دانشگاه سیستان وبلوچستان

تحصیلات تکمیلی

پایان نامه کارشناسی ارشد ریاضی کاربردی

عنوان:

روش های کمترین باقیمانده تعمیم یافته ($GMRES$) و بررسی نرخ همگرایی آن در یک سیستم خطی خاص

استاد راهنما:

دکتر پرویز سرگلزایی

تحقیق و نگارش:

طاهره یزدی

(این پایان نامه از حمایت مالی معاونت پژوهشی دانشگاه سیستان و بلوچستان بهره مند شده است)

بهمن ماه ۱۳۸۹

چکیده

در این پایان نامه روش تکراری کمترین باقیمانده تعمیم یافته($GMRES$) را یکبار با استفاده از تبدیلات گیونز و بار دیگر با استفاده از مشتق مورد بررسی قرار داده و سپس آنها را از نظر تعداد اعمال حسابی و مقدار حافظه اشغال شده مورد مقایسه قرار می دهیم.

نرخ همگرایی کمترین باقیمانده تعمیم یافته ($GMRES$) را برای سیستم خطی $Ax = b$ با ماتریس تاپ لیتز سه قطری، وقتی که b ستون اول یا آخر ماتریس همانی باشد مورد بررسی قرار می دهیم.

واژگان کلیدی:

روشهای کمترین باقیمانده تعمیم یافته، ماتریس تاپ لیتز، سیستم خطی، نرخ همگرایی

فهرست مندرجات

پیشگفتار

جواب بسیاری از مسائل عملی در علوم کاربردی سرانجام وابسته به حل دستگاههای معادلات خطی می شود، که معمولاً از روش های تکراری به دست می آید. انگیزه اصلی این روشها در حقیقت سریع و کارا بودن آنهاست. یکی از روشهای تکراری، که بوسیله جنتلمن [۱] در سال ۱۹۷۳ ارائه گردید، روش کمترین باقیمانده تعمیم یافته $(GMRES)$ می باشد که در آن تبدیلات گیونز استفاده می شود.

همچنین کرانهای روش کمترین باقیمانده تعمیم یافته $(GMRES)$ برای سیستم های خطی با ماتریس تاپ لیتز سه قطری که از یک مسئله تابش امواج ناشی می شد، در سال ۲۰۰۰ توسط اپسن [۲] و لیسن [۳] و استرکس [۴] محاسبه شد و آنها پیشنهاد کردند که همگرایی به کندی صورت می گیرد.

سیستم های خطی با ماتریس تاپ لیتز کاربردهای زیادی در معادلات انتگرال و تحلیل سریهای زمانی و نظریه کنترل دارد، در سال ۲۰۰۸ لی [۵] و ژانگ [۶] نرخ همگرایی روش کمترین باقیمانده تعمیم یافته $(GMRES)$ را روی سیستم خطی $Ax = b$ با ماتریس تاپ لیتز سه قطری و چند جمله ای چبیشف نوع اول مورد بررسی قرار دادند و سپس در سال ۲۰۰۹ با استفاده از چند جمله ای چبیشف نوع دوم آنها کار را تکرار کردند.

حال در این پایان نامه روشهای تکراری کمترین باقیمانده تعمیم یافته را با استفاده از تبدیلات گیونز و مشتق مورد مقایسه قرار می دهیم.

در فصل اول به تعریف های مقدماتی مانند تعاریف چند جمله ای چبیشف نوع اول و دوم و زیرفضای کریلف

---

[۱] $Gentelman$

[۲] $Ipsen$

[۳] $Liesen$

[۴] $Strakos$

[۵] Li

[۶] $Zhang$

می پردازیم.

در فصل دوم روشهای تکراری کمترین باقیمانده تعمیم یافته را با استفاده از تبدیلات گیونز و با استفاده از مشتق می پردازیم.

در فصل سوم نرخ همگرایی روش کمترین باقیمانده تعمیم یافته را در یک سیستم خطی با ماتریس تاپ لیتز سه قطری مورد بررسی قرار می دهیم.

در فصل چهارم نتایج عددی فصلهای دوم و سوم را مورد بررسی قرار می دهیم.

فصل ۱

تعاریف و مفاهیم مقدماتی

۱—۱ مقدمه

در این فصل، تعاریف، قضایا و مفاهیمی که برای آشنایی با موضوع پایان نامه لازم و در فصل های بعدی مورد استفاده قرار می گیرند، آورده شده است.

