



دانشکده علوم ریاضی
گروه ریاضی کاربردی

پایان نامه

برای دریافت درجه کارشناسی ارشد در رشته
ریاضی کاربردی، گرایش آنالیز عددی

عنوان

روش عناصر متناهی تطبیقی برای مسایل بیضوی با ضرایب ناپیوسته

استاد راهنما

دکتر علی مس فروش

استاد مشاور

دکتر مهدی قوتمند جزئی

دانشجو

سولماز انکاری

آذر ۱۳۹۳

تقدیم به خداوند یکتا

و دو فرشته اش، پدر و مادر عزیزم که بال پروانه به من دادند
و یگانه شریک و همراه جاده زندگی، همسرم

سپاس‌گزاری...

سپاس و ثنا یگانه خالص را که ذات وجودم در تلالو حضورش نورانی می‌شود و نگاه خسته‌ام از جوشش مهرش جان می‌گیرد، سپاس می‌گویم او را که یگانه‌ترین در عظمت، تنهاترین در اوج و پاک‌ترین در وجود است.

اکنون که به لطف حق تعالی نگارش این پایان نامه به اتمام رسیده، وظیفه خود می‌دانم مراتب سپاس‌گزاری خویش را حضور:

- استاد ارجمندم آقای دکتر علی مس‌فروش به عنوان استاد راهنما چون چراغی روشن‌گر راهم بودند و با راهنمایی‌های ارزشمند خود در انجام هر چه بهتر این پژوهش مرا یاری نمودند،
- جناب آقای دکتر مهدی قوتمند به عنوان استاد مشاور،
- اساتید بزرگوار آقایان دکتر حجت احسنی طهرانی و دکتر علیرضا ناظمی که زحمت داوری این پایان نامه را تقبل فرمودند،
- پدر و مادرو همسر مهربانم که همواره حامی من بوده‌اند،
- تمامی دوستان عزیزی که با محبت‌های صادقانه‌شان همراهم بودند،

اعلام نمایم.

سولماز انخاری
آذر ۱۳۹۳

تعمدنامه

اینجانب سولماز انکاری دانشجوی کارشناسی ارشد رشته ریاضی کاربردی دانشکده علوم ریاضی دانشگاه شاهرود، نویسنده پایان نامه با عنوان روش عناصر متناهی تطبیقی برای مسایل بیضوی با ضرایب ناپیوسته، تحت راهنمایی دکتر علی مس فروش متعهد می شوم:

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش های دیگر پژوهشگران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان نامه، تاکنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه شاهرود تعلق دارد، و مقالات مستخرج با نام “دانشگاه شاهرود” یا “Shahrood University” به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آوردن نتایج اصلی پایان نامه تاثیرگذار بوده اند، در مقالات مستخرج از پایان نامه رعایت می گردد.
- در تمام مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافت های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده) شده است، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

سولماز انکاری
آذر ۱۳۹۳

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزارها و تجهیزات ساخته شده) متعلق به دانشگاه شاهرود می باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان نامه بدون ذکر منبع مجاز نمی باشد.

چکیده

معادلات دیفرانسیل جزئی بیضوی با ضرایب انتشار ناپیوسته در دامنه‌های کاربردی همچون انتشار از طریق رسانه متخلخل، انتشار میدان الکترومغناطیس در رسانه‌های ناهمگن، و پروسه‌های انتشار در سطوح خشن روی می‌دهند. روش استاندارد برای حل عددی این مسایل با استفاده از روش‌های عناصر متناهی عبارت است از فرض این واقعیت که، ناپیوستگی در مرزهای سلول‌های مثلث اولیه بوجود آمده است. اما، این مساله با کاربرد ناپیوستگی در منحنی‌ها، سطوح، یا چند شاخه‌ای‌ها مطابقت نداشته و از قبل قابل شناسایی نیست. یکی از موانع حل اینگونه مسایل ناپیوسته این است که نظریه اختلال برای معادلات دیفرانسیل با مشتقات جزئی بیضوی، مرزهایی برای اختلال ضرایب در نورم L_∞ در نظر می‌گیرد، این مساله مستلزم این است که ناپیوستگی به طور کامل همسان بوده، اما ضرایب تقریبی باشند. روش جدید را براساس اختلال ضرایب در نورم L_q با $q < \infty$ ارایه می‌کنیم که به همین دلیل نیازی به تطابق دقیق ضرایب ندارد. از این نظریه اختلال جدید برای فرموله کردن روش‌های عناصر محدود تطبیقی جدید استفاده می‌کنیم تا مسایل ناپیوستگی را حل کنیم.

کلمات کلیدی: مسایل بیضوی، ضرایب ناپیوسته، تخمین اختلال، روش‌های عناصر متناهی تطبیقی، نسبت بهینگی از همگرایی

لیست مقالات مستخرج از پایان نامه

۱. انکاری.س، مس فروش.ع، ”روش های عناصر محدود تطبیقی برای مسایل بیضوی با ضرایب ناپیوسته”، چهل و پنجمین کنفرانس ریاضی سمنان، شهریور ۱۳۹۳.

فهرست مطالب

۱	تعاریف و پیش‌نیاز	۱
۱	۱.۱ مقدمه	۱
۲	۲.۱ تعاریف و مفاهیم اولیه	۲
۷	۳.۱ معادلات دیفرانسیل با مشتقات جزئی	۷
۹	۱.۳.۱ روش های عددی حل معادلات دیفرانسیل با مشتقات جزئی	۹
۱۱	۴.۱ روش عناصر متناهی	۱۱
۱۱	۵.۱ تحلیل روش عناصر متناهی برای مسایل دو بعدی	۱۱
۱۲	۱.۵.۱ تجزیه دامنه	۱۲
۱۲	۲.۵.۱ معادلات المان	۱۲
۱۳	۳.۵.۱ محاسبه ماتریس‌های المان	۱۳
۱۴	۴.۵.۱ ساخت ماتریس‌های المان	۱۴
۱۴	۵.۵.۱ حل معادلات تولید شده	۱۴
۱۴	۶.۱ انواع خطاها	۱۴
۱۷	۲ مسایل بیضوی	۱۷
۱۷	۱.۲ مقدمه	۱۷
۱۸	۱.۱.۲ مثال‌هایی از معادله‌های بیضوی	۱۸
۱۹	۲.۱.۲ جواب اساسی و تابع گرین	۱۹
۲۰	۲.۲ روش عناصر متناهی برای معادله‌های بیضوی	۲۰
۲۰	۱.۲.۲ مسأله‌ای در صفحه	۲۰
۲۱	۲.۲.۲ گسسته سازی	۲۱
۲۳	۳.۲ برآورد خطا در نورم H^1	۲۳
۲۴	۴.۲ خطای پسین	۲۴
۲۷	۳ آنالیز تطبیقی	۲۷
۲۷	۱.۳ مقدمه	۲۷
۲۷	۲.۳ انواع روش‌های آنالیز تطبیقی در حالت کلی	۲۷

۲۸	تولید شبکه‌ی عناصر متناهی	۳.۳
۲۹	حل تطبیقی در عناصر متناهی	۴.۳
۳۰	دوبخشی جدیدترین راس	۵.۳
۳۲	تقریب تطبیقی در نورم $H^1(\Omega)$	۶.۳
۳۳	تقریب تطبیقی در نورم $H^{-1}(\Omega)$	۷.۳
۳۵	روش عناصر متناهی تطبیقی برای مسایل بیضوی با ضرایب ناپیوسته	۴
۳۵	مقدمه	۱.۴
۴۰	برهان اختلال	۲.۴
۴۰	قضیه اختلال	۱.۲.۴
۴۲	شرایط کافی برای ∇u که در L_p باشد	۲.۲.۴
۴۴	روش‌های عناصر متناهی تطبیقی	۳.۴
۴۴	افرازاها و فضای عناصر متناهی	۱.۳.۴
۴۵	ساختار AFEM	۲.۳.۴
۴۹	اندازه‌گیری عملکرد AFEM	۳.۳.۴
۵۳	تقریب ε و عملکرد بهینه رده	۴.۳.۴
۵۴	AFEM برای ماتریس‌های انتشار ناپیوسته (DISC)	۴.۴
۵۸	الگوریتم‌های RHS و COEFF	۵.۴
۵۹	الگوریتم‌های بهینه برای تقریب تطبیقی تابع	۱.۵.۴
۶۰	الگوریتم‌های بهینه برای COEFF	۲.۵.۴
۶۵	آزمایشات عددی	۵
۶۶	نتایج عددی	۱.۵
۶۶	مثال ۱: دامنه L شکل.	۱.۱.۵
۶۸	مثال ۲: صفحه چکرز	۲.۱.۵
۷۱	نتیجه‌گیری	۲.۵
۷۲	آ کدهای متلب	
۷۲	کد متلب مثال ۱	۱.آ
۸۹	کد متلب مثال ۲	۲.آ
۹۰	مراجع	
۹۳	واژه‌نامه فارسی به انگلیسی	
۹۸	واژه‌نامه انگلیسی به فارسی	

فصل ۱

تعاریف و پیش‌نیاز

۱.۱ مقدمه

وقتی معادله دیفرانسیل با مشتقات جزئی^۱ بیضوی گسسته می‌شود، یک دستگاه معادله خطی حاصل می‌شود. روش‌های عددی متفاوتی برای حل دستگاه معادلات تفاضل محدود حاصل از معادله دیفرانسیل با مشتقات جزئی بیضوی وجود دارد.

در بسیاری از کاربردهای عملی، معادلات دیفرانسیل با مشتقات جزئی با ضرایب ناپیوسته به طور عددی حل شده است. این مسایل معمولاً از طریق روش عناصر متناهی، گسسته می‌شوند. اساس کار این روش با حذف کامل معادلات دیفرانسیل یا ساده‌سازی آنها به معادلات دیفرانسیل معمولی، که با روش‌های عددی مثل اویلر حل می‌شوند، می‌باشد. ایده اولیه روش عناصر متناهی از لزوم تحلیل سازه هواپیما شکل گرفت. ریشه‌های توسعه این روش را باید در اوایل دهه ۱۹۴۰ میلادی جستجو کرد. در روش‌های عددی، به ویژه روش عناصر متناهی، ضمن بیان مبانی تخمین خطا و حل تطبیقی در عناصر متناهی، از روش عناصر متناهی تطبیقی به عنوان ابزاری قوی برای غلبه بر محدودیت‌های روش عناصر متناهی معمولی استفاده شد. در این پایان‌نامه از روش عناصر متناهی تطبیقی برای حل مسایل عددی استفاده می‌کنیم که هزینه محاسباتی کمتر خواهد شد.

در این فصل به معرفی مفاهیمی که در سراسر این پایان‌نامه مورد استفاده قرار می‌گیرند، می‌پردازیم. ابتدا مفاهیم و تعاریف اولیه، را بیان نموده، مروری گذرا بر فضاها، باناخ، هیلبرت، L^p ، سوبولوف و قضایای مرتبط با آنها خواهیم داشت. در زیربخش اول از این فصل، معادلات دیفرانسیل با مشتقات جزئی و در زیر بخش دوم، روش عناصر متناهی را معرفی خواهیم کرد. در فصل دوم، مسایل بیضوی و روش عناصر متناهی برای این مسایل را داریم. در فصل سوم روش عناصر متناهی تطبیقی را معرفی می‌کنیم. در فصل چهارم روش عناصر متناهی تطبیقی برای مسایل بیضوی با ضرایب ناپیوسته معرفی می‌شوند. در فصل پنجم برخی نتایج عددی ارائه می‌شود.

برای گردآوری این فصل از مراجع [۱]، [۱۷]، [۲۱]، [۲۲]، [۲۳]، [۲۷]، استفاده شده است.

^۱Partial Differential Equation

۲.۱ تعاریف و مفاهیم اولیه

تعریف ۱.۲.۱. فضای خطی V (یا فضای برداری) با اسکالرهایی حقیقی نامیده می‌شود هرگاه:

$$\forall u, v \in V, \quad \forall \mu, \lambda \in \mathbb{R}, \quad \lambda u + \mu v \in V,$$

تعریف ۲.۲.۱. تابع $L: V \rightarrow \mathbb{R}$ را تابع خطی می‌نامیم هرگاه،

$$L(\lambda u + \mu v) = \lambda L(u) + \mu L(v), \quad u, v \in V, \quad \lambda, \mu \in \mathbb{R}.$$

برقرار باشد. $L(\cdot)$ را کران‌دار گوییم هرگاه ثابت C مستقل از u وجود داشته باشد به طوری که،

$$|L(u)| \leq C \|u\|.$$

تعریف ۳.۲.۱. فرم دو خطی $a(\cdot, \cdot)$ روی V تابعی مانند $a: V \times V \rightarrow \mathbb{R}$ است، که نسبت به هر مولفه خود خطی است، یعنی برای هر $u, v, w \in V$ و $\lambda, \mu \in \mathbb{R}$ روابط زیر برقرار باشند:

$$a(\lambda u + \mu v, w) = \lambda a(u, w) + \mu a(v, w) \quad \bullet$$

$$a(w, \lambda u + \mu v) = \lambda a(w, u) + \mu a(w, v) \quad \bullet$$

تعریف ۴.۲.۱. فرم دوخطی $a(\cdot, \cdot)$ را متقارن می‌گوییم هرگاه:

$$a(w, v) = a(v, w), \quad \forall v, w \in V,$$

و معین مثبت گوییم هرگاه،

$$a(v, v) > 0, \quad \forall v \in V, \quad v \neq 0.$$

تعریف ۵.۲.۱. فرم دوخطی که روی فضای برداری V مثبت معین و متقارن باشد، ضرب داخلی^۲ روی V نامیده می‌شود. فضای خطی V به همراه ضرب داخلی، فضای ضرب داخلی نامیده می‌شود.

تعریف ۶.۲.۱. اگر V فضای ضرب داخلی و $\langle \cdot, \cdot \rangle$ ضرب داخلی روی V باشد، نرم متناظر با این ضرب داخلی به شکل زیر تعریف می‌شود:

$$\|v\| = \langle v, v \rangle^{\frac{1}{2}}, \quad \forall v \in V,$$

تعریف ۷.۲.۱. ماتریس A را متقارن می‌نامیم هرگاه،

$$A^T = A,$$

و آن را معین مثبت گوییم هرگاه،

$$v^T A v > 0, \quad \forall v \neq 0.$$

باشد. اگر به جای علامت $>$ علامت $\geq$ باشد، ماتریس را نیمه معین مثبت می‌گوییم.

لم ۸.۲.۱. برای هر $v, w \in V$ ، نامساوی کوشی-شوارتز^۳ به صورت زیر است:

$$|\langle w, v \rangle| \leq \|w\| \|v\|.$$

^۲Inner Product

^۳Cauchy Schwarz

لم ۹.۲.۱. برای هر $v, w \in V$ ، نامساوی مثلثی به فرم زیر تعریف می‌شود:

$$\|w + v\| \leq \|w\| + \|v\|.$$

تعریف ۱۰.۲.۱. دنباله‌ی نامتناهی $\{v_i\}_{i=1}^{\infty}$ را در V همگرا گوئیم و با نماد

$$\lim_{i \rightarrow \infty} v_i = v$$

نشان می‌دهیم هرگاه،

$$\|v_i - v\| \rightarrow 0, \quad i \rightarrow \infty.$$

تعریف ۱۱.۲.۱. دنباله‌ی $\{v_i\}_{i=1}^{\infty}$ کوشی نامیده می‌شود هرگاه،

$$\|v_i - v_j\| \rightarrow 0, \quad i, j \rightarrow \infty$$

تعریف ۱۲.۲.۱. اگر هر دنباله‌ی کوشی در V همگرا باشد، فضای ضرب داخلی V ، کامل نامیده می‌شود.

تعریف ۱۳.۲.۱. فضای ضرب داخلی کامل، فضای هیلبرت^۴ نام دارد.

تعریف ۱۴.۲.۱. نورم را در فضای خطی V به شکل تابعی مانند $\|\cdot\| : V \rightarrow \mathbb{R}^+$ تعریف می‌کنیم به گونه‌ای که خواص زیر را داشته باشد:

$$\|v\| > 0 \quad \forall v \in V, \quad v \neq 0.$$

$$\|\lambda v\| = |\lambda| \|v\| \quad \forall \lambda \in \mathbb{R}, \quad v \in V.$$

$$\|v + w\| \leq \|v\| + \|w\| \quad \forall v, w \in V.$$

تابع $|\cdot|$ نیم نورم نامیده می‌شود، هرگاه شرایط سه گانه بالا برقرار باشند، با این تفاوت در شرط اول که

$$|v| \geq 0, \quad \forall v \in V.$$

تعریف ۱۵.۲.۱. فضای خطی به همراه نورم، فضای خطی نورم‌دار نامیده می‌شود.

تعریف ۱۶.۲.۱. فضای نورم‌دار کامل، فضای باناخ^۵ نامیده می‌شود.

تعریف ۱۷.۲.۱. مجموعه همه تابع‌های خطی کراندار روی V ، فضای دوگان V نامیده شده و با V^* نمایش داده می‌شود. نورم در V^* به شکل زیر است:

$$\|L\|_{V^*} = \sup_{v \in V \setminus \{0\}} \frac{|L(v)|}{\|v\|_V}$$

تعریف ۱۸.۲.۱. فرم دو خطی $a(\cdot, \cdot)$ را کراندار گوئیم هرگاه مقدار ثابت M به گونه‌ای موجود باشد که

$$|a(v, w)| \leq M \|w\|_V \|v\|_V, \quad \forall v, w \in V.$$

^۴Hilbert Space

^۵Banach Space

تعریف ۱۹.۲.۱. فرم دوخطی متقارن $a(\cdot, \cdot)$ در V کواریسیو^۶ گفته می‌شود هرگاه،

$$a(v, v) \geq \alpha \|v\|_V^2, \quad \forall v \in V, \quad \alpha > 0.$$

تعریف ۲۰.۲.۱. اگر $a(\cdot, \cdot)$ فرم دو خطی متقارنی باشد که در V ، کواریسیو و کراندار می‌باشد، آنگاه نورم انرژی $\|\cdot\|_a$ را به شکل

$$\|v\|_a = a(v, v)^{\frac{1}{2}}, \quad \forall v \in V,$$

تعریف می‌کنیم.

لم ۲۱.۲.۱. نورم انرژی $\|\cdot\|_a$ روی V با $\|\cdot\|_V$ هم ارز است. بنا به خاصیت کواریسیو و کراندار بودن نورم انرژی داریم:

$$\sqrt{\alpha} \|v\|_V \leq \|v\|_a \leq \sqrt{M} \|v\|_V.$$

قضیه ۲۲.۲.۱. (Lax – Milgram). اگر فرم دو خطی $a(\cdot, \cdot)$ در فضای هیلبرت V ، کراندار و کواریسیو باشد و L تابعی خطی کرانداری روی V باشد، آنگاه عضو یکتای $u \in V$ موجود است به گونه‌ای که،

$$a(u, v) = L(v), \quad \forall v \in V.$$

تعریف ۲۳.۲.۱. فرض کنید $\mathbb{R}$ مجموعه اعداد حقیقی و

$$\mathbb{R}^d = \{(x_1, x_2, \dots, x_d), x_i \in \mathbb{R}, i = 1, 2, \dots, d\}$$

باشد، Ω دامنه نامیده می‌شود هرگاه، $\Omega \subset \mathbb{R}^d$ باز و محدب باشد. این دامنه برای $d = 1$ بازه، برای $d = 2$ بخشی از صفحه، و برای $d = 3$ ، زیرمجموعه‌ای از فضا است. مرز Ω را با نماد $\partial\Omega$ یا Γ نمایش می‌دهیم. بستار Ω به شکل $\bar{\Omega} = \Omega \cup \Gamma$ تعریف می‌شود. طول، سطح و حجم Ω با نماد $|\Omega|$ نمایش داده می‌شود.

