

אנליזה נומרית 1

© ארזים

4 במאי 2017

1 שיטת Steepest Descent

ניזכר: הגדרנו $f(x) = x^T A x - b^T x$. מתקיים $\nabla f(x) = Ax - b$, ולכן אנחנו מחפשים מינימום של f . סימנו

$$\begin{aligned} r_i &= b - Ax_i = Ax - Ax_i = A(x - x_i) \\ r_i^T &= (x - x_i)^T A^T = (x - x_i)^T A \end{aligned}$$

האלגוריתם שלנו הוא בשיטת נקודות שבת:

$$\begin{aligned} \alpha_i &= \frac{r_i^T r_i}{r_i^T A r_i} \\ x_{i+1} &= x_i + \alpha_i r_i \\ r_{i+1} &= r_i - A r_i \end{aligned}$$

הגדרנו

$$\|x\|_A = \sqrt{x^T A x}$$

ננסה לחסום את $\|x - x_{i+1}\|_A^2$ מתקיים

$$\begin{aligned} \|x - x_{i+1}\|_A^2 &= \|x - x_i - \alpha_i r_i\|_A^2 = \\ &= (x - x_i - \alpha_i r_i)^T A (x - x_i - \alpha_i r_i) = \\ &= \|x - x_i\|_A^2 + \alpha_i^2 \|r_i\|_A^2 - 2\alpha_i (x - x_i)^T A r_i = \\ &= \|x - x_i\|_A^2 + \frac{(r_i^T r_i)^2}{(r_i^T A r_i)^2} (r_i^T A r_i) - 2 \frac{r_i^T r_i}{r_i^T A r_i} r_i^T r_i = \\ &= \|x - x_i\|_A^2 - \frac{(r_i^T r_i)^2}{\|r_i\|_A^2} \\ p_i &= \frac{(r_i^T r_i)^2}{r_i^T A r_i} \text{ נחסום מלמטה את } \end{aligned}$$

למה 1.1 (תהיה בתרגול) תהי A מטריצה סימטרית ויהי $\lambda_{\max}$ הערך העצמי המקסימלי שלה. אזי

$$\lambda_{\max} = \max_{x \neq 0} \frac{x^T A x}{x^T x}$$

יהי $\lambda_{\min}$ הערך העצמי המינימלי של A . אזי

$$\lambda_{\min} = \min_{x \neq 0} \frac{x^T A x}{x^T x}$$

כעת,

$$p_i = \frac{r_i^T r_i}{r_i^T A r_i} r_i^T r_i \geq \frac{1}{\lambda_{\max}(A)} r_i^T r_i$$

כעת נותר רק לחסום את $r_i^T r_i$ מתקיים

$$\begin{aligned} r_i^T r_i &= (x - x_i)^T A^T A (x - x_i) = (x - x_i)^T A^2 (x - x_i) = \\ &= \frac{(x - x_i)^T A (x - x_i)}{(x - x_i)^T A (x - x_i)} (x - x_i)^T A^2 (x - x_i) = \\ &= \frac{\|x - x_i\|_A^2}{(x - x_i)^T A (x - x_i)} (x - x_i)^T A^2 (x - x_i) \end{aligned}$$

נלכסן את A , כלומר $A = U^T \Lambda U$, כאשר Λ אלכסונית, ונגדיר $q_i = \Lambda^{\frac{1}{2}} U (x - x_i)$ (אפשר להוציא שורש של Λ כי A מוגדרת חיובית). כעת,

$$\begin{aligned} q_i^T \Lambda q_i &= (x - x_i)^T U^T \Lambda^{\frac{1}{2}} \Lambda \Lambda^{\frac{1}{2}} U (x - x_i) = (x - x_i)^T U^T \Lambda^2 U (x - x_i) = (x - x_i)^T A^2 (x - x_i) \\ q_i^T q_i &= (x - x_i)^T U^T \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}} U (x - x_i) = (x - x_i)^T U^T \Lambda U (x - x_i) = (x - x_i)^T A (x - x_i) \end{aligned}$$