۱—۲ تعاریف

تعریف ۱.۱. فرض کنید A ماتریسی مربعی و از مرتبه n باشد. اسکالر مختلط λ را مقدار ویژه A گویند هر گاه بردار مخالف صفری از $\mathbb{C}^n$ مانند x موجود باشد بطوریکه $Ax = \lambda x$.

بردار x را بردار ویژه نظیر مقدار ویژه λ گویند. همچنین λ مقدار ویژه A است اگر و فقط اگر $det(\lambda I - A) = ۰$ باشد.

همچنین مجموعه تمامی مقادیر ویژه A را با نماد $\sigma(A)$ و بزرگترین مقدار ویژه از نظر قدر مطلق (شعاع طیفی) را با نماد $\rho(A)$ نشان می دهند.

تعریف ۲.۱. فرض کنید $A = (a_{ij})$ ماتریسی از مرتبه $n \times m$ باشد. ترانهاده A را با A^T نشان می دهند و به صورت زیر تعریف می شود:

$$A^T = (a_{ji}) \quad ۱ \leq j \leq m, ۱ \leq i \leq n$$

تعریف ۳.۱. ماتریس A را متقارن گویند هر گاه $A^T = A$ و پادمتقارن گویند هر گاه $A^T = -A$.

تعریف ۴.۱. فرض کنید A ماتریس مختلط و از مرتبه n باشد. ترانهاده مزدوج A را با نماد A^* (یا A^H) نشان می‌دهند، به عبارت دیگر $A^* = (\bar{A})^T$.

تعریف ۵.۱. A را یک ماتریس نرمال گویند هرگاه $A^*A = AA^*$.

تعریف ٦.۱. ماتریس A را یکانی گویند هرگاه $A^*A = I$.

اگر A ماتریسی حقیقی باشد، به جای لفظ یکانی، کلمه متعامد به کار برده می‌شود. به عبارت دیگر ماتریس حقیقی A متعامد است هرگاه $A^TA = I$.

تعریف ۷.۱. فرض کنید A ماتریسی مربعی و از مرتبهٔ n باشد. A را معین مثبت گویند هرگاه

$$\forall x \in \mathbb{C}^n, \quad x \neq 0 \qquad x^*Ax > 0$$

که در آن $x^* = (\bar{x})^T$. در حالت بزرگتر یا مساوی، A را نیمه معین مثبت گویند.

تعریف ۸.۱. تابع $\|.\| : \mathbb{C}^n \to \mathbb{R}$ را یک نرم برداری گویند هرگاه به ازاء هر $x, y \in \mathbb{C}^n$ و هر $\alpha \in \mathbb{C}$، دارای خواص زیر باشد:

۱) $\|x\| \geq 0$

۲) $\|x\| = 0 \leftrightarrow x = 0$

۳) $\|x + y\| \leq \|x\| + \|y\|$

۴) $\|\alpha x\| = |\alpha| \, \|x\|$

چند نرم متداول برداری در زیر آورده شده‌است:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \qquad \text{(نرم ۱)}$$

$$\|x\|_E = (\sum_{i=1}^{n} |x_i|^2)^{\frac{1}{2}} \qquad \text{(نرم ۲ یا نرم اقلیدسی)}$$

$$\|x\|_M = max|x_i| \qquad \text{(نرم ماکزیمم یا چبیشف[1])}$$

تعریف ۹.۱. اگر A ماتریس مختلط و از مرتبه $n \times n$ باشد، نرم ماتریسی A در حالت کلی به صورت زیر تعریف می‌شود:

$$\|A\|_{pq} = \max \frac{\|Ax\|_p}{\|x\|_q} \qquad x \neq 0,\ x \in \mathbb{C}^n$$

یا اینکه

$$\|A\|_{pq} = \max \|Ax\|_p \qquad \|x\|_q = 1$$

نرم ماتریسی، همهٔ ویژگیهای ذکر شده برای نرم برداری را داراست. علاوه براین در نرم ماتریسی داریم:

$$\|AB\| \leq \|A\| . \|B\| \qquad \forall\ A, B \in \mathbb{C}^{n \times n}$$

چند نرم ماتریسی متداول در ذیل آورده شده است:

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{ij}| \qquad \text{(نرم ۱)}$$

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}| \qquad \text{(نرم بینهایت)}$$

$$\|A\|_2 = (\rho(A^*A))^{\frac{1}{2}} = (\rho(AA^*))^{\frac{1}{2}} \qquad \text{(نرم ۲ یا نرم طیفی)}$$

---

[1] Chebyshev

تعریف ۱۰.۱. عدد حالت(وضعیت) یک ماتریس نامنفرد A برای $||.||_p$ بصورت زیر تعریف می شود:

$$\kappa_p = ||A||_p||A^{-1}||_p \quad \text{یا} \quad cond\ A = ||A||||A^{-1}||$$

همچنین عدد حالت یک ماتریس میزانی از نحوه حساسیت دستگاههای معادلات با ماتریس ضرایب A نسبت به آشفتگی های کوچک مانند آشفتگی هایی که در اثر گرد کردن حاصل می شوند را معین می کند. هر گاه $\kappa(A) >> 1$ باشد ماتریس A بدرفتار است.