تعریف ۲۴.۲.۱. (بردار گرادیان). اگر u در $\mathbb{R}^n$ تعریف شده باشد، گرادیان u در $x = (x_1, \dots, x_n)^T$ برداری است در $\mathbb{R}^n$ که با $\nabla u = \text{gradu} = (\frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_n})$ تعریف می‌شود.

تعریف ۲۵.۲.۱. (دیورژانس). اگر $u = (u_1, \dots, u_n)^T$ یک میدان برداری باشد، دیورژانس u در $x = (x_1, \dots, x_n)^T$ به صورت زیر تعریف می‌شود:

$$\text{div} u = \nabla \cdot u = \frac{\partial u_1}{\partial x_1} + \dots + \frac{\partial u_n}{\partial x_n} = \sum_{j=1}^n \frac{\partial u_j}{\partial x_j}.$$

تعریف ۲۶.۲.۱. مجموعه همه توابع پیوسته روی Ω را با $C^k(\Omega)$ نشان می‌دهیم. برای $k \in \mathbb{N}$ ، $C^k(\Omega)$ نشان‌دهنده توابعی هستند که همه مشتقات تا مرتبه k آنها روی Ω پیوسته است. $C^\infty(\Omega)$ کلاس همه توابعی هستند که برای هر عدد طبیعی k متعلق به $C^k(\Omega)$ باشد.

قضیه ۲۷.۲.۱. (قضیه دیورژانس). فرض کنید $\mathbf{v} = (v_1, \dots, v_n)^T$ میدان برداری روی $\Omega \subset \mathbb{R}^n$ باشد، به طوری که $v_j \in C^1(\bar{\Omega})$ و $\mathbf{v} \in C^1(\bar{\Omega}, \mathbb{R}^n)$ آنگاه:

$$\int_{\partial\Omega} \mathbf{v} \cdot \mathbf{n} ds = \int_{\Omega} \text{div} \mathbf{v} dx,$$

^۶Coercive

که در آن $\mathbf{n}$ بردار نرمال خارجی عمود بر سطح $\partial\Omega$ و ds نمایش دهنده عنصر $(n-1)$ بعدی در $\partial\Omega$ می‌باشد. به ویژه، اگر u تابعی در $C^2(\bar{\Omega})$ باشد با جایگذاری $\mathbf{v} = \Delta u$ در رابطه بالا داریم:

$$\int_{\Omega} \Delta u dx = \int_{\partial\Omega} \nabla u \cdot \mathbf{n} ds = \int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} ds,$$

تعریف ۲۸.۲.۱. (اتحادهای گرین). فرض کنید دامنه کراندار با مرز هموار و $u, v \in C^2(\bar{\Omega})$ باشد، آنگاه اگر در قضیه دیورژانس قرار دهیم $v = v \nabla u$ ، اتحاد اول گرین به صورت زیر بدست می‌آید:

$$\int_{\partial\Omega} v \frac{\partial u}{\partial \mathbf{n}} ds = \int_{\Omega} (v \Delta u + \nabla u \cdot \nabla v) dx,$$

اگر در اتحاد اول گرین $u = v$ باشد داریم:

$$\int_{\partial\Omega} u \frac{\partial u}{\partial \mathbf{n}} ds = \int_{\Omega} (u \Delta u + |\nabla u|^2) dx,$$

حال اگر $u \in C^2(\Omega) \cap C^1(\bar{\Omega})$ و روی Ω $\Delta u = 0$ باشد و داشته باشیم $u = 0$ یا $\frac{\partial u}{\partial \mathbf{n}} = 0$ روی $\partial\Omega$ ، آنگاه:

$$|\nabla u| = 0.$$

یعنی u در Ω ثابت است. اگر در اتحاد اول گرین $v = 1$ باشد داریم،

$$\int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} ds = \int_{\Omega} \Delta u dx.$$

با تعویض u و v در اتحاد اول گرین و کم کردن این دو رابطه اتحاد دوم گرین به صورت زیر بدست می‌آید:

$$\int_{\partial\Omega} \left(v \frac{\partial u}{\partial \mathbf{n}} - u \frac{\partial v}{\partial \mathbf{n}} \right) ds = \int_{\Omega} (v \Delta u - u \Delta v) dx.$$

قضیه ۲۹.۲.۱. (فرمول گرین).

$$\int_{\Omega} \Delta u v dx = \int_{\Gamma} \frac{\partial u}{\partial \mathbf{n}} v ds - \int_{\Omega} \nabla u \cdot \nabla v dx$$

تعریف ۳۰.۲.۱. (سوپریمم اساسی). فرض کنید u تابعی اندازه‌پذیر روی Ω باشد. می‌گوییم u به طور اساسی کراندار است هرگاه ثابت $\alpha \in \mathbb{R}$ وجود داشته باشد به طوری که رابطه $|u(x)| \leq \alpha$ تقریباً در Ω برقرار باشد. به بزرگترین کران پایین چنین α ‌هایی سوپریمم اساسی می‌گوییم و آن را با نماد زیر نشان می‌دهیم:

$$\operatorname{esssup}_{x \in D} |u(x)| = \inf \{ \alpha : \mu(\{x : |u(x)| > \alpha\}) = 0 \}$$

تعریف ۳۱.۲.۱. فضای L_P به شکل زیر تعریف می‌شود:

$$L_P = L_{p(\Omega)} = \{v : \|v\|_{L_P} < \infty\},$$

که در آن،

$$\|v\|_{L_P} = \begin{cases} \left(\int_{\Omega} |v(x)|^p dx \right)^{\frac{1}{p}}, & 1 \leq p < \infty, \\ \operatorname{esssup}_{\Omega} |v(x)|, & p = \infty. \end{cases}$$

در مورد L_2 ، ضرب داخلی به صورت،

$$\langle v, w \rangle = \langle v, w \rangle_{L_2(\Omega)} = \int_{\Omega} v w dx,$$

تعریف می‌شود که پایه نورم زیر است:

$$\|v\| = \|v\|_{L_2(\Omega)} = \langle v, v \rangle_{L_2(\Omega)}^{\frac{1}{2}} = \left(\int_{\Omega} v^2 dx \right)^{\frac{1}{2}}$$

قضیه ۳۲.۲.۱. (نامساوی هولدر^۷). فرض کنید $\frac{1}{p} + \frac{1}{q} = 1$ و $1 < p, q < \infty$ ، باشد. اگر $f, g \in L_1(R)$ آنگاه $f \in L_p(R)$ ، $g \in L_q(R)$ و

$$\|fg\|_{L_1} \leq \|f\|_{L_p} \|g\|_{L_q}.$$

تعریف ۳۳.۲.۱. فضای سوبولوف^۸ روی دامنه Ω برای $k \geq 1$ با نماد $H^k(\Omega)$ نمایش داده می‌شود.

$$H^k = H^k(\Omega) = \{v \in L_2(\Omega) : D^{\alpha}v \in L_2(\Omega), |v| \leq k\}$$

که $\alpha = (\alpha_1, \dots, \alpha_d)$ بردار d تایی از عناصر صحیح نامنفی است و چند اندیشه^۹ گفته می‌شود، مشتق مرتبه α ام و طول α به صورت زیر تعریف می‌شوند:

$$D^{\alpha}v = \frac{\partial^{|\alpha|}v}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}, \quad |\alpha| = \sum_{i=1}^d \alpha_i$$

ضرب داخلی برای این فضا به صورت زیر تعریف می‌شود:

$$\langle v, w \rangle_k = \langle v, w \rangle_{H^k} = \sum_{|\alpha| \leq k} \int_{\Omega} D^{\alpha}v D^{\alpha}w dx$$

نورم و نیم نورم در فضای سوبولوف به صورت زیر تعریف می‌شوند:

$$\|v\|_k = \|v\|_{H^k} = \langle v, v \rangle_k^{\frac{1}{2}} = \left(\sum_{|\alpha| \leq k} \|D^{\alpha}v\|_{L_2}^2 \right)^{\frac{1}{2}}.$$

$$|v|_k = |v|_{H^k} = \left(\sum_{|\alpha|=k} \|D^{\alpha}v\|_{L_2}^2 \right)^{\frac{1}{2}} = \|\nabla v\|.$$

در حالت‌های خاص $k = 1$ و $k = 0$ داریم:

$$H^0 = H = L_2, \quad \|v\|_{H^0} = \|v\|_{L_2}.$$

$$H^1 = \{v \in L_2, D^{\alpha}v \in L_2, |\alpha| \leq 1\}.$$

$$H_0^1 = \{v \in H^1, v|_{\Gamma} = 0\}.$$

تعریف ۳۴.۲.۱. یک مساله پایدار نامیده می‌شود هرگاه تغییرات کوچک در داده‌ها، منجر به تغییرات بزرگ در جواب نشود.

^۷Holder Inequality

^۸Sobolev Space

^۹Multi Index

تعریف ۳۵.۲.۱. (نورم اقلیدسی)

$$\|p\|_2 = \sqrt{p_1^2 + p_2^2 + \cdots + p_n^2} = \sqrt{p \cdot p}$$

p بردار اقلیدسی نامیده می‌شود.

تعریف ۳۶.۲.۱. (دامنه لیپ شیتز^{۱۰}). فرض کنید Ω زیرمجموعه باز و کرانداری از $\mathbb{R}^n$ باشد و $n \in N$ فرض شود. $\partial\Omega$ کران Ω را نشان می‌دهد. به Ω کران لیپ شیتز گفته می‌شود و همچنین دامنه لیپ شیتز نامیده می‌شود اگر برای هر نقطه $p \in \partial\Omega$ ، $r > 0$ و نگاشت $h_p : B_r(p) \rightarrow \varphi$ وجود داشته باشد به‌گونه‌ای که،

• h_p یک نگاشت است،

• h_p, h_p^{-1} هر دو توابع پیوسته لیپ شیتز هستند،

• $h_p(\partial\Omega \cap B_r(p)) = \varphi_0$.

• $h_p(\Omega \cap B_r(p)) = \varphi_+$.

که $B_r(p) := \{x \in \mathbb{R}^n \mid \|x - p\| < r\}$ ، φ_0 توابعی با مقدار صفر و φ_+ توابعی با مقدار مثبت خواهد بود.

تعریف ۳۷.۲.۱. (نامساوی فریدریچز^{۱۱})

$$\|v\|_{L^2_r(\Omega)} \leq C(\|\nabla v\|_{L^2_r(\Omega)}^2 + \|v\|_{L^2_r(\Gamma)}^2)^{\frac{1}{2}}, \quad v \in \mathcal{C}^1$$

لم ۳۸.۲.۱. (لم پوانکاره^{۱۲}) اگر $\Omega \subset \mathbb{R}^n$ ، دامنه کراندار باشد. آنگاه ثابت $C = C(\Omega)$ وجود دارد به طوری که،

$$\int_{\Omega} |u(x)|^2 dx \leq C \int_{\Omega} |\nabla u(x)|^2 dx.$$

۳.۱ معادلات دیفرانسیل با مشتقات جزئی

این معادلات، زبانی برای فرمول‌بندی بسیاری از مساله‌های مهندسی و علمی است، که در پیش‌بینی و کنترل خواص دینامیکی و استاتیکی ساختمان‌ها، گرم کردن و ذوب کردن فلزات، زدودن آلودگی هوا و آب در تسهیلات شهری، عکس برداری تشدید مغناطیسی و ... به کار می‌رود.

معادلات دیفرانسیل با مشتقات جزئی^{۱۳} دسته‌ای از معادلات دیفرانسیل هستند که در آن، توابع مجهول بر حسب چند متغیر مستقل به همراه مشتق جزئی توابع نسبت به آن متغیرها، وجود داشته باشد، یعنی

^{۱۰}Lipschitz

^{۱۱}Friedrichs inequality

^{۱۲}Poincare lemma

^{۱۳}Partial Differential Equation

هر رابطه بین متغیرهای $x_1, \dots, x_n$ و متغیر تابع u و مشتقات متغیر تابع نسبت به متغیرهای مستقل را یک معادله دیفرانسیل جزئی گویند. اگر $u = f(x_1, \dots, x_n)$ یک تابع چند متغیره باشد، مشتق مرتبه k ام نسبت به مولفه x_i را به صورت $\frac{\partial^k u}{\partial x_i^k}$ نشان می‌دهیم. هرگاه بزرگترین مرتبه مشتق ظاهر شده k باشد، معادله دیفرانسیل از مرتبه k است. به این معادلات به اختصار PDE گفته می‌شود. بیشتر PDEها فقط شامل مشتقات جزئی مرتبه اول و دوم بوده و PDEهای مرتبه دوم نامیده می‌شوند. اگر در معادله‌ای متغیر وابسته یا مشتقات آن به توان نرسیده یا در هم ضرب نشده باشد، معادله خطی، و در غیر این صورت غیر خطی است. ضرایب این معادلات می‌تواند ثابت یا متغیر باشد. فرض کنید Ω دامنه و $\mathcal{O}$ مجموعه بازی در Ω باشد. نشان می‌دهیم که $\mathcal{O} \subseteq \Omega$ و $\mathcal{O}$ کامل است. شکل PDE مرتبه دوم با n متغیر مستقل $z = (z_1, \dots, z_n)$ به صورت زیر است:

$$Au := - \sum_{i,j=1}^n \frac{\partial}{\partial z_i} (a_{ij}(z) \frac{\partial u}{\partial z_j}) + \sum_{i=1}^n (\frac{\partial}{\partial z_i} (b_i(z)u) + c_i(z) \frac{\partial u}{\partial z_i}) + a_0(z)u = f(z), \quad (1.1)$$

که $f \in C^1(\mathcal{O})$, $a_0 \in C^1(\mathcal{O})$, $c_i \in C^1(\mathcal{O})$, $b_i \in C^1(\mathcal{O})$, $a_{ij} \in C^1(\mathcal{O})$, $u \in C^2(\mathcal{O})$ و $A(z) = \{a_{ij}\}_{i,j=1}^n$ است.

در معادله (۱.۱) به ازای $z \in \mathcal{O}$ اگر تابع برداری $A(z)$ معین مثبت باشد معادله را بیضوی گوئیم، اگر نیمه معین مثبت باشد اما معین مثبت نباشد و رتبه‌ی $A(z)$ ، $b(z)$ و $c(z)$ مساوی n باشد معادله را سهموی می‌نامیم، به ازای $z \in \mathcal{O}$ اگر $A(z)$ یک مقدار ویژه‌ی منفی و $n-1$ مقدار ویژه‌ی مثبت داشته باشد معادله را هذلولوی گوئیم.

اگر معادله (۱.۱)، مستقل از زمان باشد، قرار می‌دهیم $n = d$ و $z = x$ ، که d بعد فضا و x متغیر فضا است.

تعریف ۱.۳.۱. معادله دیفرانسیل جزئی

• خطی نامیده می‌شود اگر

$$\sum_{|\alpha| \leq k} a_\alpha(x) D^\alpha u(x) = f(x),$$

که a_α با $|\alpha| \leq k$ چند اندیسی است و f تابع داده شده است. این PDE خطی همگن است اگر $f \equiv 0$.

• نیمه خطی است اگر

$$\sum_{|\alpha|=k} a_\alpha(x) D^\alpha u(x) + E_0(x, u(x), Du(x), \dots, D^{k-1}u(x)) = 0.$$

• شبه خطی است اگر

$$\sum_{|\alpha|=k} a_\alpha(x, u(x), Du(x), \dots, D^{k-1}u(x)) D^\alpha u(x) + E_0(x, u(x), Du(x), \dots, D^{k-1}u(x)) = 0.$$

در زیر چند معادله دیفرانسیل با مشتقات جزئی مرتبه دوم که کاربرد بیشتری دارند آورده شده است.

- معادله لاپلاس (خطی، همگن)

$$\Delta u = \operatorname{div} Du = \sum_{i=1}^n D_i D_i u = 0, \quad \text{در } \Omega.$$

- معادله گرما (خطی، همگن)

$$D_t u = \Delta u, \quad \text{در } \mathbb{R}^+ \times \Omega.$$

- معادله سطح مینیمال (شبه خطی)

$$\operatorname{div} \left(\frac{Du}{(1 + |Du|^2)^{\frac{1}{2}}} \right) = 0, \quad \text{در } \Omega.$$

تعریف ۲.۳.۱. مساله‌های داده شده از PDE خوش وضع است اگر

(۱) مساله جواب دقیق داشته باشد.

(۲) جواب یکتا باشد.

(۳) جواب به پیوستگی در اطلاعات داده شده در مساله بستگی داشته باشد.

۱.۳.۱ روش‌های عددی حل معادلات دیفرانسیل با مشتقات جزئی

- تفاضلات محدود

- المان‌های محدود

- حجم‌های محدود

- روش انتگرال‌گیری عددی

سه عمگر اصلی در معادلات دیفرانسیل:

- عملگر لاپلاس $\Delta_n = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$

- عملگر پخش یا انتشار $\frac{d}{dt} - \Delta_n$

- عملگر دالامبرت $\frac{d^2}{dt^2} - \Delta_n$

به طور کلی، معادلات دیفرانسیل به سه رده‌ی کلی زیر تقسیم می‌شود:

۱. مسایل مقدار اولیه (IVP)^{۱۴}

در این دسته از مسایل همواره معادله دیفرانسیل با شرط اولیه همراه است، این شرط اولیه به

شکل $u(x_0) = u_0$ است.

^{۱۴}Initial Value Problems

مثال ۳.۳.۱.

$$\begin{cases} u_{tt} - u_{xx} = 0, & -\infty < x < \infty, \quad t > 0, \\ u(x, 0) = g(x) \quad u_t(x, 0) = 0, & -\infty < x < \infty, \quad t = 0. \end{cases}$$

۲. مسایل مقدار مرزی (BVP)^{۱۵}

- در $\mathbb{R}$: یک BVP در $\mathbb{R}$ باید در مقادیر روی مرز، یعنی نقطه ابتدا و انتهای بازه تعریف و مقدار u مشخص باشد.

مثال ۴.۳.۱.

$$\begin{cases} -(a(x)u'(x))' = f, & 0 < x < 1, \\ u(0) = u(1) = 0, & x \in \partial\Omega. \end{cases}$$

- در $\mathbb{R}^n$: مقدار u باید روی سراسر مرز ناحیه $\Omega \subset \mathbb{R}^n$ که آن را با نماد $\partial\Omega$ نشان می‌دهیم، معلوم باشد.

مثال ۵.۳.۱.

$$\begin{cases} u - \Delta u = 0, & x \in \Omega, \quad t > 0, \\ u(x, t) = g(x), & x \in \partial\Omega, \quad t > 0. \end{cases}$$

برای مسایل مقدار مرزی ۳ نوع شرط داریم:

۱- شرط مرزی دیریکله^{۱۶}

مثال ۶.۳.۱.

$$u(x, t) = g(x), \quad x \in \partial\Omega,$$

۲- شرط مرزی نوین^{۱۷}

مثال ۷.۳.۱.

$$\frac{\partial u}{\partial \mathbf{n}} = \mathbf{n} \cdot \nabla u = g(x) \quad x \in \partial\Omega,$$

که $\mathbf{n}$ بردار یکه‌ی خارجی عمود بر Ω در $x \in \partial\Omega$ است.

