אנליזה נומרית 1

© ארזים

7 במאי 2017

1 שיטת Conjugate direction

אנחנו רוצים לפתור את $Ax = b$, כאשר A מאוד גדולה. נבצע סדרה של עדכונים מהצורה

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

ואז

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

(ניסיון ראשון) נבחר את d_i להיות אורתוגונליים. נפתח את e_0 בבסיס d_i:

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

לפי צעד העדכון:

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

ולכן נבחר את α_0כך שהווקטור d_0 יהיה ניצב לווקטור e_1.

$$0 = d_0^T e_1 = (\delta_0 + \alpha_0) d_0^T d_0$$

נקבל

$$e_1 = \sum_{j=1}^{n-1} \delta_j d_j$$

ובאופן כללי

$$e_i = \sum_{j=i}^{n-1} \delta_j d_j$$

לכן נקבל כי e_i ניצב לכל d_j כאשר $j < i$. כמו כן, אחרי n צעדים, $e_n = 0$ וסיימנו.
נמצא α_i שיקיים

$$d_i^T e_{i+1} = 0$$

כלומר

$$d_i^T \left(e_i + \alpha_i d_i\right) = 0$$
$$\alpha_i = -\frac{d_i^T e_i}{d_i^T d_i}$$

הבעיה - אי אפשר לחשב את α_i, כי e_i לא ידוע.
במקום לבחור כיוונים אורתוגונליים, נבחר אותם להיות A אורתוגונליים, כלומר אורתוגונליים ביחס למכפלה הפנימית שמגדירה A:

$$d_i^T A d_j = 0$$

לכל i, j. נדרוש כעת כי e_{i+1} יהיה A אורתוגונלי לווקטור d_i. כמו קודם

$$e_0 = \sum_{j=0}^{n-1} \delta_j d_j$$
$$e_1 = e_0 + \alpha_0 d_0 = \left(\delta_0 + \alpha_0\right) d_0 + \sum_{j=1}^{n-1} \delta_j d_j$$

נדרוש $d_0^T A e_1 = 0$:

$$0 = d_0^T A e_1 = \left(\delta_0 + \alpha_0\right) \underbrace{d_0^T A d_0}_{\neq 0} + \sum_{j=1}^{n-1} \delta_j \cancelto{0}{d_0^T A d_j}$$

נקבל

$$\delta_0 + \alpha_0 = 0$$

ולכן

$$e_1 = \sum_{j=1}^{n-1} \delta_j d_j$$

ובאופן כללי

$$e_i = \sum_{j=i}^{n-1} \delta_j d_j$$

קיבלנו לכל $j < i$ שמתקיים

$$d_j^T A e_i = 0$$

ואחרי n צעדים נקבל את הפתרון המדוייק. נחלץ את α_i:

$$\begin{aligned} e_{i+1} &= e_i + \alpha_i d_i \\ 0 = d_i^T A e_{i+1} &= d_i^T A e_i + \alpha_i d_i^T A d_i \\ \alpha_i &= -\frac{d_i^T A e_i}{d_i^T A d_i} \end{aligned}$$

כעת

$$Ae_i = A(x_i - x) = Ax_i - Ax = Ax_i - b = -r_i$$

ולכן

$$\alpha_i = \frac{d_i^T r_i}{d_i^T A d_j}$$

נותר לנו רק לבחור את הכיוונים d_i, שיקיימו $d_i^T A d_j = 0$. נבצע גרם־שמידט:
ניקח קבוצה של ווקטורים בלתי תלויים $u_0, \ldots, u_{n-1}$ וכדי לבנות את d_j נחסר את כל הרכיבים של u_j שאינם A אורתוגונליים לכיוונים $d_0, \ldots, d_{j-1}$. כלומר, נקבע $d_0 = u_0$, ועבור $j > 0$:

$$d_j = u_j + \sum_{k=0}^{j-1} \beta_{jk} d_k$$

נדרוש

$$d_j^T A d_l = 0$$

עבור $l < j$:

$$\begin{aligned}0 = d_j^T A d_l &= u_j^T A d_l + \sum_{k=0}^{j-1} \beta_{jk} d_k^T A d_l = \\ &= u_j^T A d_l + \beta_{jl} d_l^T A d_l \\ \beta_{jk} &= -\frac{u_j^T A d_l}{d_l^T A d_l}\end{aligned}$$

החיסרון בשיטה זו הוא שצריך בשלב j לזכור את כל אותם d_l, עבור $l < j$. כמו כן, נדרשות $O\left(n^3\right)$ פעולות כדי לחשב את הכיוונים d_j.

1.1 שיטת Conjugate graduents

נבחר $u_i = r_i$ ואז נקבל שאין צורך לעשות גרם שמידט (בתרגול).

2 מציאת ערכים עצמיים

2.1 שיטת החזקה

נתונה מטריצה A. נרצה למצוא את הערך העצמי הגדול ביותר של A בערך מוחלט, ואת הווקטור העצמי המתאים. נניח כי A ניתנת ללכסון עם ערכים עצמיים $|\lambda|_1 > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$ עם ווקטורים עצמיים $u_1, u_2, \ldots, u_n$. האלגוריתם (שיטת החזקה): ניקח $q^{(0)} \in \mathbb{R}^n$ עם $\left|\left|q^{(0)}\right|\right|_2 = 1$. כעת, עבור כל $1 \leq k \leq n$, נבצע:

$$\begin{aligned}z^{(k)} &= Aq^{(k-1)} \\ q^{(k)} &= \frac{z^{(k)}}{\left|\left|z^{(k)}\right|\right|_2} \\ \lambda^{(k)} &= \left(q^{(k)}\right)^T Aq^{(x)}\end{aligned}$$

חשוב: הערך העצמי הגדול של A הוא יחיד.