تعریف ۱۱.۱. زیرفضای کریلف[۲] با بعد m بصورت زیر تعریف می شود:

$$\kappa_m(A,b) = Span\{b, Ab, A^2b, ..., A^{m-1}b\}$$

و همچنین ماتریس کریلف به صورت $K_m = [b\ Ab\ ...\ A^{m-1}b]$ تعریف می شود.

تعریف ۱۲.۱. فرض کنید A ماتریس $m \times n$ $(m \geq n)$ باشد. آنگاه مقادیر ویژه ماتریس A^TA، حقیقی و نامنفی هستند. فرض کنید این مقادیر ویژه، توسط σ_i^2 نمایش داده شوند و $\sigma_n^2 \leq \ldots \leq \sigma_2^2 \leq \sigma_1^2$ آنگاه $\sigma_1, \sigma_2, \ldots, \sigma_n$ مقادیر منفرد A نامیده می شوند.

تعریف ۱۳.۱. چند جمله ای چبیشف نوع اول به صورت زیر تعریف می شود:[۸]

$$\begin{aligned} T_m(t) &= \cos(m\ arc\cos t) && \text{حقیقی}\, t, \quad |t| \leq 1 \\ &= \frac{1}{2}(t+\sqrt{t^2-1})^m + \frac{1}{2}(t-\sqrt{t^2-1})^m && |t| \geq 1 \end{aligned}$$

---

[۲] $Krylov$

تعریف ۱۴.۱. چند جمله ای چبیشف نوع دوم به صورت زیر تعریف می شود:[٦]

$$U_m(t) = \frac{\sin((m+1)\arccos t)}{\sin(\arccos t)} \qquad t \text{ حقیقی}, \quad |t| \leq 1$$

$$= \frac{\left(t+\sqrt{t^2-1}\right)^{m+1} - \left(t-\sqrt{t^2-1}\right)^{m+1}}{2\sqrt{t^2-1}} \qquad |t| \geq 1$$

تعریف ۱۵.۱. مجموعه همه ترکیبات خطی از مجموعه بردارهای $G = \{a_1, a_2, \ldots, a_q\} \subseteq \mathbb{C}^n$ زیر فضای برداری تولید شده توسط G نامیده می شوند و با نماد $span\{G\}$ نمایش داده می شود. یعنی

$$span\{G\} = span\{a_1, a_2, ..., a_q\} =$$

$$\{z \in C^n | z = \sum_{i=1}^{q} \alpha_i a_i, \{\alpha_i\} \in C^q\}$$

اگر a_i ها مستقل خطی باشند، مجموعه G پایه زیر فضا نامیده می شود.

تعریف ۱۶.۱. دو زیر فضای مهم متناظر با ماتریس $A \in C^{n\times m}$، که $R(A)$ تصویر و هسته A با $Ker(A)$ به صورت زیر نمایش داده می شوند:

$$R(A) = \{Ax | x \in C^n\}$$

$$Ker(A) = \{x \in C^m | Ax = 0\}$$

تعریف ۱.۱۷. مجموعه بردارهای $G = \{a_1, a_2, \ldots, a_n\}$ متعامد نامیده می شود، اگر $(a_i, a_j) = 0$ برای هر $i \neq j$ بعلاوه اگر نرم (نرم ۲) هر بردار از G، یک باشد آنگاه G متعامد یکه است،(نماد$(.,.)$ نشانگر ضرب داخلی است).

تعریف ۱.۱۸. ماتریس از مرتبه $m \times n$، $A \in \mathbb{R}^{m\times n}$ دارای رتبه ستونی کامل است، اگر ستونهای آن مستقل خطی باشند و دارای رتبه سطری کامل می باشد در صورتی که سطرهای آن مستقل خطی باشند.

به طور کلی ماتریس A دارای رتبه کامل است اگر یا دارای رتبه سطری کامل و یا رتبه ستونی کامل باشد. اگر دارای رتبه کامل نباشد، دارای رتبه ناقص است.