۳- شرط مرزی روبین^{۱۸}

مثال ۸.۳.۱.

$$\frac{\partial u}{\partial \mathbf{n}} + k \cdot u(x, t) = g(x), \quad k > 0, \quad x \in \partial\Omega,$$

۳. مسایل مقدار ویژه

^{۱۵}Boundary Value Problems

^{۱۶}Dirichlet Boundary Condition

^{۱۷}Neumanns Boundary Condition

^{۱۸}Robin Boundary Condition

تعریف ۹.۳.۱. (روش گالرکین). اساس روش گالرکین مبتنی بر یافتن تقریبی برای جواب است به گونه‌ای که:

- به سادگی بتوان از آن مشتق گیری یا انتگرال گرفت.
- توسط مجموعه‌ای از توابع پایه‌ای تقریباً متعامد با بعد متناهی تولید شوند.

۴.۱ روش عناصر متناهی

امروزه استفاده از روش‌های عددی به ویژه روش عناصر متناهی، در حل مسایل متفاوت کاربرد فراوانی دارد. چون این روش‌ها تقریبی هستند، داشتن اطلاعاتی از میزان و نحوه‌ی توزیع خطاها اهمیت بسیار زیادی دارد. روش عناصر متناهی یا روش المان محدود که به اختصار FEM^{۱۹} نامیده می‌شود، روشی عددی برای حل تقریبی معادلات دیفرانسیل جزئی و حل معادله‌های انتگرالی است. در روش عناصر متناهی، محیط پیوسته به اجزای هندسی ساده و کوچکتری که المان^{۲۰} نامیده می‌شود تقسیم می‌شود. این عمل را گسسته سازی^{۲۱} و نقطه مشترک المان‌ها را گره^{۲۲} می‌نامند. در روش عناصر متناهی به دو فضا نیاز داریم که به صورت زیر هستند:

فضای آزمایش^{۲۳}: مجموعه‌ای از توابع، که جواب تقریبی را در آن جستجو می‌کنیم.

فضای آزمون^{۲۴}: فضای تابعی که برای شرط تعامد از آن استفاده می‌کنیم.

۵.۱ تحلیل روش عناصر متناهی برای مسایل دو بعدی

مسایل دو بعدی با معادلات مشتقات جزئی مطرح می‌شوند. در حالت کلی مرز در اینگونه مسایل، یک منحنی است. لذا نه تنها دنبال جوابی تقریبی برای معادله هستیم بلکه باید دامنه را با مش مناسب تقریب بزنیم. تونایی نمایش دامنه‌های نامنظم مهندسی، با مجموعه‌ای از عناصر متناهی باعث می‌شود که روش به عنوان ابزار علمی با ارزش برای حل مسایل مقدار مرزی مطرح باشد. فرض کنیم مساله پیدا کردن جواب u از معادله مشتق جزئی مرتبه دوم،

$$-\frac{\partial}{\partial x}(a_{11}\frac{\partial u}{\partial x} + a_{12}\frac{\partial u}{\partial y}) - \frac{\partial}{\partial y}(a_{21}\frac{\partial u}{\partial x} + a_{22}\frac{\partial u}{\partial y}) + a_{00}u - f = 0 \quad (2.1)$$

^{۱۹}Finite element method

^{۲۰}Element

^{۲۱}Discretization

^{۲۲}Joint

^{۲۳}Test Space

^{۲۴}Trial Space

به ازای (a_{ij}) ، a_{∞} ، f و شرایط مرزی مشخص داده شده می‌باشد.
 اگر $a_{11} = a_{22} = a$ و $a_{12} = a_{21} = a_{\infty} = 0$ باشد، آنگاه معادله پواسون را داریم:

$$-\nabla(a\nabla u) = f \quad \Omega \text{ در.} \quad (3.1)$$

۱.۵.۱ تجزیه دامنه

دامنه $\bar{\Omega} = \Omega \cup \partial\Omega$ به N زیرمجموعه $\bar{\Omega}^e$ که عناصر متناهی نامیده می‌شوند، (المان‌ها دارای شکل‌های هندسی ساده مانند مثلث، مستطیل و ... هستند) تقسیم شده بطوری که،

۱. هر یک از $\bar{\Omega}^e$ بسته و ناتهی است.

$$(e \neq f) \quad \Omega^e \cap \Omega^f = \emptyset \quad ۱.$$

$$\bar{\Omega} \approx \bar{\Omega}_h = \bigcup_{e=1}^N \bar{\Omega}^e \quad ۱.$$

h اندازه مش است. طول ضلع در هر المان حداکثر برابر h است.

۲.۵.۱ معادلات المان

فرض کنیم Ω^e یک المان باشد. به منظور بدست آوردن فرم وردشی، طرفین (۲.۱) را در تابع آزمایش $(v \in S^h \subset H_0^1(\Omega))$ ضرب کرده و روی Ω^e انتگرال‌گیری می‌کنیم.

$$0 = \int_{\Omega^e} v \left[-\frac{\partial}{\partial x}(F_1) - \frac{\partial}{\partial y}(F_2) + a_{\infty}u - f \right] dx dy, \quad (4.1)$$

که

$$F_1 = a_{11} \frac{\partial u}{\partial x} + a_{12} \frac{\partial u}{\partial y}, \quad F_2 = a_{21} \frac{\partial u}{\partial x} + a_{22} \frac{\partial u}{\partial y}. \quad (5.1)$$

با استفاده از انتگرال‌گیری جزیه‌جز و قضیه دیورژانس داریم،

$$\begin{aligned} 0 &= \int_{\Omega^e} \left[\frac{\partial v}{\partial x} \left(a_{11} \frac{\partial u}{\partial x} + a_{12} \frac{\partial u}{\partial y} \right) - \frac{\partial v}{\partial y} \left(a_{21} \frac{\partial u}{\partial x} + a_{22} \frac{\partial u}{\partial y} \right) + a_{\infty}vu - vf \right] dx dy \\ &\quad - \int_{\Gamma^e} v \left[n_x \left(a_{11} \frac{\partial u}{\partial x} + a_{12} \frac{\partial u}{\partial y} \right) - n_y \left(a_{21} \frac{\partial u}{\partial x} + a_{22} \frac{\partial u}{\partial y} \right) \right] ds. \end{aligned} \quad (6.1)$$

در این صورت مقدار مشخص u شرط مرزی اساسی (u متغیر اولیه) و

$$q_n = n_x \left(a_{11} \frac{\partial u}{\partial x} + a_{12} \frac{\partial u}{\partial y} \right) + n_y \left(a_{21} \frac{\partial u}{\partial x} + a_{22} \frac{\partial u}{\partial y} \right). \quad (7.1)$$

شرط مرزی طبیعی (q_n متغیر ثانویه) را نشان می‌دهند.

تابع $q_n = q_n(x)$ تصویر بردار $F = F_1 i \rightarrow F_2 j$ در طول بردار نرمال $\mathbf{n}$ است. اگر $a_{11} = a_{22} = a$ و $a_{12} = a_{21} = 0$ آنگاه q_n مساوی حاصل ضرب و مشتق جهتی در جهت بردار نرمال است. در بیشتر مسائلی از نظر فیزیکی اهمیت دارند.

برای مثال در حالت انتقال حرارت، جریان گرمای گذرنده از مرز المان را نشان می‌دهد. بنابراین فرم تغییراتی معادله عبارتست از:

$$\begin{aligned} \circ &= \int_{\Omega^e} \left[\frac{\partial v}{\partial x} \left(a_{11} \frac{\partial u}{\partial x} + a_{12} \frac{\partial u}{\partial y} \right) - \frac{\partial v}{\partial y} \left(a_{21} \frac{\partial u}{\partial x} + a_{22} \frac{\partial u}{\partial y} \right) + a_{33} v u - v f \right] dx dy \\ &- \int_{\Gamma} v q_n ds \end{aligned} \quad (۸.۱)$$

این معادله اساس مدل FE از معادله اصلی را تشکیل می‌دهد. فرض می‌کنیم:

$$u = \sum_{j=1}^n u_j \psi_j, \quad (۹.۱)$$

که u_j مقادیر u در (x_j, y_j) و ψ_j توابع درونیاب با ویژگی زیر باشند.

$$\sum_{j=1}^n \psi_j = 1, \quad \psi_j(x_i, y_i) = \delta_{ij}, \quad (۱۰.۱)$$

معادله (۹.۱) را به جای u و ψ_j را به جای v در معادله (۸.۱) جایگزین می‌کنیم، داریم:

$$\begin{aligned} \circ &= \sum_{j=1}^n \left\{ \int_{\Omega^e} \left[\frac{\partial \psi_i}{\partial x} \left(a_{11} \frac{\partial \psi_j}{\partial x} + a_{12} \frac{\partial \psi_j}{\partial y} \right) + \frac{\partial \psi_i}{\partial y} \left(a_{21} \frac{\partial \psi_j}{\partial x} + a_{22} \frac{\partial \psi_j}{\partial y} \right) + a_{33} \psi_i \psi_j \right] dx dy \right\} \\ &u_j - \int_{\Omega^e} f \psi_i dx dy - \int_{\Gamma^e} \psi_i q_n ds, \quad i = 1, 2, \dots, n, \end{aligned} \quad (۱۱.۱)$$

که،

$$\begin{aligned} \sum_{i=1}^n k_{ij}^{(e)} u_j^{(e)} &= F_i^{(e)}, \quad ([k^{(e)}] \{u^{(e)}\} = \{F^{(e)}\}), \\ k_{ij}^{(e)} &= \int_{\Omega^e} \left[\frac{\partial \psi_i}{\partial x} \left(a_{11} \frac{\partial \psi_j}{\partial x} + a_{12} \frac{\partial \psi_j}{\partial y} \right) - \frac{\partial \psi_i}{\partial y} \left(a_{21} \frac{\partial \psi_j}{\partial x} + a_{22} \frac{\partial \psi_j}{\partial y} \right) + a_{33} \psi_i \psi_j \right], \\ F_i^{(e)} &= \int_{\Omega^e} f \psi_i dx dy + \int_{\Gamma^e} q_n \psi_i ds = f_i^{(e)} + Q_i^{(e)}. \end{aligned} \quad (۱۲.۱)$$

اگر $a_{12} = a_{21}$ ، آنگاه $[k]$ متقارن است. یعنی $k_{ij}^{(e)} = k_{ji}^{(e)}$.

۳.۵.۱ محاسبه ماتریس‌های المان

محاسبه ماتریس‌های $[k^{(e)}]$ و $[F^{(e)}]$ در معادلات (۱۲.۱) به وسیله روش‌های انتگرال‌گیری دقیق، در حالت کلی ساده نیست. در هر صورت، اگر a_{ij} ، a_{33} و f ثابت باشند، می‌توان انتگرال‌ها را روی المان‌های مثلثی و مستطیلی ارزیابی نمود.

برای سادگی ماتریس $[k^{(e)}]$ را به صورت چهار ماتریس اساسی $[S^{\alpha\beta}]$ ($\alpha, \beta = 1, 2$) و $[S]$ می‌نویسیم:

$$[k^{(e)}] = a_{11}[S^{11}] + a_{12}[S^{12}] + a_{21}[S^{21}] + a_{22}[S^{22}] + a_{33}[S], \quad (۱۳.۱)$$

که،

$$\begin{aligned} S_{ij}^{11} &= \int_{\Omega^e} \frac{\partial \psi_i}{\partial x} \frac{\partial \psi_j}{\partial x} dx dy, & S_{ij}^{12} &= \int_{\Omega^e} \frac{\partial \psi_i}{\partial x} \frac{\partial \psi_j}{\partial y} dx dy, \\ S_{ij}^{22} &= \int_{\Omega^e} \frac{\partial \psi_i}{\partial y} \frac{\partial \psi_j}{\partial y} dx dy, & S_{ij} &= \int_{\Omega^e} \psi_i \psi_j dx dy, \end{aligned} \quad (14.1)$$

همچنین،

$$f_i^{(e)} = \int_{\Omega^e} f \psi_i dx dy, \quad Q_i^{(e)} = \int_{\Gamma^e} q_n \psi_i ds. \quad (15.1)$$

اگر $a_{21} = a_{12}$ ، ماتریس‌های المان $[k^{(e)}]$ متقارن هستند. در ضمن این ماتریس‌ها منفرد بوده و معادلات (۱۱.۱) قابل حل نمی‌باشد.

۴.۵.۱ ساخت ماتریس‌های المان

دیدیم که معادله (۱۱.۱) برای یک المان نوعی برقرار است. اینک می‌خواهیم با کنار هم گذاشتن این معادلات، معادله زیر را در سراسر ناحیه بدست آوریم:

$$[k]\{u\} = \{F\}. \quad (16.1)$$

u روی مرزهای درون المانی پیوسته است. بنابراین می‌توان بین شماره‌های گره‌های موضعی^{۲۵} و شماره‌های گره‌های عمومی^{۲۶}، یک تناظر برقرار نمود. این تناظر را می‌توان به صورت آرایه‌ای نشان داد که به آن ماتریس ارتباط می‌گویند. همچنین روابط بین مقادیر u در گره‌های موضعی (u_i) و عمومی (U_i) را شرایط پیوستگی درون المانی نامند.

۵.۵.۱ حل معادلات تولید شده

برای حل دستگاه بدست آمده در (۱۰.۱) به صورت زیر عمل می‌کنیم:

$$\begin{bmatrix} [k^{11}] & [k^{12}] \\ [k^{21}] & [k^{22}] \end{bmatrix} \begin{Bmatrix} \{\Delta^1\} \\ \{\Delta^2\} \end{Bmatrix} = \begin{Bmatrix} \{F^1\} \\ \{F^2\} \end{Bmatrix},$$

که $\{\Delta^1\}$ ستون متغیرهای اولیه (U_i) معلوم و $\{\Delta^2\}$ ستون متغیرهای اولیه مجهول و $\{F^1\}$ ستون نیروهای مجهول و $\{F^2\}$ ستون نیروهای معلوم می‌باشد. به این ترتیب،

$$U(x) = \sum_{e=1}^N \sum_{i=1}^n u_i^{(e)} \psi_i^{(e)} = \sum_{I=1}^M U_I \phi_i, \quad (17.1)$$

می‌توانیم مقادیر U را در نقاط مختلف دامنه محاسبه کنیم.

۶.۱ انواع خطاها

مهم‌ترین خطاهایی که در حل یک مساله روش عناصر متناهی وجود دارد به صورت زیر دسته‌بندی می‌گردد:

^{۲۵}Local

^{۲۶}Global

۱- خطای ناشی از نامعلومی اطلاعات ورودی: شامل تمام خطاهای ناشی از هندسه مساله و شرایط مرزی می‌باشد.

۲- خطای ناشی از گسسته سازی: شامل خطای ناشی از انتخاب شبکه المان و تعریف شرایط مرزی است. در واقع این خطا به اندازه المان‌ها بستگی دارد. در حل مسایل مهندسی با شبکه المان نامناسب حتی تا صد درصد خطای محاسباتی مشاهده شده است که ناشی از عدم شناخت نواحی بحرانی است.

۳- خطای ناشی از گرد کردن: این خطا، ناشی از گرد کردن اعداد توسط کامپیوتر است که مقدار آن به دقت مثبت اعداد در کامپیوتر بستگی دارد.

شبکه المان مناسب، شبکه‌ای است که خطای ناشی از گسسته سازی در آن ناچیز و بسیار کم باشد. این خطا عبارت از اختلاف بین جواب دقیق و جواب عددی است که به روش عناصر متناهی به دست آمده است. برای بیشتر مسایلی که معادله دیفرانسیل حاکم بر آن پاره‌ای است، حل دقیق قابل دسترسی نیست. بنابراین خطای گسسته سازی به صورت تقریبی محاسبه می‌گردد. اهمیت این نوع خطا می‌تواند به وسیله نتیجه اشتباه حل یک مساله به روش عناصر متناهی درک شود.

به طور معمول کیفیت و درستی روش‌های تعیین خطا به وسیله مقایسه بین معیار خاصی از خطای واقعی و خطای تقریبی بررسی می‌گردد. البته امکان محاسبه خطای واقعی برای برخی مسایل خاص که حل واقعی آن در دسترس است وجود دارد.

با استفاده از مقدار خطای محاسبه شده توسط یکی از روش‌های برآورد خطا می‌توان شبکه المان جدیدی ایجاد کرد و با استفاده از این شبکه المان دوباره مساله را به روش عناصر متناهی آنالیز نمود و مقدار خطا را محاسبه کرد. این فرایند تا زمانی که خطا به حد مجاز خود برسد، ادامه می‌یابد. به این روش، آنالیز تطبیقی^{۲۷} گفته می‌شود. برای بررسی نرخ^{۲۸} همگرایی می‌توان اندازه المانها را کاهش داد و خطا را محاسبه نمود.

^{۲۷}Adaptivity

^{۲۸}Rate

فصل ۲

مسائل بیضوی

۱.۲ مقدمه

معادلات بیضوی در بسیاری از مدل‌ها ایجاد می‌شوند. به طور مثال، چندین دامنه پتانسیل (گرانشی، الکتریسیته ساکن، الکترومغناطیس و ...)، جریان گرمای پایا، پدیده‌های بیولوژی و همچنین پدیده‌های حالت پایدار را توصیف می‌کنند. برای گردآوری این فصل از مراجع [۱]، [۱۸] استفاده شده است.

تعریف ۱.۱.۲. فرض کنید H فضای هیلبرت باشد. فرم دوخطی $a : H \times H \rightarrow \mathbb{R}$ بیضوی نامیده می‌شود اگر کراندار، متقارن و کواریسو باشد.

تعریف ۲.۱.۲. (عملگرهای بیضوی). فرض کنید Ω ناحیه هموار و کراندار در $\mathbb{R}^n$ باشد. مساله‌ای به صورت زیر را در نظر بگیرید:

$$\begin{cases} Lu(x) = f(x), & x \in \Omega, \\ Bu(x) = 0, & x \in \partial\Omega, \end{cases} \quad (1.2)$$

که در آن عملگر دیفرانسیلی L به صورت زیر تعریف می‌شود:

$$L = - \sum_{i,j=1}^N a_{ij} \partial_i \partial_j + \sum_{i=1}^N a_i \partial_i + c \quad (2.2)$$

عملگر فوق را برای هر بردار حقیقی $(\xi_1, \dots, \xi_N)^T$ در نقطه $x = (x_1, \dots, x_n)^T$ بیضوی گوئیم اگر و تنها اگر ضریب مثبت $\mu(x)$ موجود باشد به‌طوری‌که

$$\sum_{i,j=1}^N a_{ij}(x) \xi_i \xi_j \geq \mu(x) \sum_{i=1}^N \xi_i^2.$$

عملگر L را بر Ω بیضوی گوئیم هرگاه بر هر نقطه از Ω بیضوی باشد. همچنین این عملگر را بیضوی یکنواخت می‌نامیم اگر در هر نقطه از Ω بیضوی بوده و ثابت μ_0 به طوری موجود باشد که برای هر $x \in \Omega$

نامساوی $\mu(x) \geq \mu_0$ برقرار باشد.

عملگر بیضوی یکنواخت را از نوع دیورژانس گوئیم اگر

$$Lu = \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial}{\partial x_j} u) + cu,$$

با ضرایب کراندار $a_{ij} = a_{ji}$ باشد.