נניח כי $q^{(0)}$ מכיל רכיב שאינו 0 בכיוון של u_1. נראה שהאלגוריתם מתכנס לווקטור בכיוון של u_1. קל לראות:

$$q^{(k)} = \frac{A^k q^{(0)}}{\left|\left|A^k q^{(0)}\right|\right|_2}$$

כיוון שהמטריצה A ניתנת ללכסון, נכתוב את $q^{(0)}$ בבסיס הווקטורים העצמיים:

$$q^{(0)} = \sum_{i=1}^{n} a_i u_i$$

כאשר $a_1 \neq 0$. לכן

$$A^k q^{(0)} = \sum_{i=1}^{n} a_i \lambda_i^k u_i = a_1 \lambda_1^k \left(u_1 + \sum_{j=2}^{n} \frac{a_j}{a_1} \left(\frac{\lambda_j}{\lambda_1}\right)^k u_j \right)$$

משפט 2.1 תהי A מטריצה $n \times n$ ניתנת ללכסון שהערכים העצמיים שלה מקיימים $|\lambda_1| > |\lambda_2| \geq \dots$. נניח כי $a_1 \neq 0$. אזי קיים קבוע c עבורו

$$\left|\left|\tilde{q}^{(k)} - u_1\right|\right|_2 \leq c \left|\frac{\lambda_2}{\lambda_1}\right|^k$$

כאשר

$$\tilde{q}^{(k)} = \frac{q^{(k)} \left|\left|A^k q^{(0)}\right|\right|}{a_1 \lambda_1^k}$$

הוכחה: בלי הגבלת הכלליות $||u_i||_2 = 1$. מתקיים

$$\tilde{q}^{(k)} = u_1 + \sum_{j=2}^{n} \frac{a_j}{a_1} \left(\frac{\lambda_j}{\lambda_1}\right)^k u_j$$

ולכן

$$\left|\left|\tilde{q}^{(k)} - u_1\right|\right|_2 = \left|\left|\sum_{j=2}^{n} \frac{a_j}{a_1} \left(\frac{\lambda_j}{\lambda_1}\right)^k u_j\right|\right|_2 \leq \sum_{j=2}^{n} \left|\frac{a_j}{a_1}\right| \left|\frac{\lambda_j}{\lambda_1}\right|^k ||u_j||_2 \leq$$
$$\leq \left|\frac{\lambda_2}{\lambda_1}\right|^k \underbrace{\sum_{j=2}^{n} \frac{|a_j|}{|a_1|}}_{c}$$

■

דוגמה דירוג דפים על ידי גוגל. אם דף i מצביע לדף j אז הוא בעצם נותן לו נקודות. לכל דף יש קול אחד שהוא מחלק בין כל הדפים שאליהם הוא מצביע. מגדירים

$$r(p_i) = \sum_{p_j \to p_i} \frac{r(p_j)}{|p_j|}$$

כאשר $p_j \to p_i$ אומר שדף j מצביע על דף i, $|p_j|$ הוא מספר הדפים שעליהם הדף j מצביע. נגדיר את המטריצה

$$p_{i,j} = \begin{cases} \frac{1}{|p_i|} & p_i \to p_j \\ 0 & otherwise \end{cases}$$

נציב את הווקטור $r\left(p_i\right)$ (ווקטור pagerank) באופן הבא:

$$r\left(p_i\right)=\sum_{p_j\to p_i}\frac{r\left(p_j\right)}{\left|p_j\right|}=\sum_j r\left(p_j\right)p_{j,i}=\left(r^Tp\right)_i$$

קיבלנו שהווקטור צריך לקיים $r^T=r^TP$, כלומר $P^Tr=r$. כלומר r הוא ווקטור עצמי של P^T שמתאים לערך העצמי 1.

2.2 שיטת החזקה ההפוכה

מה עושים אם רוצים לחשב ערך עצמי שאינו הגדול ביותר?

נפתור ראשית בעיה אחרת. ניקח $\mu\in\mathbb{R}$ שאינו ערך עצמי. נתבונן במטריצה $(A-\mu I)^{-1}$. הווקטורים העצמיים שלה זהים לאלה של A, והערכים העצמיים שלה הןא $(\lambda_i-\mu)^{-1}$. מדוע? עבור $A-\mu I$, אם $A=UDU^{-1}$, אזי

$$A-\mu I=UDU^{-1}\mu UU^{-1}=U\left(D-\mu I\right)D^{-1}$$

כלומר הערכים העצמיים של $A-\mu I$ הם $\lambda_i-\mu$. כעת

$$A^{-1}=UD^{-1}U^{-1}$$

ולכן הערכים העצמיים של A^{-1} הם λ_i^{-1} ומכאן שהערכים העצמיים של $(A-\mu I)^{-1}$ הם $(\lambda_i-\mu)^{-1}$.