تعریف ۱.۱۹. فرض کنید $A = (a_{ij})$ یک ماتریس مربعی باشد. اگر به ازای هر $i > j + 1$، $a_{ij} = 0$ باشد آنگاه ماتریس مربعی A بالا هسنبرگی است.

در صورتیکه به ازای هر $j > i + 1$، $a_{ij} = 0$ باشد آنگاه ماتریس مربعی A پایین هسنبرگی است.

توجه: منظور از درایه های خالی در ماتریس های این پایان نامه عدد صفر می باشد.

تعریف ۱.۲۰. ماتریس $A = (a_{ij})$ را در نظر بگیرید. اگر ماتریس A دارای خاصیت زیر باشد، یک ماتریس تاپ لیتز [3] نامیده می شود.[۴]

$$a_{ij} = a_{i-1,j-1}$$

[3] $toeplitz$

و ماتریس B را که به صورت زیر تعریف می شود را ماتریس تاپ لیتز سه قطری گویند:

$$\begin{bmatrix} a & b & & & \\ c & a & b & & \\ & c & a & \ddots & \\ & & \ddots & \ddots & b \\ & & & c & a \end{bmatrix}$$

تعریف ۲۱.۱. فرض کنید $\{x_n\}$ یک دنباله دلخواه باشد. حد بالایی $\{x_n\}$ که با $limsup\ x_n$ نشان داده می شود، به صورت زیر تعریف می گردد:

$$\limsup x_n = \inf\{\sup_{m\geq n} x_m; n\in\mathbb{N}\}$$

و حد پایینی دنباله $\{x_n\}$ که با $lim\ inf\ x_n$ نشان داده می شود، به صورت زیر تعریف می گردد:

$$\liminf x_n = \sup\{\inf_{m\geq n} x_m; n\in\mathbb{N}\}$$

تعریف ۲۲.۱. مسئله کمترین مربعات به صورت زیر تعریف می شود:

فرض کنید A یک ماتریس حقیقی $m\times n$ و b یک بردار باشد، بردار حقیقی x به قسمی تعیین می شود که تابع $||r(x)|| = ||Ax-b||_۲$ مینیمم شود.

تعریف ۲۳.۱. فرض کنید S یک زیر فضا از صفحه مختلط $\mathbb{C}^n$ باشد. آنگاه S یک زیرفضای پایا نامیده می شود اگر $x\in S$ آنگاه $Ax\in S$ باشد که در آن $A=(a_{ij})$ است.

تعریف ۲۴.۱. فرض کنید (d,X) یک فضای متری و A زیر مجموعه ای از فضای X باشد، نقطه $x \in X$ را یک نقطه حدی مجموعه A گوییم در صورتی که هر گوی باز به مرکز x لااقل یک نقطه A متمایز از x را در بر داشته باشد و به صورت زیر نمایش می دهیم

$$\forall r > \circ, B_r(x) \cap A - \{x\} \neq \emptyset$$

قضیه ۲۵.۱. هر گاه $\{x_n\}$ یک دنباله کراندار باشد، آنگاه $liminf\ x_n$ و $limsup\ x_n$ کوچکترین و بزرگترین نقطه حدی $\{x_n\}$ هستند و داریم:

$$liminf\ x_n \leq limsup\ x_n$$

اثبات. [1]

قضیه ۲۶.۱. دنباله کراندار $\{x_n\}$ از اعداد حقیقی همگراست اگر و فقط اگر

$$limsup\ x_n = liminf\ x_n = x$$

در این حالت $\lim x_n = x$.

اثبات. [1]

فصل ۲

روشهای کمترین باقیمانده تعمیم یافته

$(GMRES)$

۲ـــ۱ مقدمه

در این فصل ابتدا روشهای $GMRES-Giv$[۱] و $GMRES-Aya$[۲]، که روش هایی تکراری برای حل مسئله کمترین مربعات هستند، ارائه می دهیم که در روش $GMRES-Giv$ از تبدیلات گیونز استفاده می شود ولی در روش $GMRES-Aya$ از مشتق استفاده می شود.

سپس دو روش ارائه داده شده را از لحاظ تعداد اعمال حسابی، سرعت و ظرفیت حافظه با هم مقایسه می کنیم.