عملگر مرزی B به صورت زیر تعریف می‌شود:

$$B = b_0 + b_1 \frac{\partial}{\partial \mathbf{n}}.$$

$\frac{\partial}{\partial \mathbf{n}}$ بیانگر مشتق در جهت بردار یکه نرمال خارجی بر Ω می‌باشد. در حالتی که $b_1 = 0$ و $b_0 = 1$ عملگر (۱.۲) را معادله بیضوی با شرط مرزی دیریکله، $b_1 = 1$ و $b_0 = 1$ عملگر (۱.۲) را معادله بیضوی با شرط مرزی نیومن و اگر $b_1 = 1$ و $b_0 \in \mathbb{R}$ آن را معادله بیضوی با شرط مرزی نوع سوم یا رابین گوئیم.

۱.۱.۲ مثال‌هایی از معادله‌های بیضوی

• معادله لاپلاس^۱

$$\Delta u := \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad \text{در } \Omega.$$

• معادله پواسون^۲

$$-\Delta u = f \quad \text{در } \Omega.$$

• معادله انتشار واکنش^۳

$$-\Delta u + cu = f \quad \text{در } \Omega.$$

• معادله انتشار همرفت (انتقال گرما)^۴

$$-\varepsilon \Delta u + b_1 \frac{\partial u}{\partial x} + b_2 \frac{\partial u}{\partial y} = f \quad \text{در } \Omega, \quad \varepsilon > 0.$$

ابتدا مساله مقدار مرزی را برای معادله دیفرانسیل جزئی بیضوی بررسی می‌کنیم. در حالت خاص، ماتریس $A(x) = a_{ij}(x)$ به ضرب اسکالر aI از ماتریس یکتا تبدیل می‌شود که $A(x) = a_{ij}(x)$ ماتریس معین مثبت هموار است.

ابتدا مساله دیریکله را در نظر بگیرید:

$$\begin{cases} \mathcal{A}u := -\nabla \cdot (a \nabla u) + b \cdot \nabla u + cu = f, & x \in \Omega, \\ u = g, & x \in \Gamma, \end{cases} \quad (3.2)$$

که در آن،

^۱Laplas Equation

^۲Poisson Equation

^۳Reaction-Diffusion Equation

^۴Heat Transfer

• $\Omega \subset \mathbb{R}^d$ دامنه‌ای با مرز هموار Γ است.

• ضرایب $a = a(x)$, $b = b(x)$, $c = c(x)$ هموار بوده و در روابط زیر صدق می‌کنند:

$$a(x) \geq a_* > 0, \quad c(x) \geq 0, \quad \forall x \in \Omega. \quad (۴.۲)$$

• f, g توابع داده شده‌ای روی Γ و Ω هستند.

در حالت خاص که $a = 1$, $b = 0$ و $c = 0$ است معادله پواسون را خواهیم داشت یعنی،

$$-\Delta u := -\sum_{j=1}^d \frac{\partial^2 u}{\partial x_j^2} = f. \quad (۵.۲)$$

زمانی که $f = 0$ است این معادله به معادله لاپلاس تبدیل می‌شود و جواب آن، توابع هارمونیک نامیده می‌شود.

قضیه ۳.۱.۲. فرض کنید $u \in C^2(\bar{\Omega})$ ، آنگاه مقدار ثابت c به گونه‌ای موجود است که،

$$\|u\|_{C(\bar{\Omega})} \leq \|u\|_{C(\Gamma)} + c\|Au\|_{C(\bar{\Omega})}.$$

برهان. فرض کنید ϕ تابعی باشد که

$$\phi \geq 0, \quad A\phi \leq -1, \quad x \in \Omega,$$

دو تابع به فرم زیر را تعریف می‌کنیم:

$$v_{\pm}(x) = \pm u(x) + \|Au\|_{C(\bar{\Omega})}\phi(x),$$

بنابراین،

$$Av_{\pm} = \pm Au + \|Au\|_{C(\bar{\Omega})}A\phi(x) \leq 0, \quad x \in \Omega,$$

بنابراین $v_{\pm}$ ماکزیمم خود را روی Γ خواهد داشت:

$$v_{\pm}(x) \leq \max_{\Gamma}(v_{\pm}) \leq \max_{\Gamma}(\pm u) + \|Au\|_{C(\bar{\Omega})}\|\phi\|_{C(\Gamma)} \leq \|u\|_{C(\Gamma)} + c\|Au\|_{C(\bar{\Omega})},$$

□

با انتخاب $c = \|\phi\|_{C(\Gamma)}$ حکم ثابت می‌شود.

۲.۱.۲ جواب اساسی و تابع گرین

معادله ناهمگن $Au = f$ را در نظر بگیرید که در آن

$$A = -\nabla \cdot (a\nabla) + b \cdot \nabla + c,$$

است. در حالت خاص $b = 0$ فرض می‌شود،

$$A = -\nabla \cdot (a\nabla) + c.$$

برای شروع، معادله $Au = f$ را در $C_0^\infty(\mathbb{R}^d)$ ضرب کرده و روی Ω انتگرال گیری کنید.

$$\langle Au, \varphi \rangle = \langle f, \varphi \rangle.$$

با دو بار انتگرال گیری جزیه جز از سمت چپ تساوی داریم،

$$\begin{aligned}
 \int_{\Omega} Au\varphi dx &= \int_{\Omega} -\nabla \cdot (a\nabla u)\varphi dx + \int_{\Omega} cu\varphi dx \\
 &= -\varphi(a\nabla u)|_{\Gamma} + \int_{\Omega} a\nabla u \nabla \varphi dx + \int_{\Omega} cu\varphi dx \\
 &= \int_{\Omega} a\nabla u \nabla \varphi dx + \int_{\Omega} cu\varphi dx \\
 &= a\nabla \varphi u|_{\Gamma} - \int_{\Omega} \nabla \cdot (a\nabla \varphi)u dx + \int_{\Omega} cu\varphi dx \\
 &= - \int_{\Omega} \nabla \cdot (a\nabla \varphi)u dx + \int_{\Omega} cu\varphi dx \\
 &= \int_{\Omega} (-\nabla \cdot (a\nabla \varphi) + c\varphi)u dx \\
 &= \langle u, A\varphi \rangle.
 \end{aligned}$$

بنابراین خواهیم داشت،

$$\langle u, A\varphi \rangle = \langle f, \varphi \rangle = \int_{R^d} f(x)\varphi(x)dx, \quad \forall \varphi \in C_0^\infty(R^d).$$

تعریف ۴.۱.۲. u جواب اساسی معادله $Au = f$ است، هرگاه برای $x \neq \circ$ هموار بوده و در $x = \circ$ دارای تکینگی (یعنی مقدار تابع در نقطه تکین تعریف نشده باشد) باشد به گونه‌ای که،

$$u \in L_1(B), \quad B = \{x \in R^d : |x| < 1\} \cdot$$

$$|D^\alpha u(x)| \leq c_\alpha |x|^{2-d-|\alpha|}, \quad |\alpha| \neq \circ \cdot$$

$$\langle u, A\varphi \rangle = \varphi(\circ), \quad \forall \varphi \in C_0^\infty(R^d) \cdot$$

۲.۲ روش عناصر متناهی برای معادله‌های بیضوی

هم گام با رشد علوم و فناوری، مسایل مهندسی نیز روز به روز پیچیده‌تر می‌شوند. با پیچیده‌تر شدن مسایل و لزوم حل سریع‌تر و دقیق‌تر آن‌ها، روش‌های تحلیلی، دیگر جواب‌گوی نیازهای روز افزون جوامع نیستند. با چنین نگرشی، محققان همواره سعی می‌کنند در کنار توسعه مبانی علوم، روش‌های عددی را نیز توسعه بخشند.

روش عناصر متناهی، یکی از روش‌هایی است که کاربرد فراوانی در حل مسایل بسیاری از رشته‌های مهندسی و به خصوص مسایل مکانیک جامدات دارد. همچنین این روش برای حل معادله‌های بیضوی به کار برده می‌شود.

۱.۲.۲ مسأله‌ای در صفحه

مسأله زیر را در نظر بگیرید:

$$Au = -\nabla(a\nabla u) = f, \quad x \in \Omega, \quad \text{و} \quad u = \circ, \quad x \in \Gamma,$$

که در آن $\Omega \subset \mathbb{R}^2$ دامنه‌ای به شکل چندضلعی است و $\Gamma = \partial\Omega$ کرانی برای آن می‌باشد. همچنین فرض می‌کنیم که $a = a(x)$ هموار، $a(x) \geq a_0 > 0$ و $f \in L_2$ است. فرمول‌بندی وردشی، شکل ضعیف یا به اختصار VF^5 این مسأله به شکل زیر می‌باشد: $u \in H_0^1$ را به گونه‌ای پیدا کنید که

$$a(u, v) = \langle f, v \rangle, \quad \forall v \in H_0^1(\Omega),$$

که در آن

$$a(w, v) = \int a \nabla v \cdot \nabla w dx, \quad \langle f, v \rangle = \int f v dx.$$

این مسأله جواب یکتایی در H_0^1 دارد. همچنین می‌خواهیم جواب تقریبی برای معادله به دست آوریم. به این منظور ابتدا باید ناحیه چندضلعی Ω را مثلثی سازی^۶ کنیم. فرض کنید $\mathcal{T}_h = \{T\}$ مجموعه‌ای از مثلث‌های بسته باشد به این مجموعه مثلثی سازی گفته می‌شود هرگاه

$$\bar{\Omega} = \bigcup_{T \in \mathcal{T}_h} T, \quad h_T = \text{diam}(T), \quad h = \max_{T \in \mathcal{T}_h} h_T.$$

رئوس p از مثلث‌های $T \in \mathcal{T}_n$ گره‌های مثلثی سازی نامیده می‌شود. اشتراک دو مثلث می‌تواند یکی از سه حالت تهی، گره یا یال مشترک باشد. لازم به ذکر است هیچ گره‌ای نمی‌تواند داخل یک یال قرار گیرد.

۲.۲.۲ گسسته سازی

فضای S_h را به شکل زیر تعریف می‌کنیم:

$$S_h = \{v \in \mathcal{C}(\bar{\Omega}) : v = 0 \text{ on } \Gamma\}.$$

این فضا، زیر فضایی از H_0^1 است. همچنین فرض کنید که $\{p_i\}_{i=0}^{M_h}$ مجموعه نقاط گره‌ای درونی (نه روی مرز) باشند. لذا تابعی در S_h با مقادیر آن در p_j و توابع هرمی $\{\varphi_i\}_{i=0}^{M_h}$ که به شکل

$$\varphi_i(p_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

تعریف می‌شود، مشخص می‌گردد. این توابع تشکیل پایه‌ای برای S_h می‌دهند، لذا هر عضو S_h را می‌توان به صورت ترکیب خطی

$$v(x) = \sum_{i=1}^{M_h} v_i \varphi_i(x),$$

نوشت، که در آن $v_i = v(p_i)$ می‌باشد.

حال می‌توان مسأله عناصر متناهی را به شکل زیر نوشت:

^۵Varitional Formulation

^۶Triangulation

$u_h \in S_h$ را به گونه‌ای بیابید که:

$$a(u_h, \chi) = \langle f, \chi \rangle, \quad \forall \chi \in S_h.$$

چون $u_h \in S_h$ پس،

$$u_h(\chi) = \sum_{i=1}^{M_h} u_i \varphi_i(x)$$

با قرار دادن این جواب در مسأله عناصر متناهی داریم:

$$a\left(\sum_{i=1}^{M_h} u_i \varphi_i(x), \chi\right) = \langle f, \chi \rangle \quad \forall \chi \in S_h.$$

با انتخاب $\chi = \varphi_i$, $i = 1, 2, \dots, M_h$ داریم:

$$\sum_{j=1}^{M_h} a(u_j \varphi_j, \varphi_i) = \langle f, \varphi_i \rangle \quad i = 1, 2, \dots, M_h,$$

$$\sum_{j=1}^{M_h} u_j a(\varphi_j, \varphi_i) = \langle f, \varphi_i \rangle \quad i = 1, 2, \dots, M_h.$$

شکل ماتریسی این معادلات به صورت زیر است:

$$A\mathbf{u} = \mathbf{b}$$

که در آن،

$$\mathbf{u} = (u_j), \quad A = (a_{ij}) = a(\varphi_j, \varphi_i), \quad \mathbf{b} = (b_i) = \langle f, \varphi_i \rangle$$

A ماتریس سختی و $\mathbf{b}$ بردار بار است. چون ماتریس A متقارن و مثبت معین است، لذا دستگاه $A\mathbf{u} = \mathbf{b}$ جواب یکتایی دارد.

اگر مش ظریف انتخاب شود، آنگاه A ماتریس اسپارس (ماتریسی که بیشتر عناصرش صفر باشد) و بزرگ خواهد بود. در ماتریس اسپارس A تمام عناصر مساوی صفر است مگر آنکه p_i, p_j همسایه باشند. در هنگام مثلی سازی اگر مرز به شکل چند ضلعی نباشد، ولی دامنه محدب باشد، می‌توان نقاطی را روی مرز انتخاب کرد تا مرز به شکل چندضلعی در آید. می‌توان با انتخاب تعداد نقاط در درون دامنه، به مثلی سازی اقدام کرد. هر چه مش ظریف‌تر باشد جواب تقریبی دقیق‌تر است. فرض کنید:

- $\tilde{S}_h$ مجموعه‌ای از توابع قطعه وار خطی پیوسته روی T_h باشد، که لزوماً روی مرز Γ صفر نیستند.
- $\{p_j\}_{j=1}^{N_h}$ مجموعه تمام گره‌های T_h هستند. این مجموعه نقاط گره‌ای واقع بر مرز Γ را نیز شامل می‌شود.
- $\{\varphi_j\}_{j=1}^{N_h}$ توابع هرمی متناظر با این نقاط هستند.

تعریف ۱.۲.۲. عملگر درونیاب $\tilde{I}_h : C(\bar{\Omega}) \rightarrow \tilde{S}_h$ به شکل زیر تعریف می‌شود:

$$(I_h v)(x) = \sum_{i=1}^{N_h} v_i \varphi_i(x), \quad v_i = v(p_i)$$

تعریف ۲.۲.۲. تصویر متعامد^۷ که به شکل $P_h : L_2(\Omega) \rightarrow S_h$ نمایش داده می‌شود، به صورت زیر تعریف می‌شود:

$$\langle P_h v - v, \chi \rangle = 0, \quad \forall \chi \in S_h, \quad \forall v \in L_2(\Omega).$$

۳.۲ برآورد خطا در نورم H^1

قضیه ۱.۳.۲. فرض کنید که u و u_h جواب‌های

$$a(u_h, \chi) = \langle f, \chi \rangle, \quad \forall \chi \in S_h \quad (۶.۲)$$

$$a(u, v) = \langle f, v \rangle, \quad \forall v \in H_0^1 \quad (۷.۲)$$

باشند، آنگاه

$$\|u_h - u\|_a = \min_{\chi \in S_h} \|\chi - u\| \bullet$$

$$|u_h - u|_1 \leq ch \|u\|_2 \bullet$$

از قبل می‌دانیم که نورم انرژی به صورت زیر تعریف می‌شود:

$$\|v\|_a = a(v, v)^{\frac{1}{2}} = \left(\int_{\Omega} a |\nabla v|^2 dx \right)^{\frac{1}{2}}$$

برهان. چون $S_h \subset H_0^1$ است، لذا در (۷.۲) قرار می‌دهیم، $v = \chi \in S_h$ تا رابطه $a(u, \chi) = \langle f, \chi \rangle$ حاصل شود. و همچنین با کم کردن رابطه اخیر از (۶.۲) داریم:

$$a(u_h - u, \chi) = 0, \quad \forall \chi \in S_h.$$

با توجه به تعریف P_h در می‌یابیم که $u_h = P_h u$ است. به این معنا که، u_h تصویر متعامد u نسبت به $a(\cdot, \cdot)$ در S_h است.

برای اثبات نامساوی اول داریم،

$$\|u_h - u\|_a^2 = a(u_h - u, u_h - u) = a(u_h - u, u_h) - a(u_h - u, u)$$

$$a(u_h - u, \chi) \text{ را به رابطه بالا اضافه و کم می‌کنیم}$$

$$\|u_h - u\|_a^2 = a(u_h - u, u_h) - a(u_h - u, u) + a(u_h - u, \chi) - a(u_h - u, \chi)$$

$$= a(u_h - u, u_h - \chi) + a(u_h - u, \chi - u)$$

$$= a(u_h - u, \chi - u)$$

$$\leq \|u_h - u\|_a \|\chi - u\|_a$$

بنابراین داریم،

$$\|u_h - u\|_a \leq \|\chi - u\|_a, \quad \forall \chi \in S_h$$

^۷Orthogonal Projection

که در آن،

$$\|u_h - u\|_a = \min_{\chi \in S_h} \|\chi - u\|_a.$$

برای اثبات نامساوی دوم توجه کنید که $|\cdot|_1$ و $\|\cdot\|_a$ هم ارز هستند یعنی،

$$c|v|_1 \leq \|v\|_a \leq C|v|_1.$$

در نامساوی بالا به جای v قرار می‌دهیم: $v = u_h - u$

$$c|u_h - u|_1 \leq \|u_h - u\|_a \leq C|u_h - u|_1.$$

بنابراین نامساوی اول داریم،

$$\|u_h - u\|_a = \min_{\chi \in S_h} \|\chi - u\|_a \leq C \min_{\chi \in S_h} |\chi - u|_1,$$

بنابراین،

$$|u_h - u|_1 \leq C \min_{\chi \in S_h} |\chi - u|_1.$$

حال به جای χ قرار می‌دهیم $\chi = I_h u$ ، که I_h عملگر درونیاب است. همچنین با استفاده از رابطه $|I_h v - v|_1 \leq ch\|v\|_2$ خواهیم داشت،

$$|u_h - u|_1 \leq C \min_{\chi \in S_h} |\chi - u|_1 \leq c|I_h u - u|_1 \leq ch\|u\|_2$$

□

این قضیه به این معناست که u_h بهترین تقریب برای u در S_h است.

۴.۲ خطای پسین

لم ۱.۴.۲. فرض کنید $\Omega = (0, 1) \times (0, 1)$ دامنه‌ای مربعی باشد که طول هر ضلع آن L است. آنگاه

$$\|v\|_{L^2(\Gamma)} \leq c(L^{-1}\|v\|_{L^2(\Gamma)}^2 + L\|\nabla v\|_{L^2(\Gamma)}^2)^{\frac{1}{2}}, \quad \forall v \in C^1(\bar{\Omega}),$$

داریم،

$$\|w\|_{\partial k} \leq c \left(h_k^{-\frac{1}{2}} \|w\|_k + h_k^{\frac{1}{2}} \|\nabla w\|_k \right)$$

برای بدست آوردن خطای پسین^۸ خواهیم داشت:

قضیه ۲.۴.۲. [۱۷] فرض کنید که Ω دامنه‌ای به شکل چندضلعی محدب در صفحه باشد. همچنین فرض کنید که u_h و u جواب معادلات (۶.۲) و (۷.۲) باشند، آنگاه

$$\|u_h - u\| \leq c \left(\sum_{T \in T_h} R_T^2 \right)^{\frac{1}{2}} \quad (۸.۲)$$

که در آن

$$R_T = h_T^{\frac{1}{2}} \|Au_h - f\|_k + h_k^{\frac{1}{2}} \|a[n \cdot \nabla u_h]\|_{\frac{\partial k}{T}}$$

که $[n \cdot \nabla u_h]$ بیانگر پرش در جهت بردار نرمال $n \cdot \nabla u_h$ می‌باشد.