לכן נקבל

$$\begin{aligned} r_i^T r_i &= \|x - x_i\|_A^2 \frac{(x - x_i)^T A^2 (x - x_i)}{(x - x_i)^T A (x - x_i)} = \\ &= \frac{q_i^T \Lambda q_i}{q_i^T q_i} \|x - x_i\|_A^2 \geq \lambda_{\min}(A) \|x - x_i\|_A^2 \end{aligned}$$

נאחד את מה שקיבלנו:

$$p_i \geq \frac{\lambda_{\min}}{\lambda_{\max}} \|x - x_i\|_A^2 = \frac{1}{k_2(A)} \|x - x_i\|_A^2$$

המעבר האחרון נכון כי $\|A\|_2 = \lambda_{\max}, \|A^{-1}\|_2 = \frac{1}{\lambda_{\min}}$ ולכן

$$\|x - x_{i+1}\|_A^2 = \|x - x_i\|_A^2 - \frac{(r_i^T r_i)^2}{\|r_i\|_A^2} \leq \|x - x_i\|_A^2 \left(1 - \frac{1}{k_2(A)}\right)$$

אחרי n צעדים של האלגוריתם, נקבל

$$\|x - x_n\|_A^2 \leq \left(1 - \frac{1}{k_2(A)}\right)^n \|x - x_0\|_A^2$$

נחשב את כמות השלבים הנדרשת כדי להקטין את השגיאה לפחות מאשר ε : נרצה

$$\begin{aligned} \left(1 - \frac{1}{k_2(A)}\right)^n &\leq \varepsilon \\ -n \log \left(1 - \frac{1}{k_2(A)}\right) &\geq \log \left(\frac{1}{\varepsilon}\right) \end{aligned}$$

ניזכר שמתקיים

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots = -\sum_{i=1}^{\infty} \frac{(-1)^i x^i}{i}$$

לכן נוכל להעריך:

$$\begin{aligned} -n \log \left(1 - \frac{1}{k_2(A)}\right) &\approx -n \cdot -\frac{1}{k_2(A)} = \frac{n}{k_2(A)} \geq \log \left(\frac{1}{\varepsilon}\right) \\ n &\geq k_2(A) \log \left(\frac{1}{\varepsilon}\right) = -k_2(A) \log(\varepsilon) \end{aligned}$$

נרצה לבדוק איך אפשר להפוך את האלגוריתם לטוב יותר.

2 שיטת Conjugate Direction

צעדי העדכון של האלגוריתם הקודם הם מהצורה $x_{i+1} = x_i + \alpha_i d_i$. הרעיון כאן הוא לבחור כיוונים $d_0, d_1, \dots, d_{n-1}$ כך שבכל כיוון נשתמש פעם אחת בלבד. בכל שלב נבצע עדכון מהצורה

$$x_{i+1} = x_i + \alpha_i d_i$$

נחסר x משני האגפים ונקבל

$$x_{i+1} - x = x_i - x + \alpha_i d_i$$

נסמן $e_i = x_i - x$ ונקבל

$$e_{i+1} = e_i + \alpha_i d_i$$

נכתוב את e_0 בבסיס d_i :

$$e_0 = \sum_{j=0}^{n-1} \delta_j d_j$$

ואז

$$\begin{aligned} e_1 &= e_0 + \alpha_0 d_0 = \sum_{j=0}^{n-1} \delta_j d_j + \alpha_0 d_0 = \\ &= (\delta_0 + \alpha_0) d_0 + \sum_{j=1}^{n-1} \delta_j d_j \end{aligned}$$

נניח כי d_j אורתוגונליים וכן α_0 מקיים $d_0^T e_1 = \delta_0 + \alpha_0$, ואז בשגיאה אחרי שלב אחד לא נצטרך יותר את הכיוון d_0 .