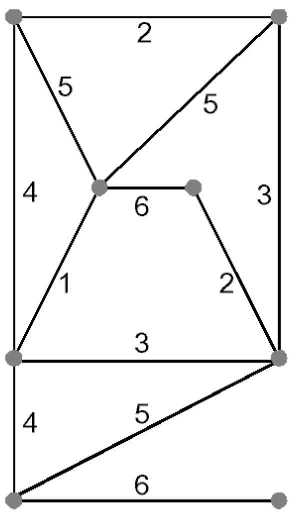
אלגוריתמים איטרטיביים , *Preconditioners*

ברצוננו לפתור *Ax=b* כאשר המטריצה *A* היא מטריצה סימטרית חיובית (כל הערכים העצמיים שלה חיוביים) ודלילה (רוב הערכים בה הם אפס). דרך אחת לפתור מערכת לינארית כזאת היא הדרך הישירה: מציאת *L* משולשית עבורה $A=LL^T$, ואח"כ פתרון *Ly=b* ו – $L^Tx=y$. אולם דרך זאת אינה תמיד מעשית. הזמן הדרוש לחישוב *L* תלוי בגרף של *A* המוגדר להלן:

גרף של מטריצה סימטרית: כל מטריצה סימטרית ניתן לייצג בעזרת גרף שבו יש צומת עבור כל אחד מהמשתנים, וקשת בין צומת *i* לצומת *j* לכל איבר a_{ij} שונה מאפס. משקל קשת זאת הוא a_{ij} -. האיור הבא מתאר מטריצה *A* והגרף המתאים לה.



$$A=\begin{bmatrix} 12 & -2 & -5 & & -4 & & & \\ -2 & 13 & -5 & & & -6 & & \\ -5 & -5 & 17 & -6 & -1 & & & \\ & & -6 & 8 & & -2 & & \\ -4 & & -1 & & 12 & -3 & -4 & \\ & -6 & & -2 & -3 & 16 & -5 & \\ & & & & -4 & -5 & 15 & -6 \\ & & & & & & -6 & 6 \end{bmatrix}$$

יהי G_A הגרף של *A*. מספר הפעולות הדרוש לפירוק *A* ל – LL^T עבור *L* משולשית חסום מלמטה על-ידי גודל מפריד בגרף בחזקת 3. מפריד הוא קבוצת צמתים שמחלקת את הגרף לשני רכיבי קשירות בגודל דומה, שאינם מחוברים בקשתות.
נניח ברצוננו לפתור מערכת לינארית שהגרף שלה הוא שריג תלת מימדי 100*x*100*x*100. גודל מפריד בגרף הוא 100*x*100. לכן זמן הפתרון הוא לפחות $\left(10^4\right)^3=10^{12}$. זמן הפתרון במחשב שמבצע 10^8=100 מגה-פלופס בשניה, הוא $10^4=10000$ שניות, פחות מ – 3 שעות. אם נפתור בעיה בגודל 200*x*200*x*200 הדבר ייקח פי 64 זמן, למעלה משבוע.

בעיה נוספת היא בעיית זכרון: גם אם המטריצה *A* היא דלילה ויש די מקום לאכסן אותה בזכרון, ייתכן כי הפירוק שלה *L* הוא מאוד לא דליל ולא נכנס בזכרון.

כאשר אין אפשרות לפתור בדרך ישירה, מקובל להשתמש בדרכים איטרטיביות. דרך איטרטיבית נפוצה, כאשר *A* היא מטריצה סימטרית וחיובית, היא Conjugate Gradients (CG).
שיטה זאת מחפשת באיטרציה ה – *i*, את הפתרון הקרוב ביותר (בהתייחס לנורמה מסוימת) מתוך המרחב $\{b, Ab, A^2b, A^3b, \ldots, A^ib\}$Span, שנקרא Krylov Subspace. מסתבר שמרחב זה מכיל, לעיתים קרובות, קירוב טוב עבור הפתרון, גם עבור *i* קטן.

קצב ההתכנסות של CG תלוי בהתפלגות הערכים העצמיים של A. אם הערכים העצמיים של A מתפלגים בצורה פחות או יותר אחידה בתחום גדול, ההתכנסות תהיה איטית. אם הם נוטים להיות מקובצים בתחום יותר קטן, ההתכנסות יותר מהירה.

באופן כללי, מספר האיטרציות הדרוש להתכנסות CG חסום על-ידי קבוע כפול שורש ה – condition number של A, שהוא היחס בין הערך העצמי הגדול לערך העצמי הקטן של A. חסם זה אינו בהכרח הדוק.

Preconditioner היא מטריצה שעושים בה שימוש לצורך האצת אלגוריתם איטרטיבי. יהי M^{-1} ה – preconditioner שלנו. במקום לפתור $Ax=b$ נוכל לפתור $M^{-1}Ax=M^{-1}b$. דבר זה עשוי להועיל אם התפלגות הערכים העצמיים של $M^{-1}A$ טובה מזאת של A. אחת הדרכים היא לנסות לבחור מטריצה M כך של - $M^{-1}A$ יש condition number טוב מזה של A. אם M^{-1} קירוב ל – A^{-1}, אז $M^{-1}A$ הוא קירוב למטריצת היחידה, לה יש condition number אחד.

אלגוריתם Preconditioned Conjugate Gradient (PCG), הוא אלגוריתם שפותר מערכת לינארית $Ax=b$ בעזרת פתרון מערכת לינארית אחרת שבה המטריצת המקדמים היא $M^{-\frac{1}{2}}AM^{-\frac{1}{2}}$ (ולא $M^{-1}A$ - כך נשמרת הסימטריה). מסתבר שניתן לעשות זאת מבלי לחשב את המטריצות $M^{\frac{1}{2}}$ או $M^{-\frac{1}{2}}$.

האלגוריתם מבצע בכל איטרציה פתרון משוואה לינארית עם המטריצה M. על מנת שה-preconditioner יהיה יעיל, הזמן שנחסך מכך שההתכנסות היא בפחות איטרציות צריך להיות רב יותר מהזמן שמבוזבז בפיתרון משוואות עם M בכל איטרציה (וגם הפירוק של M). זמן החישוב הכולל של האלגוריתם הוא: זמן חישוב M + זמן פירוק M + זמן שלב איטרציות.

אם ניקח $M=I$, אלגוריתם זה פועל בדיוק כמו Conjugate Gradient ללא preconditioner. הפירוק של I מיידי, פתרון בעזרתה טריוויאלי, אולם קצב ההתכנסות לא משתנה. אין שיפור של ספקטרום הערכים העצמיים.
אם ניקח את $M=A$, הספקטרום של $M^{-1}A$ הוא מושלם. ההתכנסות תהיה באיטרציה אחת. אולם זה יאלץ אותנו לפרק את A, בדיוק הדבר ממנו רצינו להמנע.

Preconditioner טוב הוא מטריצה M שהיא איזשהו קירוב של A, מהבחינה שהספקטרום של $M^{-1}A$ טוב מזה של A, וכך שהפירוק המשולשי של M לא יקר מדי.