^۸A posteriori error estimate

با استفاده از برآوردهای خطای پسین می‌توان برای حل عددی یک مسأله مقدار مرزی، از انطباق^۹ استفاده کرد. به این معنا که ابتدا با مثلث‌های بزرگ شروع کرده، مقدار خطا را در هر مثلث محاسبه می‌کنیم اگر خطای حاصل از خطای قابل تحمل کمتر باشد، به مثلث دیگری می‌رویم، ولی اگر خطا بیشتر از خطای قابل تحمل باشد، هر مثلث را به تعداد بیشتری مثلث تقسیم می‌کنیم. این روند آنقدر ادامه می‌یابد تا میزان خطا روی تمام مش از خطای قابل تحمل کوچکتر شود.

^۹Adaptivity

فصل ۳

آنالیز تطبیقی

۱.۳ مقدمه

یکی از اهداف مهم روش آنالیز تطبیقی پیدا کردن شبکه‌ی المان بهینه است. منظور از شبکه‌ی المان بهینه، شبکه‌ای است که با تعداد المان‌های ثابت حداقل خطای حل به روش عناصر متناهی را داشته باشد، اولین تلاش برای رسیدن به شبکه‌ی المان بهینه توسط نایس^۱ و مارکل^۲ در سال ۱۹۷۴ صورت گرفت.

۲.۳ انواع روش‌های آنالیز تطبیقی در حالت کلی

هدف اصلی آنالیز تطبیقی بدست آوردن بهترین جواب با صرف کمترین هزینه محاسباتی است، همانطور که قبلاً اشاره شد در این روش پس از حل به روش عناصر متناهی عادی در هر مرحله حل، از اطلاعات مرحله قبل استفاده می‌گردد. در این فرآیند جواب اولیه به روش عناصر متناهی عادی نقش بسیار مهمی در تقریب جواب و مقدار خطای محاسبه شده دارد. برای این منظور می‌توان از روش‌های زیر استفاده نمود.

- انتخاب یک شبکه المان جدید با اندازه المان‌های متفاوت و ثابت نگه داشتن درجه تابع درونیاب. به این روش از اصلاح شبکه المان، روش h گفته می‌شود. در این روش اندازه المان‌های جدید با استفاده از خطای محاسبه شده برای شبکه قبلی و اندازه المان در شبکه قبلی تعیین می‌گردد. به دو روش متفاوت می‌توان شبکه المان را تصحیح نمود.

الف- اصلاح شبکه المان به وسیله تقسیم المان قبلی به چند المان کوچکتر

^۱Neice

^۲Marcal

ب- توابع اصلاح شبکه المان با تعریف یک شبکه المان جدید که اساساً اندازه و موقعیت آنها با شبکه المان قبلی متفاوت باشد. در روش اول الگوی شبکه المان بدون تغییر باقی می‌ماند و تنها المان‌هایی که از خطای مجاز تجاوز می‌کنند به المان‌های کوچکتر تقسیم می‌شوند. در روش دوم تمام المان‌های موجود در شبکه قبلی مورد بررسی قرار می‌گیرند و با استفاده از آن شبکه المان جدید به دست می‌آید.

- تغییر دادن مرتبه توابع شکل المان یا تابع درونیاب، بدون تغییر در اندازه المان‌ها. به این روش از اصلاح شبکه المان، روش p گفته می‌شود. در این حالت الگوی شبکه المان قبلی بدون تغییر می‌ماند ولی تعداد گره‌های موجود در شبکه المان افزایش می‌یابد.
- روش دیگری که مورد استفاده قرار می‌گیرد، ترکیب دو روش فوق است یعنی علاوه بر تغییر بعد المان‌ها، توان توابع شکل نیز تغییر می‌نماید. به این روش از اصلاح شبکه المان، روش $h - p$ می‌گویند.

در این روش موقعیت گره‌ها به صورت نامعلوم در نظر گرفته می‌شود. در روش‌های آنالیز تطبیقی، شبکه المانی بهینه است که معیار خطای انرژی به صورت منطبق روی کل ناحیه توزیع شده باشد. این تعریف اولین بار در مورد مسایل یک بعدی صورت گرفت. در حال حاضر این تعریف در همه روش‌های آنالیز تطبیقی استفاده می‌شود. توجه به این نکته ضروری است که اگر از معیار خطای دیگری استفاده شود، شبکه المان بهینه تغییر خواهد نمود. به عبارت دیگر تعریف شبکه المان بهینه منحصر بفرد نیست. با توجه به مطالب فوق آشکار است که یک محاسبه‌گر خطای موثر و قابل اطمینان باید معیار خاص و مناسبی از خطا را برآورد نماید (روی کل دامنه) سپس، با یک سری عملیات معقول و مناسب، اندازه المان را محاسبه کند. در مورد مسایل غیرخطی برای یک روش آنالیز تطبیقی موفق و کارآمد باید دو موضوع اصلی در نظر گرفته شود.

- انتخاب روش مناسب آنالیز خطا
 - انتخاب مناسب یک اپراتور انتقالی متغیرها هنگامی که شبکه المان اصلاح می‌گردد.
- پس از برآورد خطا و محاسبه اندازه المان‌های جدید بایستی شبکه المان جدید تولید گردد و اطلاعات از شبکه المان قبلی به شبکه المان جدید منتقل شود. با انتقال این اطلاعات بر روی المان‌های جدید نه تنها امکان نقض روابط بنیادی وجود دارد بلکه ممکن است معادلات تعادل نیز برقرار نگردند. بنابراین بایستی در اینگونه مسایل روش‌های خاصی به کار گرفته شود.

۳.۳ تولید شبکه‌ی عناصر متناهی

مش‌بندی در تولید المان‌ها یکی از مراحل مهم در حل معادلات دیفرانسیل با مشتقات جزئی با روش عناصر متناهی در مسایل مکانیک جامدات است.

تولید مش به مفهوم تقسیم بندی و جایگزینی شکل هندسی اولیه با قطعات ساده تر و کوچک تر می باشد. برای حل مسایل سازه ای با استفاده از روش عناصر متناهی بایستی، مدل هندسی اولیه به یک مدل محاسباتی تبدیل شود. به این منظور مهم ترین بخش کار، تولید مش می باشد. تولید شبکه ای اجزای محدود به صورت اتوماتیک و با استفاده از المان های مثلثی شش گروهی و چهار ضلعی هشت ایزوپارامتریک انجام می گیرد. روش های تولید مش سازمان نیافته، روش هایی هستند که برای هر گره و المان موجود در داخل دامنه، تعداد گره ها و المان های موجود در همسایگی آن ها ثابت نمی باشد و تولید آنها به طور همزمان صورت می پذیرد. لازم به ذکر است که بهبود محلی شبکه ای اجزای محدود تنها به کمک این نوع روش مش بندی امکان پذیر است.

۴.۳ حل تطبیقی در عناصر متناهی

با توجه به اینکه در روش عناصر متناهی معمولی، هیچ گونه ابزاری به جز قضاوت مهندسی برای تعیین مناسب بودن اندازه المان به کار رفته و نیز دقت جواب، در دسترس مهندسین نمی باشد و انتخاب اندازه المان بیشتر بر اساس توصیه های داده شده صورت می گیرد، در روش های عددی، به ویژه روش عناصر متناهی، ضمن بیان مبانی تخمین خطا و حل تطبیقی در عناصر متناهی، از روش عناصر متناهی تطبیقی به عنوان ابزاری قوی برای غلبه بر محدودیت های روش عناصر متناهی معمولی استفاده شد. فرض کنید u جواب دقیق و u_h جواب تقریبی حاصل از روش عناصر متناهی باشد و $R(u_h) = u - u_h$ خطای مانده، e خطای پسین و tol خطای قابل چشم پوشی باشد. در حل معادلات دیفرانسیل با مشتقات جزئی به روش عناصر متناهی اگر از مش ریز استفاده کنیم جواب دقیق خواهد شد، ولی دو ایراد کلی بروز می کند.

۱. زمان زیادی برای محاسبات نیاز است.

۲. در برخی از نقاط، جواب با مش های درشت تر، دقت مورد نظر تامین شود.

برای رفع این مشکل از مفهوم انطباق^۳ استفاده می کنیم که ایده ی کلی آن به شکل زیر است: با یک مش درشت آغاز می کنیم. پس از محاسبه جواب، اگر جواب محاسبه شده دقت مورد نظر را در سرتاسر دامنه تامین کند که تمام است. ولی اگر در برخی از نواحی مش بندی شده، دقت تامین شود و در برخی تامین نشود خواهیم داشت:

• در نواحی درست، مش را بدون تغییر نگه می داریم.

• در نواحی با دقت کم، مش را ظریف می کنیم.

ساختار روش عناصر متناهی تطبیقی^۴ به صورت حلقه تکراری زیر است:

$AFEM : SOLVE \rightarrow ESTIMATE \rightarrow MARK \rightarrow REFINE,$

^۳Adaptivity

^۴Adaptive Finite Element Method

- $U_k \in V(\mathcal{T}_k), \quad U_k = \text{SOLVE}(\mathcal{T}_k),$ محاسبه جواب گالرکین
 - $\varepsilon_k = \text{ESTIMATE}(\mathcal{T}_k, U_k, f)$ محاسبه شاخص‌های خطای محلی $e(z)$
 - $M_k = \text{MARK}(\varepsilon_k, \mathcal{T}_k)$ انتخاب $M_k \subset \mathcal{T}_k$ با استفاده از علامت گذاری *Dorfler*
 - $\varepsilon_k(M_k) \geq \theta \varepsilon_k(\mathcal{T}_k) \quad \text{for } 0 < \theta < 1$
 - $\mathcal{T}_{k+1} = \text{REFINE}(\mathcal{T}_k, M_k)$ تعریف عناصر علامت گذاری شده M_k و خروجی مش منطبق $\mathcal{T}_{k+1}$ از $\mathcal{T}_k$
- برای پیاده سازی انطباق به شکل زیر عمل کنید:
- فرض کنید می‌خواهیم خطای مساله یعنی $\|e(x)\|_E$ از مقدار قابل تحمل TOL کمتر باشد.
- ۱- مش اولیه را در نظر بگیرید، این مش می‌تواند مش درشتی باشد.
 - ۲- $u_{\mathcal{T}}$ را براساس مش بالا محاسبه کنید.
 - ۳- خطای مانده $R(u_{\mathcal{T}})$ را روی هر بخش از مش حساب کنید.
 - ۴- اگر خطای مانده از TOL کمتر باشد که تمام است، در غیر اینصورت همان بخش از مش را ظریف کنید. و مراحل ۲، ۳، ۴ را تکرار کنید.

۵.۳ دوبخشی جدیدترین راس

فرم خاصی از تقسیمات، که به نام دوبخشی جدیدترین راس^۵ است را معرفی می‌کنیم. فرض کنید Ω دامنه چندوجهی در $\mathbb{R}^2$ باشد. $\mathcal{T}$ را به عنوان افرازی از Ω داخل عناصر مثلث T نشان می‌دهیم. که این به معنای $\Omega = \cup_{T \in \mathcal{T}} T$ است. همچنین با داشتن $T, T' \in \mathcal{T}$ خواهیم داشت $(T \cap T') = \emptyset$.

با داشتن افراز $\mathcal{T}$ ، فرض می‌کنیم که $\mathcal{S}_{\mathcal{T}}$ فضای توابع خطی تکه‌ای پیوسته، وابسته به $\mathcal{T}$ است که روی مرز یعنی $\partial\Omega$ به صفر میل می‌کند. تابع S در $\mathcal{S}_{\mathcal{T}}$ وجود دارد اگر و تنها اگر S تابع خطی روی هر $T \in \mathcal{T}$ باشد و همچنین S روی Ω پیوسته و روی $\partial\Omega$ به صفر میل کند، یعنی $\mathcal{S}_{\mathcal{T}} \subset H_0^1(\Omega)$ است. با $\varepsilon_{\mathcal{T}}$ مجموعه‌ای از یال‌های T و با $\mathcal{V}_{\mathcal{T}}$ مجموعه‌ای از راس‌های T را نشان می‌دهیم. در فرایند تطبیقی خانواده‌ای از افرازهای مجاز را ایجاد می‌کنیم.

روش عناصر متناهی تطبیقی دنباله‌ای از افرازهای $\mathcal{T}_0, \mathcal{T}_1, \dots, \mathcal{T}_n$ را با استفاده از قانون‌هایی برای مثلث‌های تقسیم شده ایجاد می‌کند. با M_k مجموعه‌ای از عناصر علامت گذاری شده را نمایش می‌دهیم. عناصر علامت گذاری شده با استفاده از قانون‌های تقسیم معین، تعریف شده است. این فرآیند، به‌رحال گره‌های آویز را ایجاد خواهد کرد. گوییم $v \in \mathcal{V}_{\mathcal{T}}$ گره آویز برای $T \in \mathcal{T}$ است اگر در درون یکی از

^۵Newest vertex bisection

یال‌های T ظاهر شود. افراز T_n ، افراز مجاز نهایی وابسته به این عملکرد از الگوریتم تطبیقی است. در اینجا خود را به روش خاصی از تقسیم شناخته شده به نام دوبخشی جدیدترین راس محدود می‌کنیم. افراز اولیه $\mathcal{T}$ از Ω داده شده است، برای هر $T \in \mathcal{T}$ ، به طور دقیق یکی از آن راس‌های $v(T)$ را به عنوان جدیدترین راس به آن عنصر اختصاص می‌دهیم. شماره‌گذاری اولیه می‌تواند به یک روش اختیاری در نظر گرفته شود. یالی در T که متضاد $v(T)$ است را با $E(T)$ نمایش می‌دهیم. اگر عنصری تقسیم شود، حاصل این تقسیمات دوبخشی ساده‌ای است که از طریق ارتباط جدیدترین راس و یال متضاد بدست آمده است.

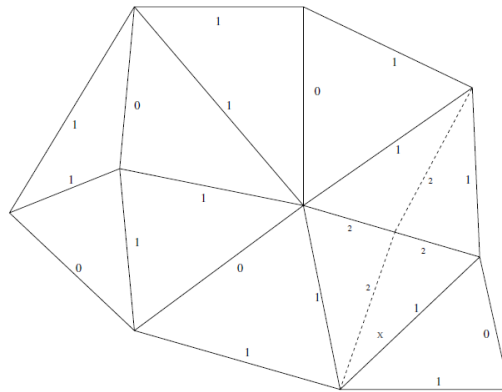
افرازهای بوجود آمده زمانی از دوبخشی جدیدترین راس استفاده می‌کنند که در شرط زاویه مینیمال منطبق صدق کنند. اگر افراز $\mathcal{T}$ توسط دنباله‌ای از دوبخشی جدیدترین راس بوجود آمده و همچنین دارای گره آویز نباشد، آن را افراز منطبق می‌نامیم. بنابراین این افراز مجاز خواهد بود. می‌توانیم دوبخشی جدیدترین راس را با درخت دودویی نامتناهی T^* (که آنرا درخت اصلی می‌نامیم) نمایش دهیم. درخت اصلی T^* شامل همه عناصر مثلثی است که می‌توانند توسط دنباله‌ای از تقسیمات بدست آیند. ریشه‌های درخت اصلی عناصر مثلثی در $\mathcal{T}$ هستند. زمانی که عنصر T تقسیم می‌شود، حاصل آن دو سلول جدید است که فرزندان T نامیده می‌شوند. جدیدترین راس وابسته یکتاست و فقط به T بوجود آمده در فرایند تقسیم بستگی دارد. یعنی آن به دنباله قبلی از تقسیمات بستگی ندارد. دلیل این امر اینست که هر تقسیم فقط جدیدترین راس را برای عناصر مثلثی جدید تولید شده توسط تقسیمات و تغییر ندادن هر تخصیص قبلی اختصاص می‌دهد.

برای زیر درخت معین T ، $N(T)$ را تعداد تقسیمات برای تولید T تعریف می‌کنیم. هر افراز $\mathcal{T} = T_n$ ای که توسط فرایند تطبیقی براساس دوبخشی جدیدترین راس بدست می‌آید، می‌تواند مربوط به زیر درخت معین $T = T(\mathcal{T})$ از T^* ، که شامل همه عناصر مثلثی که در طول الگوریتم ایجاد می‌شوند باشد. یعنی همه عناصر در $T_n, \dots, T_0$.

$T = T(\mathcal{T})$ مجاز است اگر $\mathcal{T}$ مجاز باشد. کلاس همه درخت‌های معین را با $\mathcal{T}$ و همه درخت‌های مجاز را با $\mathcal{T}^a$ نشان می‌دهیم. فرض می‌کنیم $\mathcal{T}_n$ مجموعه‌ای از همه درخت‌های معین T با $N(T) = n$ و $\mathcal{T}_n^a$ کلاس متناظر از درخت‌های مجاز از $\mathcal{T}_n$ باشد.

برای هر افراز اولیه $\mathcal{T}$ شماره‌گذاری یال‌ها در $\mathcal{T}$ به گونه‌ای است که هر یالی دارای برچسب $\circ$ یا $\bullet$ است. در مثلث $T \in \mathcal{T}$ ، دوتا از یال‌ها برچسب $\bullet$ و دیگری برچسب $\circ$ خواهد داشت. حال قانونی را برای شماره‌گذاری یال‌هایی که از فرایند کامل تقسیم به وجود می‌آید را ارایه می‌کنیم. دو خاصیت اصلی از این شماره‌گذاری وجود دارد. اول اینکه هر عنصر مثلثی، یال‌هایی با برچسب‌های $(i, i, i-1)$ برای برخی اعداد صحیح مثبت i دارد. دوم اینکه، جدیدترین راس برای این عنصر، راس متضاد با یالی است که دارای کمترین شماره می‌باشد.

فرض کنید چنین شماره‌گذاری را برای یال‌های در $\mathcal{T}_k$ داریم و همچنین این شماره‌گذاری را برای یال‌های در $\mathcal{T}_{k+1}$ توصیف خواهیم کرد. فرض کنید عنصر $T \in \mathcal{T}_k$ ، یال‌هایی را با برچسب $(i, i, i-1)$ دارد و جدیدترین راس برای این عنصر، متضاد با یالی است که دارای برچسب $i-1$ می‌باشد. زمانی که این عنصر از طریق دوبخشی جدیدترین راس تقسیم می‌شود، یالی که دارای برچسب $i-1$ است نصف شده و



شکل ۱.۳: برچسب یال جدید

برچسب هر کدام از دو یال جدید $i+1$ خواهد بود. همچنین نیمساز را هم با $i+1$ شماره گذاری می‌کنیم. بنابراین هر مثلث جدید، دارای یال‌هایی با برچسب $(i, i+1, i+1)$ می‌باشد. همچنین جدیدترین راس متضاد با یالی است که کمترین شماره را دارد.

