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

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

נגזור ונקבל $\nabla f\left(x\right)=Ax-b$. לכן מספיק לנו למצוא אקסטרמום של הפונקציה. נרצה להניחוש 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:

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

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

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

הערה 1.1 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

$$||x_{i+1} - x||_A^2 \leq \left(1 - \frac{1}{k_2(A)}\right) ||x_i - x||_A^2$$
$$k_2(A) = ||A||_2 \left|\left|A^{-1}\right|\right|_2$$

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 $u_1, \ldots, u_n$ בסיס אורתונורמלי עם $A = U^T DU$, כאשר D אלכסונית.

$$M = \max_{||x||_2 \leq 1} \left\langle U^T DUx, x \right\rangle = \max_{||x||_2 \leq 1} \langle DUx, Ux \rangle = \max_{||z||_2 \leq 1} \langle Dz, z \rangle = \max_i \alpha_i$$

המעבר הלפני אחרון נכון כי $||Ux||_2 \leq 1$ לכל $||x||_2 \leq 1$ (סימנו $Ux = z$). 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

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

ואז

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

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 $d_0, d_1, \ldots, d_{n-1}$ 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 $u_0, \ldots, u_{n-1}$ בלתי תלויים לינארית, ומבצעים את תהליך גרם שמידט (ביחס למכפלה הפנימית $\langle w, z\rangle_A = \langle Aw, z\rangle$). אם $u_0, \ldots, u_{n-1}$ הם סתם ווקטורים, גרם שמידט לוקח $O\left(n^3\right)$, כמו פירוק LU.

בפעם הבאה נדבר על שיטת Conjugate Gradients, שמתבסס על בחירה נכונה של $u_0, \ldots, u_{n-1}$ כך שמתקיים:

1. החישוב של הכיוונים זול יותר.
2. אפשר לעשות תוך כדי האופטימיזציה.