T והשורות של Y הן מערכת אורתונורמלית קירובים למטריצה A מדרגה נמוכה. איך מודדים קירבה בין מטריצות? מה מבטאים הערכים הסינגולריים בסיטואציות ממשיות?