۶.۳ تقریب تطبیقی در نورم $H^1(\Omega)$

تابع $w \in H^1_0(\Omega)$ و افراز $\mathcal{T}$ داده شده است. تعریف می‌کنیم،

$$E(w, \mathcal{T})_{H^1(\Omega)} := \inf_{s \in \mathcal{T}} \|w - s\|_{H^1(\Omega)}, \quad (1.3)$$

که کوچکترین خطایی است که می‌توانیم با تقریب w در نورم $H^1(\Omega)$ با عناصری از $\mathcal{T}$ انجام دهیم. در موردی که $w = u$ جواب $-\Delta u = f$ و $\mathcal{T}$ مجاز باشد داریم،

$$\|u - u_{\mathcal{T}}\| \leq E(u, \mathcal{T})_{H^1(\Omega)} \leq C_{\Omega}^{-1} \|u - u_{\mathcal{T}}\|. \quad (2.3)$$

که $u_{\mathcal{T}}$ تقریب گالرکین وابسته به $\mathcal{T}$ است.

به مورد کلی یعنی $w \in H^1_0(\Omega)$ برمی‌گردیم. همه افرازهای $\mathcal{T} \in \mathcal{T}_n$ را با هم مقایسه کرده و خطای زیر را از بهترین تقریب تطبیقی معرفی می‌کنیم:

$$\sigma_n(w) := \inf_{\mathcal{T} \in \mathcal{T}_n} E(w, \mathcal{T})_{H^1(\Omega)}. \quad (3.3)$$

برای هر $s > 0$ ، کلاس $\mathcal{A}^s := \mathcal{A}^s(H^1_0(\Omega))$ از توابع $w \in H^1_0(\Omega)$ را معرفی می‌کنیم. به گونه‌ای که،

$$\sigma_n(w) \leq M n^{-s}, \quad n = 1, 2, \dots, \quad (4.3)$$

کوچکترین M برای اینکه در (۴.۳) صدق کند، نورمی در $\mathcal{A}^s$ است.

$$\|w\|_{\mathcal{A}^s} := \sum_{n \geq 1} n^s \sigma_n(w). \quad (5.3)$$

زمانی که خود را به افرازهای مجاز محدود می‌کنیم، تقریب مشابه‌ای خواهیم داشت:

$$\sigma_n^a(w) := \inf_{\mathcal{T} \in \mathcal{T}_n^a} E(w, \mathcal{T})_{H^1(\Omega)}. \quad (6.3)$$

۷.۳ تقریب تطبیقی در نورم $H^{-1}(\Omega)$

برای تجزیه سمت راست f در الگوریتم عددی $-\Delta u = f$ به تقریب ثابت تکه‌ای نیاز داریم. تقریب فرض شده در نورم $H^{-1}(\Omega)$ برای توزیع بازگشتی g با دوگان زیر تعریف شده است.

$$\|g\|_{H^{-1}(\Omega)} := \sup_{\phi \in H_0^1(\Omega)} \frac{\langle g, \phi \rangle}{\|\phi\|}, \quad (۷.۳)$$

که $\langle \cdot, \cdot \rangle$ نمایانگر دوگان دوطرفی تولید شده توسط ضرب داخلی L_2 استاندارد است. افراز $\mathcal{T}$ داده شده، فرض می‌کنیم $\mathcal{S}_{\mathcal{T}}^{\circ}$ کلاس توابع ثابت تکه‌ای وابسته به $\mathcal{T}$ را نشان می‌دهد. برای تابع $g \in H^{-1}(\Omega)$ داریم

$$E(g, \mathcal{S}_{\mathcal{T}}^{\circ})_{H^{-1}(\Omega)} := \inf_{s \in \mathcal{S}_{\mathcal{T}}^{\circ}} \|g - s\|_{H^{-1}(\Omega)}, \quad (۸.۳)$$

که بهترین خطاست. می‌توانیم به تقریب f در نورم $H^{-1}(\Omega)$ با عناصری از $\mathcal{S}_{\mathcal{T}}^{\circ}$ دست پیدا کنیم. به طور مشابه داریم:

$$\sigma_n(g)_{H^{-1}(\Omega)} := \inf_{\mathcal{T} \in \mathcal{T}_n} E(g, \mathcal{S}_{\mathcal{T}}^{\circ})_{H^{-1}(\Omega)}, \quad (۹.۳)$$

که خطایی از بهترین تقریب غیرخطی نسبت به g با ثابت‌های تکه‌ای است. در مورد تقریب خطی تکه‌ای برای هر $s > 0$ کلاس تقریب $\mathcal{A}^s(H^{-1}(\Omega))$ و نورم آن به طور دقیق در (۴.۳) آمده است. از $\sigma_n(g)_{(H^{-1}(\Omega))}$ به جای $\sigma_n(w)$ استفاده می‌کنیم. کلاس تقریب در این مورد را با $\mathcal{B}^s(H^{-1}(\Omega))$ نشان می‌دهیم.

فرض کنید $g \in L_2(\Omega)$ ، بنابراین $g \in H^{-1}(\Omega)$ است. اگر $\mathcal{T}$ افرازی از Ω و $T \in \mathcal{T}$ باشد، تعریف می‌کنیم:

$$g_T := \frac{1}{|T|} \int_T g, \quad (۱۰.۳)$$

g_T میانگینی از g روی T است. همچنین g_T بهترین تقریب g در $L_2(T)$ با توابع ثابت است. بنابراین

$$\mathcal{S}_{\mathcal{T}}^{\circ}(g) := \sum_{T \in \mathcal{T}} g_T \chi_T, \quad (۱۱.۳)$$

بهترین تقریب $g \in L_2(\Omega)$ با ثابت‌های تکه‌ای وابسته به $\mathcal{T}$ است.

فصل ۴

روش عناصر متناهی تطبیقی برای مسایل بیضوی با ضرایب ناپیوسته

۱.۴ مقدمه

معادله دیفرانسیل با مشتقات جزئی بیضوی زیر را در نظر می‌گیریم:

$$\begin{cases} -\operatorname{div}(A\nabla u) = f, & x \in \Omega, \\ u = 0, & x \in \partial\Omega. \end{cases} \quad (1.4)$$

که در آن Ω دامنه چند وجهی در $\mathbb{R}^d$ است. که $d \geq 1$ عدد صحیح می‌باشد، و $A = (a_{ij})_{i,j=1}^\infty$ ماتریس $d \times d$ معین مثبتی است که در برگرنده توابعی از فضای L_∞ می‌باشد. فرض می‌کنیم که $|\cdot|$ نورم اقلیدسی روی $\mathbb{R}^d$ باشد. هنگامیکه $w : \Omega \rightarrow \mathbb{R}^d$ تابع برداری روی Ω در نظر گرفته شود، برای هر $0 < p < \infty$ داریم:

$$\|w\|_{L_p(\Omega)} := \| |w| \|_{L_p(\Omega)}. \quad (2.4)$$

به طور مشابه، اگر B ماتریس $d \times d$ باشد، آنگاه $\|B\|$ نشانگر نورم طیفی (نورم آن به عنوان اپراتور از $L_2(\mathbb{R}^d)$ به خودش است) آن خواهد بود. همچنین با توصیف مجدد a_{ij} روی مجموعه‌ای با اندازه صفر، فرض می‌کنیم که a_{ij} در هر جایی روی Ω تعریف می‌شود و داریم:

$$\|A(x)\| \leq \|A\|_{L_\infty(\Omega)} \|x\|_{L_\infty(\Omega)}.$$

چون x به صورت بردار در نظر گرفته شده است فرض می‌کنیم $\|x\|_{L_\infty(\Omega)} = 1$ باشد. بنابراین خواهیم داشت:

$$\|A(x)\| \leq \|A\|_{L_\infty(\Omega)}, \quad x \in \Omega, \quad (3.4)$$

رابطه (۱.۴) را در معنای ضعیف تفسیر کرده و از قضیه Lax-Milgram برای وجود و یکتایی آن استفاده می‌کنیم. ابتدا $H_0^1(\Omega)$ را فضای سوبولوف از توابع حقیقی مقدار روی Ω در نظر می‌گیریم. فرض

می‌کنیم Ω مجموعه بازی در $\mathbb{R}^n$ باشد. فضای سوبولوف $W^{1,2}(\Omega)$ با $H^1(\Omega)$ نمایش داده می‌شود. نورم سوبولوف به صورت زیر تعریف می‌شود:

$$\|v\|_{H^1} = \left(\int_{\Omega} |v|^2 + |\nabla v|^2 \right)^{\frac{1}{2}}.$$

زمانی که Ω کراندار است، $H_0^1(\Omega)$ به صورت فضایی از توابع در $H^1(\Omega)$ که در مرز به صفر میل می‌کند، توصیف می‌شود. بنابراین خواهیم داشت:

$$\|v\|_{H_0^1(\Omega)} := \|\nabla v\|_{L^2(\Omega)}. \quad (4.4)$$

فرم دوخطی $a(\cdot, \cdot)$ را روی فضای خطی V به صورت زیر در نظر می‌گیریم:

$$a : V \times V \rightarrow \mathbb{R}, \quad V = H_0^1(\Omega).$$

برای پیدا کردن فرم وردشی (۱.۴)، از تابع تست $v \in H_0^1(\Omega)$ و فرمول گرین استفاده می‌کنیم. ابتدا طرفین معادله (۱.۴) را در v ضرب کرده و سپس روی Ω انتگرال می‌گیریم:

$$\int_{\Omega} -\operatorname{div}(A \nabla u) v dx = \int_{\Omega} f v dx.$$

با بکارگیری انتگرالگیری جزیه‌جزیه داریم:

$$\int_{\Omega} A \nabla u \nabla v dx - \int_{\partial \Omega} a \frac{\partial u}{\partial n} v ds = \int_{\Omega} f v dx,$$

چون $v \in H_0^1(\Omega)$ است، پس روی مرز به صفر میل می‌کند. بنابراین رابطه دوم برابر با صفر، و خواهیم داشت:

$$\int_{\Omega} A \nabla u \nabla v dx = \int_{\Omega} f v dx, \quad (5.4)$$

فرم دوخطی (۵.۴) به شکل زیر است:

$$a(u, v) = \langle f, v \rangle,$$

که در آن

$$a(u, v) = \int_{\Omega} (A \nabla u) \cdot \nabla v dx, \quad u, v \in H_0^1(\Omega), \quad (6.4)$$

$$\langle f, v \rangle = \int_{\Omega} f v.$$

این مساله جواب یکتایی در $H_0^1(\Omega)$ دارد. حال می‌خواهیم جواب تقریبی برای معادله بدست آوریم. توابع پایه‌ای $\psi_1, \dots, \psi_n$ از L_n (فضای خطی n بعدی) را انتخاب می‌کنیم.

$$\sum_{e=1}^m \int_{\Omega^e} (A \nabla u \cdot \nabla v) d\Omega = \sum_{e=1}^m \int_{\Omega^e} f v d\Omega,$$

که،

$$u \approx \sum_{i=1}^{M_h} \psi_i u_i.$$

M_h تعداد گره‌ها و u_i مقادیر گره‌ای u هستند. همچنین داریم:

$$\psi_i(x_1^j, \dots, x_n^j) = \delta_i^j = \begin{cases} 1 & i = j, \\ 0 & i \neq j, \end{cases}$$

با استفاده از روش گالرکین ψ_i قرار می‌دهیم،

$$\int_{\Omega^e} \left(A \nabla \psi_i \cdot \nabla \sum_{j=1}^{M_h} \psi_j u_j \right) d\Omega = \int_{\Omega^e} f \psi_i d\Omega \quad i = 1, \dots, M_h.$$

دستگاه معادلات خطی زیر حاصل می‌شود:

$$k_{ij}^e u_j^e = f_i^e,$$

که معادل آن به شکل زیر است:

$$[k^e]_{M_h \times M_h} \cdot \{u^e\}_{M_h \times 1} = \{f^e\}_{M_h \times 1}$$

که k_{ij}^e و f_i^e هر کدام به صورت زیر نمایش داده می‌شود:

$$k_{ij}^e = \int_{\Omega^e} (A \nabla \psi_i \cdot \nabla \psi_j) d\Omega,$$

$$f_i^e = \int_{\Omega^e} f \psi_i d\Omega.$$

معیارهای همگرایی مختلفی می‌توان در نظر گرفت،

$$\frac{(\sum_{i=1}^n (u_i^k - u_i^{k-1})^2)^{\frac{1}{2}}}{(\sum_{i=1}^n (u_i^k)^2)^{\frac{1}{2}}} \leq \text{TOL},$$

n در فرمول بالا تعداد تکرارها را نشان می‌دهد.

همچنین داریم،

$$[k]_{n \times n} \cdot \{u\}_{n \times 1} = \{f\}_{n \times 1} \rightarrow [k] \cdot \{u\}^k = \{f\}$$

اندیس‌های بالا در این فرمول تعداد تکرارها و اندیس‌های پایین مرتبه ماتریس را نمایش می‌دهد. به طور کلی باید از $a \cdot b$ برای تعیین ضرب داخلی بردارهای a و b استفاده کنیم. فرض می‌کنیم که A روی Ω متقارن و به طور یکنواخت معین مثبت است. بدون از دست دادن کلیت، A را به صورت مجموعه‌ای با اندازه صفر در نظر می‌گیریم. بنابراین $A(x)$ در هر مکان روی Ω اکیدا مثبت است. فرض کنید ماتریس متقارن و معین مثبت B داده شده است. کمترین مقدار ویژه آن را با $\lambda_{\min}(B)$ و بیشترین مقدار ویژه آن را با $\lambda_{\max}(B)$ نشان می‌دهیم. در حالتی که B تابعی از $x \in \Omega$ باشد، خواهیم داشت:

$$\lambda_{\min}(B) := \inf_{x \in \Omega} \lambda_{\min}(B(x)),$$

و

$$\lambda_{\max}(B) := \sup_{x \in \Omega} \lambda_{\max}(B(x)) = \|\lambda_{\max}(B(\cdot))\|_{L^\infty(\Omega)} = \|B\|_{L^\infty(\Omega)}.$$

فرض می‌کنیم λ مقدار ویژه متناظر با بردار ویژه y باشد، یعنی

$$Ay = \lambda y,$$

و

$$0 < \lambda_{\min} \leq \lambda_i(x) \leq \lambda_{\max}, \quad (7.4)$$

از طرفی چون A معین مثبت است برای هر $y \neq 0$ نامساوی $y^t A y > 0$ برقرار است. طرفین رابطه (۷.۴) را در y و y^t ضرب می‌کنیم.

$$y^t \lambda_{\min}(A) y \leq y^t \lambda(x) y \leq y^t \lambda_{\max}(A) y.$$

چون λ مقدار مثبتی دارد خواهیم داشت:

$$\lambda_{\min}(A) y^t y \leq y^t A(x) y \leq \lambda_{\max}(A) y^t y,$$

بنابراین،

$$\lambda_{\min}(A) |y|^2 \leq y^t A(x) y \leq \lambda_{\max}(A) |y|^2. \quad (۸.۴)$$

با توجه به شرط کوارسیو و کراندار بودن $a(\cdot, \cdot)$ رابطه (۸.۴) دلالت دارد بر:

$$\lambda_{\min}(A) \|v\|_{H_0^1(\Omega)}^2 \leq a(v, v) \leq \lambda_{\max}(A) \|v\|_{H_0^1(\Omega)}^2, \quad \forall v \in H_0^1(\Omega) \quad (۹.۴)$$

نورم انرژی القاء شده با $a(\cdot, \cdot)$ هم ارز با نورم H_0^1 است. یعنی،

$$\|v\|_E = a(v, v)^{\frac{1}{2}}, \quad \forall v \in V.$$

با توجه به $H_0^1(\Omega)^* := H^{-1}(\Omega)$ فضای دوگان $H_0^1(\Omega)$ ، قضیه Lax-Milgram بر وجود جواب یکتایی $u = u_f \in H_0^1(\Omega)$ دلالت دارد به گونه‌ای که،

$$a(u, v) = \langle f, v \rangle, \quad \forall v \in H_0^1(\Omega), \quad (۱۰.۴)$$

در رابطه (۱۰.۴)، $\langle f, v \rangle$ متعلق به فضاهای $H^{-1} - H_0^1$ هستند.

برای اثبات یکتایی جواب فرض می‌کنیم $u_1, u_2 \in V$ دو جواب برای رابطه (۱۰.۴) باشند، لذا

$$a(u_1, v) = \langle f, v \rangle, \quad a(u_2, v) = \langle f, v \rangle.$$

فرض می‌کنیم $v = u_1 - u_2$ باشد، با توجه به شرط کوارسیو بودن $a(\cdot, \cdot)$ داریم:

$$\begin{aligned} \alpha \|u_1 - u_2\|_v^2 &\leq a(u_1 - u_2, u_1 - u_2) \\ &= \langle f, v \rangle - \langle f, v \rangle \\ &= 0, \end{aligned}$$

بنابراین $u_1 = u_2$ جواب یکتاست.

هدف از بکارگیری الگوریتم‌های عددی برای حل (۱۰.۴)، یافتن تقریب u در $H_0^1(\Omega)$ با دقت تعیین شده ε است. شروع این الگوریتم‌ها با تقریب f به $\hat{f}$ و A به $\hat{A}$ می‌باشد. برای تحلیل عملکرد چنین الگوریتمی، باید تخمینی برای تاثیرات چنین جایگزینی داشته باشیم. شکل معمول ناشی از نتیجه چنین اختلالی را در زیر مشاهده خواهید کرد.

فرض کنید که A و $\hat{A}$ هر دو متقارن و معین مثبت باشند، و برای $0 < r \leq M < \infty$ و $0 < \hat{r} \leq \hat{M} < \infty$ در شرایط زیر صدق کنند:

$$\begin{aligned} r &\leq \lambda_{\min}(A), \lambda_{\max}(A) \leq M, \\ \hat{r} &\leq \lambda_{\min}(\hat{A}), \lambda_{\max}(\hat{A}) \leq \hat{M}. \end{aligned} \quad (۱۱.۴)$$

پس خواهیم داشت:

$$\|u - \hat{u}\|_{H_0^1(\Omega)} \leq \hat{r}^{-1} (\|f - \hat{f}\|_{H^{-1}(\Omega)} + r^{-1} \|A - \hat{A}\|_{L^\infty(\Omega)} \|f\|_{H^{-1}(\Omega)}), \quad (۱۲.۴)$$

که $\hat{u} \in H_0^1(\Omega)$ جواب (۱۰.۴) با ضریب انتشار $\hat{A}$ و سمت راست $\hat{f}$ است. اگر A دارای ناپیوستگی باشد، برای اینکه (۱۲.۴) مناسب باشد، تقریب $\hat{A}$ باید با این ناپیوستگی‌ها هماهنگ شود. به این منظور سمت راست رابطه (۱۲.۴) باید کوچک در نظر گرفته شود. در اکثر عملکردها، ناپیوستگی A از ابتدا ناشناخته است، یا در پشت منحنی‌ها و سطوح قرار گرفته و به طور دقیق قابل تشخیص نیست. این وضعیت مانع استفاده مستقیم از (۱۲.۴) در ساخت و تحلیل روش‌های عددی برای (۱.۴) می‌شود. هدف اول در این پایان نامه، توصیف نظریه اختلال است، که در قضیه ۱.۲.۴ از بخش دوم آورده شده و جایگزین (۱۲.۴) با محدودیت زیر شده است،

$$\|u - \hat{u}\|_{H_0^1(\Omega)} \leq \hat{r}^{-1} (\|f - \hat{f}\|_{H^{-1}(\Omega)} + \|\nabla u\|_{L_p(\Omega)} \|A - \hat{A}\|_{L_q(\Omega)}), \quad q = \frac{2p}{p-2} \quad (۱۳.۴)$$

با این شرط که برای برخی $p \geq 2$ ، $\nabla u \in L_p(\Omega)$ باشد. توجه کنید که اگر $p = 2$ باشد، تخمین (۱۳.۴) مشابه با برآورد (۱۲.۴) است. با توجه به (۹.۴) و (۱۱.۴) داریم:

$$\begin{aligned} \alpha \|u\|_1^2 &\leq a(u, u) = L(u) = \langle f, u \rangle \\ &\leq \|f\| \|u\| \\ &\leq c \|u\|_1 \|f\| \end{aligned}$$

با حذف $\|u\|_1$ از طرفین و جایگزینی r به جای $\frac{\alpha}{c}$ خواهیم داشت:

$$r \|u\|_1 \leq \|f\| \rightarrow \|u\|_1 \leq r^{-1} \|f\|,$$

از طرفی از برابری $\|\nabla u\| = |u|_1$ نتیجه زیر حاصل می‌گردد:

$$\|\nabla u\| = |u|_1 \leq \|u\|_1 \leq r^{-1} \|f\|.$$

با توجه به اینکه $\nabla u \in L_2(\Omega)$ است داریم:

$$\|\nabla u\|_{L_2(\Omega)} \leq r^{-1} \|f\|_{H^{-1}(\Omega)},$$

مزیت (۱۳.۴) نسبت به (۱۲.۴) در این است که نیازی به هماهنگ کردن دقیق ناپیوستگی‌های A برای کوچک کردن سمت راست (۱۲.۴) نداریم. توجه کنید که همچنان باید کرانی برای مقدار ویژه A به‌ویژه $A \in L_\infty(\Omega)$ تعیین کنیم.

