

אנליזה נומרית 1

© ארזים

11 במאי 2017

1 אלגברה לינארית

1.1 שיטת Gradient Descent

נרצה לפתור בהינתן מטריצה סימטרית מוגדרת חיובית A את המערכת $Ax = b$.
נרצה לפתור על ידי שיטה איטרטיבית פשוטה. נגדיר פונקציה

$$f(x) = \frac{1}{2}x^T Ax - b^T x = \frac{1}{2} \langle Ax - 2b, x \rangle$$

נגזור ונקבל $\nabla f(x) = Ax - b$. לכן מספיק לנו למצוא אקסטרמום של הפונקציה. נרצה להניח מנחוש התחלתי x_0 ולנוע במספר סופי של צעדים אל x^* , שהיא אקסטרמום. נרצה להבין כיצד לבחור את כיווני התנועה r_i וגדלי התנועה α_i :

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

תרגיל x^* מינימום גלובלי.

פתרון ראשית, נראה כי x^* מינימום מקומי - אפשר לראות את זה למשל על ידי גזירה פעמיים ובדיקה שההסיאן (מטריצת הנגזרות השניות) מוגדרת חיובית (היא A).
נראה כעת שזה מינימום גלובלי. יהי x כלשהו. נכתוב $x - x^* = e$, ונזכור כי $Ax^* = b$. אזי:

$$\begin{aligned} f(x) &= \frac{1}{2} \langle Ax - 2b, x \rangle = \frac{\langle (Ax^* - 2b) + Ae, x^* + e \rangle}{2} = \\ &= \frac{\langle Ax^*, x^* \rangle + \langle Ae, x^* \rangle + \langle Ax^* - 2b, e \rangle + \langle Ae, e \rangle}{2} = \\ &= f(x^*) + \frac{\langle e, Ax^* \rangle + \langle b - 2b, e \rangle + \langle Ae, e \rangle}{2} = \\ &= f(x^*) + \frac{\langle e, b \rangle - \langle b, e \rangle + \langle Ae, e \rangle}{2} = f(x^*) + \frac{1}{2} \langle Ae, e \rangle \geq f(x^*) \end{aligned}$$

וסיימנו.

לכן האלגוריתם נראה כך: הניחוש בשלב i הוא x_i , הפתרון הוא x . נגדיר $e_i = x - x_i$, $r_i = b - Ax_i$ מסימונים שלנו, $r_i = -\nabla f(x_i)$. נרצה לחשב את α_i , ואפשר לבדוק, לאורך הקו הרלוונטי, לאן כדאי ללכת, ומקבלים

$$\alpha_i = \frac{\langle r_i, r_i \rangle}{\langle Ar_i, r_i \rangle}$$

1.1 הערה 1. הדרך הטבעית למדוד התכנסות היא לא במטריקה הסטנדרטית אלא לפי

$$\|v\|_A = \sqrt{\langle Av, v \rangle}$$

2.

$$f(x_{i+1}) = f(x_i) - \frac{1}{2} \frac{\langle r_i, r_i \rangle^2}{\langle Ar_i, r_i \rangle} < f(x_i)$$

3.

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

$$k_2(A) = \|A\|_2 \|A^{-1}\|_2$$

תרגיל תהי A סימטרית מוגדרת חיובית. הראו

$$\|A\|_2 = \max_{\|x\|_2 \leq 1} \langle Ax, x \rangle := M$$

פתרון יהי x עם $\|x\|_2 \leq 1$ אזי

$$0 < M = \max_{\|x\|_2 \leq 1} \langle Ax, x \rangle \leq \|Ax\|_2 \|x\|_2 \leq \|A\|_2 \|x\|_2^2 \leq \|A\|_2$$

בכיוון השני: נראה $M \geq \|A\|_2$. A סימטרית, ולכן לכסינה אורתוגונלית, כלומר יש U אורתוגונלית עם עמודות $u_1, \dots, u_n$ בסיס אורתונורמלי עם $A = U^T D U$, כאשר D אלכסונית.

$$M = \max_{\|x\|_2 \leq 1} \langle U^T D U x, x \rangle = \max_{\|x\|_2 \leq 1} \langle D U x, U x \rangle = \max_{\|z\|_2 \leq 1} \langle D z, z \rangle = \max_i \alpha_i$$

המעבר הלפני אחרון נכון כי $\|U x\|_2 \leq 1$ לכל $\|x\|_2 \leq 1$ (סימנו $U x = z$). באחרון סימנו את הערכים העצמיים α_i . לבסוף, יהי x עם $\|x\|_2 \leq 1$. נסמן

$$x = \sum_{i=1}^n \delta_i u_i$$

אזי

$$1 \geq \|x\|_2^2 \geq \left\| \sum_{i=1}^n \delta_i u_i \right\|_2^2 \geq \sum_{i=1}^n \delta_i^2$$

ואז

$$\|Ax\|_2^2 = \left\| A \sum_{i=1}^n \delta_i u_i \right\|_2^2 = \left\| \sum_{i=1}^n \delta_i \alpha_i u_i \right\|_2^2 = \sum_{i=1}^n \delta_i^2 \alpha_i^2 \leq \max_i \alpha_i^2$$

לכן לכל x עם $\|x\|_2 \leq 1$ מתקיים

$$\|Ax\|_2 \leq \max_i \alpha_i = M$$

ולכן $\|A\|_2 \leq M$.

1.2 שיטת Conjugate Directions

לשיטה הקודמת יש חיסרון, בדמות האפשרות שנחזור על כיוונים. נרצה בסיס אורתונורמלי A (אורתונורמליים) של כיוונים $d_0, d_1, \dots, d_{n-1}$ כך שבכל כיוון משתמשים פעם אחת. שוב

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

ובנוסף מתכנסים תוך n צעדים. בוחרים את הכיוונים, וא

$$\alpha_i = -\frac{\langle Ae_i, d_i \rangle}{\langle Ad_i, d_i \rangle} = \frac{\langle r_i, d_i \rangle}{\langle Ad_i, d_i \rangle}$$

איך נמצא את הכיוונים? לוקחים בסיס כלשהו $u_0, \dots, u_{n-1}$ בלתי תלויים לינארית, ומבצעים את תהליך גרם שמידט (ביחס למכפלה הפנימית $\langle w, z \rangle_A = \langle Aw, z \rangle$). אם $u_0, \dots, u_{n-1}$ הם סתם ווקטורים, גרם שמידט לוקח $O(n^3)$, כמו פירוק LU . בפעם הבאה נדבר על שיטת Conjugate Gradients, שמתבסס על בחירה נכונה של $u_0, \dots, u_{n-1}$ כך שמתקיים:

1. החישוב של הכיוונים זול יותר.

2. אפשר לעשות תוך כדי האופטימיזציה.