۲ـــ۲ روش کمترین باقیمانده تعمیم یافته

فرض کنید $x_\circ$ یک حدس اولیه برای جواب دستگاه $Ax=b$ بوده و $r_\circ = b - Ax_\circ$، پس جواب دستگاه به صورت زیر نوشته می شود:

$$x = x_\circ + z \in x_\circ + \kappa_k(A, r_\circ) \qquad (۱-۲)$$

حالت ایده‌آل زمانی اتفاق می افتد که

$$A\left(x_\circ + z\right) = b \rightarrow Ax_\circ + Az = b \rightarrow Az = b - Ax_\circ \rightarrow Az = r_\circ \qquad (۲-۲)$$

حال z در معادله مانده (۲ـــ۲) با استفاده از روش $GMRES$[۳] محاسبه می شود. این روش z را در زیر فضای کریلف تولید می کند و به این دلیل جزء روش های زیر فضای کریلف است. برای

[۱] $GMRES-Givens$

[۲] $GMRES-Ayachar$

[۳] $Generalized\ Minimal\ Residual$

تعیین جواب های بهتر در زیر فضای کریلف $\kappa_k(A, r_\circ)$ به یک پایه مناسب نیاز است. اگر بردارهای $r_\circ, Ar_\circ, A^{\text{۲}}r_\circ, \ldots, A^{k-\text{۱}}r_\circ$ مستقل خطی باشند یک پایه برای زیرفضای کریلف ذکر شده تشکیل خواهند داد، که از پایه متعامد $V_k = [v_\text{۱}, v_\text{۲}, \ldots, v_k]$ استفاده می شود. متداول ترین روش برای تولید این پایه روش آرنولدی است. الگوریتم روش آرنولدی بصورت زیر بیان می شود:[۵]

۲ـ۲ـ۱ الگوریتم روش آرنولدی

۱. بردار $v_\text{۱}$ را طوری انتخاب کنید که نرم آن یک باشد.

۲. برای $j = \text{۱}, \text{۲}, \ldots, k$ مراحل ۳ تا ۳.۴ را انجام دهید.

۳. برای $i = \text{۱}, \text{۲}, \ldots, j$، $h_{ij} = (Av_j, v_i)$ را محاسبه کنید.

۳.۱. $w_j = Av_j - \sum_{i=\text{۱}}^{j} h_{ij}v_i$ را محاسبه کنید.

۳.۲. $h_{j+\text{۱},j} = ||w_j||_\text{۲}$.

۳.۳. اگر $h_{j+\text{۱},j} = \circ$ آنگاه توقف کنید.

۳.۴. $v_{j+\text{۱}} = \frac{w_j}{h_{j+\text{۱},j}}$.

در این الگوریتم بردار $v_\text{۱}$ طوری انتخاب می شود که نرم آن یک باشد و بردارهای $v_\text{۲}, \ldots, v_k$ طوری ساخته می شود که $\{v_\text{۱}, v_\text{۲}, \ldots, v_k\}$ پایه ای متعامد باشد.

$$j = 1\ (i = 1) : h_{11} = (Av_1, v_1)$$

$$w_1 = Av_1 - h_{11}v_1, h_{21} = ||w_1||_2$$

$$v_2 = \frac{w_1}{h_{21}} = \frac{Av_1 - h_{11}v_1}{h_{21}} \qquad (3-2)$$

$$j = 2\ (i = 1) : h_{12} = (Av_2, v_1)$$

$$(i = 2) : h_{22} = (Av_2, v_2), w_2 = Av_2 - h_{12}v_1 - h_{22}v_2$$

$$h_{32} = ||w_2||_2, v_3 = \frac{w_2}{h_{32}} = \frac{Av_2 - h_{12}v_1 - h_{22}v_2}{h_{32}} \qquad (4-2)$$

و به همین ترتیب v_i ها به دست می آیند. که در تکرار j ام داریم:

$$v_{j+1} = \frac{Av_j - h_{1j}v_1 - h_{2j}v_2 - ... - h_{jj}v_j}{h_{j+1,j}} \qquad (5-2)$$

که در آن

$$h_{j+1,j} = ||Av_j - h_{1j}v_1 - h_{2j}v_2 - ... - h_{jj}v_j||_2$$

در هر تکرار این الگوریتم یک ستون ماتریس زیر حاصل می شود.

$$\overline{H_k} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1k} \\ h_{21} & h_{22} & \cdots & h_{2k} \\ 0 & h_{32} & \cdots & h_{3k} \\ 0 & 0 & \cdots & \vdots \\ \vdots & \vdots & & h_{kk} \\ 0 & 0 & & h_{k+1,k} \end{bmatrix}$$