تخمین (۱۳.۴) نشانگر عدم تقارن در وابستگی مقادیر ویژه A و $\hat{A}$ است. برای اینکه شرط $\nabla u \in L_p(\Omega)$ را تضمین کنیم، به فرض‌های بیشتری در مورد f نیاز داریم. این موضوع در بخش دوم توضیح داده شده است. دامنه تغییراتی از $p > 2$ وجود دارد که تنها وابسته به ثابت‌های r ، M هستند، به گونه‌ای که $f \in W^{-1}(L_p(\Omega))$ (دوگان $(W_0^1(L_{\frac{p}{p-1}}(\Omega)))$ نشان می‌دهد که $\|\nabla u\| \in L_p(\Omega)$ است. بنابراین تخمین (۱۳.۴) می‌تواند برای این چنین f ‌هایی به کار برده شود. محدودیت $(W^{-1}(L_p(\Omega)))$ برای برخی $p > 2$ کاملاً مناسب است، و در همه کاربردهایی که در نظر داشتیم دیده می‌شود.

هدف دوم، توسعه روش عناصر متناهی تطبیقی (AFEM)^۱ قابل اجرا برای (۱۰.۴) است. اما زمانی که A دارای ناپیوستگی نامیزان با مش‌ها و تقریب عناصر متناهی در L_∞ امکان پذیر نباشد. روش عناصر متناهی تطبیقی را بر اساس قانون جدیدترین دوبرخی راس در بخش سه توسعه می‌دهیم.

^۱ Adaptive Finite Element Method

الگوریتم تطبیقی که ارایه می‌کنیم براساس سه الگوریتم ^۲RHS، ^۳COEFF و ^۴PDE می‌باشد. منظور از RHS الگوریتمی است که روی سمت راست معادله اصلی عمل می‌کند، یعنی تقریب f را با استفاده از چندجمله‌ای تکه‌ای بدست می‌آورد. این نوع تقریب از f کاملاً در AFEM ها استاندارد است. الگوریتم COEFF تقریب $\hat{A}$ به A را در فضای L_q ایجاد می‌کند. در بخش پنجم نشان می‌دهیم که محدودیت معین مثبت بودن $\hat{A}$ بر مرتبه تقریب L_q تاثیری ندارد. اما استخراج الگوریتم‌های قابل اجرای عددی که تضمین کننده معین مثبت بودن بوده و در شرایط تقریب $L_q(\Omega)$ عملکرد بهینه داشته باشند، یک مساله فرعی است. در نهایت بر اساس الگوریتم PDE، روش عناصر متناهی تطبیقی استاندارد را ارایه می‌کنیم [۲۰]. الگوریتم PDE را بر اساس تقریب سمت راست $\hat{f}$ و ضریب انتشار $\hat{A}$ که در الگوریتم‌های RHS و COEFF ارایه شدند، انجام می‌دهیم.

۲.۴ برهان اختلال

در این بخش قضیه اختلال را مطرح می‌کنیم، که در آن می‌توان تقریب A را به جای نرم ضعیف‌تر از L_∞ به کار برد. همان‌طور که خواهیم دید برای این کار به $\nabla u \in L_p(\Omega)$ برای برخی از $p > 2$ نیاز داریم.

۱.۲.۴ قضیه اختلال

فرض کنید که $A, \hat{A} \in [L_\infty(\Omega)]^{d \times d}$ ماتریس‌های متقارن و معین مثبت که منطبق با (۱۱.۴) برای برخی از $r, \hat{r} > 0$ و برخی از $M, \hat{M} < \infty$ باشند، همچنین فرض کنید $f, \hat{f} \in H^{-1}(\Omega)$ است. فرض کنید $u, \hat{u} \in H_0^1(\Omega)$ جواب‌های (۱۰.۴) باشند. مساله اختلال را به صورت زیر مطرح می‌کنیم:

$$\int_{\Omega} (\hat{A} \nabla \hat{u}) \cdot \nabla v = \langle \hat{f}, v \rangle, \quad \forall v \in H_0^1(\Omega). \quad (14.4)$$

اکنون پیوستگی لیپ شیتز نگاشت $A \rightarrow u$ را از $L_q(\Omega)$ به $H_0^1(\Omega)$ ثابت می‌کنیم. این نگاشت در [۱۶] آورده شده است.

قضیه ۱.۲.۴ (قضیه اختلال): برای هر $p \geq 2$ ، توابع u و $\hat{u}$ با این شرط که $\nabla u \in L_p(\Omega)$ باشد، برای $q = \frac{2p}{p-2} \in [2, +\infty]$ در رابطه‌ی زیر صدق می‌کنند،

$$\|u - \hat{u}\|_{H_0^1(\Omega)} \leq \hat{r}^{-1} \|f - \hat{f}\|_{H^{-1}(\Omega)} + \hat{r}^{-1} \|\nabla u\|_{L_p(\Omega)} \|A - \hat{A}\|_{L_q(\Omega)}, \quad (15.4)$$

برهان. فرض می‌کنیم $\bar{u}$ جواب (۱.۴) با ماتریس انتشار $\hat{A}$ و سمت راست f باشد. از برآورد اختلال (۱۲.۴) خواهیم داشت:

$$\|\hat{u} - \bar{u}\|_{H_0^1(\Omega)} \leq \hat{r}^{-1} \|f - \hat{f}\|_{H^{-1}(\Omega)} + r^{-1} \|\hat{A} - A\|_{L_\infty(\Omega)} \|f\|_{H^{-1}(\Omega)}.$$

^۲Ritght Hande Side

^۳Coefficient

^۴Partial Differential Equation

با توجه به صفر شدن $\|\hat{A} - A\|_{L_\infty(\Omega)}$ رابطه‌ی زیر حاصل می‌شود:

$$\|\hat{u} - \bar{u}\|_{H^1_*(\Omega)} \leq \hat{r}^{-1} \|f - \hat{f}\|_{H^{-1}(\Omega)}. \quad (۱۶.۴)$$

اکنون فقط کران $\|u - \bar{u}\|$ باقی می‌ماند. با توجه به تعریف u و $\bar{u}$ داریم:

$$\int_{\Omega} (A \nabla u) \cdot \nabla v = \int_{\Omega} (\hat{A} \nabla \bar{u}) \cdot \nabla v,$$

با اضافه و کم کردن $\int_{\Omega} (\hat{A} \nabla u) \cdot \nabla v$ به سمت راست تساوی بالا داریم:

$$\int_{\Omega} (A \nabla u) \cdot \nabla v = \int_{\Omega} (\hat{A} \nabla \bar{u}) \cdot \nabla v - \int_{\Omega} (\hat{A} \nabla u) \cdot \nabla v + \int_{\Omega} (\hat{A} \nabla u) \cdot \nabla v,$$

در نتیجه برای همه $v \in H^1_*(\Omega)$ خواهیم داشت:

$$\int_{\Omega} [\hat{A} \nabla (u - \bar{u})] \cdot \nabla v = \int_{\Omega} [(\hat{A} - A) \nabla u] \cdot \nabla v,$$

اگر $v = u - \bar{u}$ باشد، داریم:

$$\begin{aligned} \int_{\Omega} [\hat{A} \nabla (u - \bar{u})] \cdot \nabla (u - \bar{u}) &= \int_{\Omega} [(\hat{A} - A) \nabla u] \cdot \nabla (u - \bar{u}) \\ &\leq \|(\hat{A} - A) \nabla u\|_{L_r(\Omega)} \|\nabla (u - \bar{u})\|_{L_r(\Omega)}. \end{aligned}$$

اگر از تخمین (۱۳.۴) با این شرط که A جایگزین $\hat{A}$ شده باشد، استفاده کنیم نتیجه زیر را خواهیم داشت:

$$\begin{aligned} \hat{r} \|u - \bar{u}\|_{H^1_*(\Omega)} &\leq \lambda_{\min}(\hat{A}) \|v\|_{H^1_*(\Omega)} \\ &\leq \int_{\Omega} \hat{A} \nabla (u - \bar{u}) \cdot \nabla (u - \bar{u}) \\ &\leq \|(A - \hat{A}) \nabla u\|_{L_r(\Omega)} \|\nabla (u - \bar{u})\|_{L_r(\Omega)} \\ &\leq \|(A - \hat{A}) \nabla u\|_{L_r(\Omega)} \|u - \bar{u}\|_{H^1_*(\Omega)}. \end{aligned}$$

با حذف $\|u - \bar{u}\|_{H^1_*(\Omega)}$ از طرفین داریم:

$$\hat{r} \|u - \bar{u}\|_{H^1_*(\Omega)} \leq \|(A - \hat{A}) \nabla u\|_{L_r(\Omega)},$$

با بکارگیری نامساوی هولدر در سمت راست، با $p \geq 2$ و $q = \frac{2p}{p-2}$ ، خواهیم داشت:

$$\|u - \bar{u}\|_{H^1_*(\Omega)} \leq \hat{r}^{-1} \|\nabla u\|_{L_p(\Omega)} \|A - \hat{A}\|_{L_q(\Omega)}, \quad (۱۷.۴)$$

ادغام (۱۷.۴) با (۱۶.۴) رابطه زیر را نتیجه خواهد داد:

$$\begin{aligned} \|u - \hat{u}\|_{H^1_*(\Omega)} &\leq \|u - \bar{u}\|_{H^1_*(\Omega)} + \|\bar{u} - \hat{u}\|_{H^1_*(\Omega)} \\ &\leq \hat{r}^{-1} \|\nabla u\|_{L_p(\Omega)} \|A - \hat{A}\|_{L_q(\Omega)} + \hat{r}^{-1} \|f - \hat{f}\|_{H^{-1}(\Omega)}, \end{aligned}$$

□

لذا حکم ثابت است.

تذکر ۲۰۲.۴. (تخمین‌های اختلال محلی): دیدیم که انتخاب p در تخمین اختلال (۱۵.۴) در زیر دامنه‌های Ω با هم تفاوت دارند. برای تثبیت این ایده، فرض کنید Ω به دو دامنه Ω_1 و Ω_2 تقسیم شده است. بحث‌هایی مشابه با آنچه در قبل داشتیم مطرح می‌شوند.

$$\begin{aligned} \|u - \hat{u}\|_{H^1_*(\Omega)} &\leq \hat{r}^{-1} \|f - \hat{f}\|_{H^{-1}(\Omega)} + \hat{r}^{-1} \|\nabla u\|_{L_{p_1}(\Omega_1)} \|A - \hat{A}\|_{L_{q_1}(\Omega_1)} \\ &\quad + \hat{r}^{-1} \|\nabla u\|_{L_{p_2}(\Omega_2)} \|A - \hat{A}\|_{L_{q_2}(\Omega_2)} \end{aligned}$$

که

$$p_i \in [2, \infty), \quad q_i = \frac{p_i}{(p_i - 2)}, \quad i = 1, 2.$$

در بخش آخر خواهیم دید، این رابطه زمانی که پرش در ضرایب صورت گیرد، در دامنه فرعی Ω_i دارای جواب $u \in (W_\infty^1(\Omega_i))$ است، در نتیجه $p_i = \infty$ خواهد بود.

۲.۲.۴ شرایط کافی برای ∇u که در L_p باشد

برای درستی برهان قضیه ۱.۲.۴، باید برای برخی $p > 2$ ، ∇u را در L_p فرض کنیم. همچنین باید اطلاعات کافی از ضریب انتشار A و سمت راست f در اختیار داشته باشیم. به این منظور، شرط p را به صورت زیر بیان خواهیم کرد:

شرط p : برای هر $f \in W^{-1}(L_p(\Omega))$ جواب $u = u_f$ در رابطه زیر صدق می‌کند.

$$\alpha |u|_{W^1(L_p(\Omega))}^2 \leq a(u, u) = L(u) = |\langle f, u \rangle|,$$

$$\leq \|f\| \|u\| \leq c \|f\| |u|_{W^1(L_p(\Omega))},$$

با حذف $|u|_{W^1(L_p(\Omega))}$ از طرفین، داریم:

$$|u|_{W^1(L_p(\Omega))} \leq C_p \|f\|_{W^{-1}(L_p(\Omega))},$$

چون $|u|_{W^1(L_p(\Omega))}$ با $\|\nabla u\|_{L_p(\Omega)}$ برابر است، بنابراین خواهیم داشت:

$$|u|_{W^1(L_p(\Omega))} := \|\nabla u\|_{(L_p(\Omega))} \leq C_p \|f\|_{W^{-1}(L_p(\Omega))}, \quad (18.4)$$

که ثابت C_p مستقل از f می‌باشد.

در حالتی که $A = I$ (معادله لاپلاس) در نظر گرفته شود، درستی شرط p مساله‌ای از پیش تعیین شده در آنالیز هارمونیک است. برای هر دامنه لیبیشیتز Ω ، $P > 2$ ای وجود دارد که وابسته به Ω است. به گونه‌ای که شرط p برای همه $2 \leq p \leq P$ برقرار است. زمانی که $A = I$ است، (۱۸.۴) را بر اساس k یعنی مقدار ثابت وابسته به Ω و همچنین P بررسی می‌کنیم، بنابراین رابطه زیر را داریم:

$$\|\nabla u\|_{(L_p(\Omega))} \leq k \|f\|_{W^{-1}(L_p(\Omega))}. \quad (19.4)$$

شرط p را می‌توان با استفاده از بحث اختلال که در [۱۹] مطرح شده، ارایه کرد. نتایج [۱۹] را برای حالتی که $p > 2$ است بیان می‌کنیم.

ابتدا $\eta(p)$ را به صورت زیر تعریف می‌کنیم:

$$\eta(p) = \frac{\frac{1}{2} - \frac{1}{p}}{\frac{1}{2} - \frac{1}{P}}, \quad (20.4)$$

پس $\eta(p)$ در $p = 2$ از مقدار صفر به مقدار یک در $p = P$ افزایش می‌یابد.

بنابراین برای هر $t \in (0, 1)$ ، $P^*(t)$ را به صورت زیر تعریف می‌کنیم:

$$P^*(t) := \max\{k^{-\eta(p)} > 1 - t \quad ; \quad 2 < p < P\}. \quad (21.4)$$

با این تعاریف، نتیجه زیر را برای A و همچنین کامل شدن این بخش ارایه می‌کنیم.

گزاره ۳.۲.۴. (عضویت در $(W^1(L_p(\Omega)))$). فرض کنید f و Ω به گونه ای باشند که، برای برخی $P > 2$ و مقادیر ثابت k ، هرگاه $f \in W^{-1}(L_p(\Omega))$ باشد، جواب از مساله (۱۰.۴)، جواب مناسبی برای (۱۹.۴) باشد.

اگر (۱۳.۴) برای A برقرار باشد، جواب $u \in H^1_0(\Omega)$ برای (۱۰.۴) مناسب خواهد بود، یعنی

$$\|\nabla u\|_{L_p(\Omega)} \leq C \|f\|_{W^{-1}(L_p(\Omega))},$$

با این شرط که، روابط زیر برقرار باشد.

$$2 \leq p \leq P^*\left(\frac{r}{M}\right), \quad C := \frac{1}{M} \frac{k^{\eta(p)}}{1 - k^{\eta(p)}\left(1 - \frac{r}{M}\right)}.$$

برهان. در حالتی که $A = I$ باشد، $-\Delta u = f$ را داریم. بنابراین عملگر $T := -\Delta$ از $H^{-1}(\Omega)$ به $H^1_0(\Omega)$ قابل تغییر است، معکوسش T^{-1} کراندار با نورم یک می باشد. همچنین با توجه به (۱۹.۴)، کراندار با نورم k به صورت نگاشتی از $W^{-1}(L_p(\Omega))$ به $W^1_0(L_p(\Omega))$ است. برای داشتن روش واقعی درونیایی، برای $2 < p < P$ داریم:

$$W^1_0(L_p(\Omega)) = [H^1_0(\Omega), W^1_0(L_p(\Omega))]_{\eta(p), p}, \quad 2 < p < P,$$

که $\eta(p)$ را در (۲۰.۴) تعریف کردیم. بر اساس درونیایی T^{-1} نگاشتی کراندار از $W^{-1}(L_p(\Omega))$ به $W^1_0(L_p(\Omega))$ است. از طرفی با توجه به مقدار ثابت k رابطه ی زیر خواهیم داشت:

$$\|\nabla u\|_{L_p(\Omega)} = \|\nabla T^{-1} f\|_{L_p(\Omega)} \leq k^{\eta(p)} \|f\|_{W^{-1}(L_p(\Omega))}.$$

فرض کنید $S : W^1_0(L_p(\Omega)) \rightarrow W^{-1}(L_p(\Omega))$ عملگری مطابق $Sv = -\operatorname{div}(\frac{1}{M} A \nabla v)$ باشد. همچنین، عملگر اختلال $Q = T - S$ را تعریف می کنیم. S و Q عملگرهای کراندار از $W^1_0(L_p(\Omega))$ به $W^{-1}(L_p(\Omega))$ با نورم های زیر هستند:

$$\|S\| \leq 1, \quad \|Q\| \leq 1 - \frac{r}{M},$$

با استفاده از نگاشتی از $W^1_0(L_p(\Omega))$ به $W^1_0(L_p(\Omega))$ خواهیم داشت:

$$\|T^{-1}Q\| \leq \|T^{-1}\| \|Q\| \leq k^{\eta(p)} \left(1 - \frac{r}{M}\right).$$

به شرطی که $1 < k^{\eta(p)} \left(1 - \frac{r}{M}\right)$ و $2 \leq p \leq P^*\left(\frac{r}{M}\right)$ ، عملگر

$$S = T - Q = T(I - T^{-1}Q)$$

معکوس پذیر است. بعلاوه، بر اساس نگاشتی از $W^{-1}(L_p(\Omega))$ به $W^1_0(L_p(\Omega))$ خواهیم داشت:

$$S^{-1} = \frac{T^{-1}}{I - T^{-1}Q}$$

که

$$\|S^{-1}\| \leq \frac{\|T^{-1}\|}{\|I - T^{-1}Q\|} \leq \frac{k^{\eta(p)}}{1 - k^{\eta(p)}\left(1 - \frac{r}{M}\right)}$$

□

کران مطلوب حاصل شد.

۳.۴ روش‌های عناصر متناهی تطبیقی

روش‌های عناصر متناهی تطبیقی در حال حاضر به طور گسترده‌ای در محاسبات‌های علمی، برای رسیدن به بهترین دقت با حداقل درجه آزادی، استفاده می‌شود. الگوریتمی که ارایه می‌کنیم از نظر فرضیات در خصوص ماتریس انتشار A با سایر الگوریتم‌ها فرق دارد. هر ورودی این ماتریس، چندجمله‌ای تکه‌ای روی افراز^۵ اولیه $\mathcal{T}$ ، در نظر گرفته می‌شود. در این الگوریتم، به فرض هماهنگی ناپیوستگی A با $\mathcal{T}$ نیازی نداریم. الگوریتمی که در اینجا به کار می‌بریم، از زیربرنامه‌ای استفاده می‌کند که در روش‌های عناصر متناهی تطبیقی استاندارد است و می‌توان آن‌ها را در [۱۴، ۲۵] مشاهده کرد، یعنی همان بخشی که تقریب‌های f بررسی شده‌اند. بنابراین در این بخش الگوریتم‌های موجود را بررسی می‌کنیم.

۱.۳.۴ افرازاها و فضای عناصر متناهی

روش عناصر متناهی تطبیقی، روشی برای تقسیم بندی منطبق دامنه به المان‌های چندوجهی است. تاکنون مقالات زیادی در زمینه قوانین نظریف شده برای روش‌های عناصر متناهی تطبیقی چاپ شده است [۹].

در مبحث زیر حالت دو بعدی (مثلث‌ها) و جدیدترین روش دو نیمه کردن راس را در نظر می‌گیریم. اما نتایجی که ارایه می‌کنیم، برای $d \geq 2$ و براساس قوانین نظریف^۶ در [۹] هستند. همچنین، بحث خود را به دو نیمه کردن راس محدود می‌کنیم.

ابتدا، فرض کنید Ω دامنه چندوجهی بوده و $\mathcal{T}$ افراز اولیه از Ω باشد، که دارای تعداد مشخصی از نواحی مثلثی شکل است.

اگر بخواهید عنصری را نظریف کنید، باید آن را با استفاده از ضلع مقابل به جدیدترین راس^۷ به دو نیمه تقسیم کنید، بدین ترتیب دو عنصر فرزند خواهید داشت. عنصر $T \in \mathcal{T}$ در افراز اولیه داده شده، راس‌ها را با $V(T)$ و یال‌ها را با $E(T)$ ، مقابل راس علامت‌گذاری می‌کنیم. عنصر $T \in \mathcal{T}$ توسط الحاق $V(T)$ و نقطه میانی $E(T)$ قطع می‌شود. و این به دو عنصر T_1 و T_2 که فرزندان T هستند خاتمه می‌یابد. راس $V(T_1) = V(T_2)$ نقطه میانی $E(T)$ هست و $E(T_1)$ ، $E(T_2)$ یال‌های معکوس برای $V(T_1)$ و $V(T_2)$ می‌باشند.

قانون دو بخشی، روند نظریف و جنگل $\mathcal{F}$ استخراج شده از ریشه $\mathcal{T}$ را به ما خواهد داد. زمانی که بازه Ω افرازبندی می‌شود درون هر افراز، نظریف‌هایی صورت می‌گیرد که مجموع کل این نظریف‌ها را جنگل می‌نامیم. اگر بتوان با استفاده از دو نیمه کردن جدیدترین راس‌ها، $\mathcal{T} \in \mathcal{F}$ را از $\mathcal{T}$ استخراج کرد، آنگاه $\mathcal{T} \in \mathcal{F}$ منطبق است. میزان پیچیدگی $\mathcal{T}$ را می‌توان بر اساس تعداد $n(T)$ از دوبخشی‌ها اندازه گیری کرد. در واقع $\#T = \#\mathcal{T} + n(T)$ را داریم. با $\mathcal{F}_n$ ، $n \geq 1$ ، مجموعه همه افرازهای $\mathcal{T}$ را که می‌توان

^۵Partition

^۶Refinement

^۷Newest Vertex Bisection

از T_0 به واسطه n دوقسمتی جدیدترین راس‌ها بدست آورد، نمایش می‌دهیم. روند مثلث سازی $T \in \mathcal{F}_n$ ممکن است منطبق نباشد، یعنی شاید شامل گره‌های آویز (گوییم v گره آویز برای $T \in \mathcal{T}$ است اگر در درون یکی از یال‌های T ظاهر شود)، باشد. اگر T_k مش منطبق بدست آمده از T_0 ، توسط دوبخشی جدیدترین راس باشد و $\mathcal{R}_k$ زیر مجموعه‌ای از مثلث‌های نظریف شده باشد داریم:

$$\#T_k - \#T_0 \leq C_0 \sum_{j=0}^{k-1} \#(\mathcal{R}_j), \quad k \geq 1,$$

فرض کنید $\hat{T}$ ، مثلث غیرمنطبق به دست آمده از T_0 باشد. اگر T کوچک‌ترین نظریف منطبق از $\hat{T}$ باشد خواهیم داشت:

$$\#T - \#T_0 \leq C_0 (\#\hat{T} - \#T_0).$$

اگر $\bar{T}$ کوچک‌ترین نظریف منطبق از $\hat{T}$ باشد، رابطه زیر را داریم:

$$\#\bar{T} \leq \#T_0 + C_0 \left((\#\hat{T} - \#T_0) + (\#T - \#\hat{T}) \right).$$

اگر T منطبق نباشد، با استفاده از دوقسمتی کردن جدیدترین راس‌ها، تحت کنترل $n(T)$ ، می‌توان آن را به افزاز منطبق $\bar{T}$ به صورت زیر تبدیل کرد:

$$\#\bar{T} - \#T_0 \leq C_0 n(T). \quad (22.4)$$

در (22.4)، C_0 مقدار ثابت مطلق است که فقط به T_0 و برجسب آن بستگی دارد. افزاز منطبق T را بصورت $^{\wedge}\text{CONF}$ مشخص می‌کنند، که $\text{CONF}(T)$ کوچک‌ترین افزاز مجاز و منطبق شامل T می‌باشد.

با داشتن افزاز منطبق $T \in \mathcal{F}_n$ و درجه چندجمله‌ای $m_u \geq 1$ ، $V(T)$ را فضای عناصر متناهی از چندجمله‌ای‌های تکه‌ای پیوسته از درجه m_u وابسته به T تعریف می‌کنیم.

$$V(T) \subset H_0^1(\Omega), \quad H_0^1(\Omega) = \{v \in H^1(\Omega) \mid v|_{\partial\Omega} = 0\},$$

با داشتن ماتریس انتشار معین مثبت $A \in L_\infty$ ، و مقادیر سمت راست $f \in L_2(\Omega)$ ، تقریب گالرکین^۹ $U := U(T, A, f) = \text{GAL}(T, A, f)$ در (۱۳.۴) جوابی یکتا برای مساله زیر است.

$$U \in V(T) : \int_{\Omega} (A \nabla U) \cdot \nabla V = \int_{\Omega} f V, \quad \forall v \in V(T), \quad (23.4)$$

تابع U بهترین تقریب u از $V(T)$ در نورم انرژی القا شده با A می‌باشد، که در عین حال معادل نورم $H_0^1(\Omega)$ است.

۲.۳.۴ ساختار AFEM

روش‌های عناصر متناهی تطبیقی استاندارد، دنباله‌ای از افزازهای تودرتو منطبق $\{T_k\}_{k \geq 0}$ برای تقریب u ایجاد می‌کنند. با استفاده از یک راه‌حل تطبیقی، افزاز T_{k+1} از T_k با $k \geq 0$ بدست می‌آید.

^۸Conform

^۹Galerkin

با داشتن افراز $\mathcal{T}$ و تابع عناصر محدود $V \in V(\mathcal{T})$ ، تخمین‌گر باقیمانده به صورت زیر تعریف می‌شود:

$$\eta_{\mathcal{T}}(V, A, f; \mathcal{T}) := \left(\sum_{T \in \mathcal{T}} \eta(V, A, f; T) \right)^{\frac{1}{2}},$$

$$\eta_{\mathcal{T}}(V, A, f; T) := \text{diam}(T) \|f + \text{div}(A \nabla V)\|_{L_r(T)} + \left(\sum_{F \in \Sigma(T)} \text{diam}(F) \| [A \nabla V] \|_{L_r(F)}^2 \right)^{\frac{1}{2}}. \quad (24.4)$$

در (24.4)، $\Sigma(T)$ مجموع یال‌ها ($d = 2$)، یا سطوح ($d = 3$) است، که کران T را نشان می‌دهد. همچنین [۰] به معنی پرش نرمال در میان F است.

درستی جواب گالرکین، یعنی $U_k = \text{GAL}(\mathcal{T}_k, A, f) \in V(\mathcal{T}_k)$ را با امتحان کردن $\eta_{\mathcal{T}_k}(U_k, A, f; T)$ و علامت‌گذاری عناصر خاص در $\mathcal{T}_k$ ، جهت نشانه‌گذاری تعیین می‌کنیم.

پس از اجرای این تعریف‌ها برای برداشتن گره‌های آویز، افراز منطبق جدیدی خواهیم داشت. این فرایند تا رسیدن تخمین‌گر باقیمانده به پایین‌تر از خطای مجاز ε_k تکرار می‌شود. در حالتی که A و f چندجمله‌ای‌های تکه‌ای وابسته به $\mathcal{T}_{k+1}$ باشند، میزان تعیین شده توسط تخمین‌گر باقیمانده معادل با خطای انرژی است. یعنی ثابت‌های $C_L \leq C_U$ وابسته به جنگل $\mathcal{F}$ روی مقادیر ویژه A وجود دارند به گونه‌ای که در زیر نمایش می‌دهیم.

$$C_L \eta_{k+1}(U_{k+1}, A, f; \mathcal{T}_{k+1}) \leq \|u - U_{k+1}\|_{H^1_0(\Omega)} \leq C_U \eta_{k+1}(U_{k+1}, A, f; \mathcal{T}_{k+1}) \quad (25.4)$$

این روند را براساس الگوریتم PDE انجام می‌دهیم که به صورت زیر خواهد بود:

$$[\mathcal{T}_{k+1}, U_{k+1}] = \text{PDE}(\mathcal{T}_k, A, f, \varepsilon_k),$$

$$\eta_{k+1}(U_{k+1}, A, f; \mathcal{T}_{k+1}) \leq \varepsilon_k,$$

به عبارت دیگر ورودی‌های الگوریتم PDE، افراز $\mathcal{T}_k$ ، ماتریس A ، f و خطای ε_k است. خروجی این الگوریتم، افراز $\mathcal{T}_{k+1}$ و جواب گالرکین U_{k+1} است که در کران خطای زیر صدق می‌کند:

$$\|u - U_{k+1}\|_{H^1_0(\Omega)} \leq C_U \varepsilon_k. \quad (26.4)$$

هر حلقه داخل الگوریتم PDE، ادغام خطای انرژی با ثابت $\alpha < 1$ وابسته به C_L ، C_U و پارامترهای نشانه‌گذاری [۲۰] است. اگر $\mathcal{T}_{k+1}$ گره درونی $\mathcal{T}_k$ باشد، کران پایین گسسته زیر را از [۲۰] داریم:

$$C_U \varepsilon_k(U_k, M_k)^2 \leq \|U_{k+1} - U_k\|_{\Omega}^2$$

با اضافه و کم کردن U_k به $\|u - U_{k+1}\|_{\Omega}$ و استفاده از نامساوی مثلثی داریم:

$$\|u - U_{k+1}\|_{\Omega}^2 \leq \|u - U_k\|_{\Omega}^2 + \|U_k - U_{k+1}\|_{\Omega}^2, \quad (27.4)$$

ترکیب کران پایین گسسته و کران بالا رابطه زیر را خواهد داد:

$$\|U_{k+1} - U_k\|_{\Omega}^2 \geq C_U \varepsilon_k(U_k, M_k)^2 \geq C_U \theta^2 \varepsilon_k(U_k)^2 \geq \frac{C_U}{C_L} \theta^2 \|u - U_k\|_{\Omega}^2, \quad (28.4)$$

با جایگذاری (28.4) در (27.4) خواهیم داشت:

$$\|u - U_{k+1}\|_{\Omega}^2 \leq \|u - U_k\|_{\Omega}^2 - \frac{C_U}{C_L} \theta^2 \|u - U_k\|_{\Omega}^2,$$

برای $\alpha := (1 - \frac{C_U}{C_L} \theta^2)^{\frac{1}{2}} < 1$ داریم:

$$\|u - U_{k+1}\|_{\Omega}^2 \leq (1 - \frac{C_U}{C_L} \theta^2) \|u - U_k\|_{\Omega}^2,$$

بنابراین اگر $\hat{\varepsilon}_k = \eta_{T_k}(U_k, A, f, T_k)$ ، سطح خطا قبل از بکارگیری PDE باشد، تعداد تکرارهای i_k در PDE برای کاهش خطا تا سطح ε_k برابر است با:

$$i_k \leq \frac{\log(\frac{C_U M^{\frac{1}{p}}}{C_L r^{\frac{1}{p}}}) + \log(\frac{\hat{\varepsilon}_k}{\varepsilon_k})}{\log \alpha^{-1}} + 1, \quad (29.4)$$

این الگوریتم، به خوبی خطای موجود در جواب (۲۳.۴) را نشان نمی‌دهد. برای این مرحله به محاسبه انتگرال‌هایی نیاز داریم که حاصل، A و f با توابعی از فضای عناصر متناهی هستند. در تحلیل چنین الگوریتم‌هایی، فرض بر این است که، انتگرال‌ها به طور دقیق محاسبه شده‌اند. روش جایگزینی که در [۴، ۲۵] برای عملگر لاپلاس معرفی شده، تقریب f با چندجمله‌ای تکه‌ای f_k روی T_k است.

الگوریتم AFEM ای که ارائه می‌کنیم متفاوت از الگوریتم PDE است، چون در الگوریتم AFEM از هر دو تقریب A و f استفاده می‌کنیم. افزاز T_k و خطای ε_k ، داده شده است. الگوریتم AFEM، ابتدا افزاز منطبق T_k' را می‌یابد، که همان تقریب شده T_k است. f را می‌توانیم با چندجمله‌ای تکه‌ای f_k و همچنین A را با چندجمله‌ای تکه‌ای A_k تقریب بزنیم. به گونه‌ای که

$$\|f - f_k\|_{H^{-1}(\Omega)} \leq \varepsilon_k', \quad \|A - A_k\|_{L_q(\Omega)} \leq \varepsilon_k'. \quad (30.4)$$

روابط (۳۰.۴) برای $q = \frac{2p}{(p-2)} \in [2, \infty]$ برقرار می‌باشد. (الگوریتم‌های فعلی همیشه $q = \infty$ را در نظر می‌گیرند)

خطای ε_k' را مضربی از ε_k به صورت $\varepsilon_k' = w \varepsilon_k$ با $w > 0$ در نظر می‌گیریم. که باید تعیین شود. الگوریتم $[T_{k+1}, U_{k+1}] = \text{PDE}(T_k', A_k, f_k, \frac{\varepsilon_k}{4})$ را برای یافتن افزاز T_{k+1} که تقریب شده T_k' است، به کار می‌بریم. با استفاده از T_k و جواب گالرکین $U_{k+1} \in V(T_{k+1})$ رابطه‌ی زیر را خواهیم داشت:

$$\|u_k - U_{k+1}\|_{H_0^1(\Omega)} \leq C_U \eta_{k+1}(U_{k+1}, A_k, f_k; T_{k+1}) \leq \frac{C_U}{4} \varepsilon_k, \quad (31.4)$$

که u_k جواب (۱۰.۴) با ماتریس انتشار A_k و سمت راست f_k است. بر اساس تخمین اختلال (۱۵.۴) با $\hat{r} > 0$ که کرانی برای حداقل مقدارویژه A_k است، خواهیم داشت:

$$\begin{aligned} \|u - U_{k+1}\|_{H_0^1(\Omega)} &\leq \|u - u_k\|_{H_0^1(\Omega)} + \|u_k - U_{k+1}\|_{H_0^1(\Omega)}, \\ &\leq \hat{r}^{-1} \|f - f_k\|_{H^{-1}(\Omega)} + \hat{r}^{-1} \|\nabla u\|_{L_p(\Omega)} \|A - A_k\|_{L_q(\Omega)} + \frac{C_U}{4} \varepsilon_k. \end{aligned}$$

با بکارگیری (۱۸.۴) و کوچک گرفتن ε_k' (یا w) تا حد ممکن به کران دلخواه دست پیدا می‌کنیم.

$$\begin{aligned} \|u - U_{k+1}\|_{H_0^1(\Omega)} &\leq \hat{r}^{-1} \varepsilon_k' + \hat{r}^{-1} C_p \|f\|_{w^{-1}(L_p(\Omega))} \varepsilon_k' + \frac{C_U}{4} \varepsilon_k \\ &\leq \hat{r}^{-1} (1 + C_p \|f\|_{w^{-1}(L_p(\Omega))}) \varepsilon_k' + \frac{C_U}{4} \varepsilon_k \\ &\leq \hat{r}^{-1} (1 + C_p \|f\|_{w^{-1}(L_p(\Omega))}) w \varepsilon_k + \frac{C_U}{4} \varepsilon_k \\ &\leq C_U \varepsilon_k. \end{aligned} \quad (32.4)$$

همانطور که می‌بینید الگوریتم AFEM دارای سه زیر برنامه اصلی است. با تکرار k ، الگوریتم اولی RHS خواهد بود که تقریب f_k به f را نشان می‌دهد، دومین الگوریتم COEFF است که تقریب A_k به A را نشان می‌دهد. الگوریتم سوم PDE است که نشانه گذاری و نظریف T_{k+1} جهت کاهش خطای تقریب گالرکین از U_{k+1} به u را نشان می‌دهد.

با الگوریتم RHS تقریب f را مشخص می‌کنیم. ورودی‌های این الگوریتم تابع $f \in H^{-1}(\Omega)$ ، افزاز منطبق T و خطای ε است. خروجی‌های الگوریتم عبارتند از:

$$[\hat{T}, \hat{f}] = \text{RHS}(f, T, \varepsilon),$$

که $\hat{T}$ افزاز منطبقی است که T را نظریف می‌کند. $\hat{f}$ چندجمله‌ای تکه‌ای از درجه بزرگتر مساوی m_f وابسته به T است. به گونه‌ای که،

$$\|f - \hat{f}\|_{H^{-1}(\Omega)} \leq \varepsilon, \quad (33.4)$$

برقرار باشد. از نظر تئوری می‌توان الگوریتم RHS را ساخت، اما در عمل برای ساخت این الگوریتم به اطلاعات بیشتری از f نیاز داریم.

تاکنون در بیشتر روش‌های عناصر متناهی تطبیقی فرض بر این است که $f \in L_2(\Omega)$ ، اما اکنون تاکید بیشتر بر سمت راست $f \in H^{-1}(\Omega)$ قرار گرفته است. اگر $f \in L_2(\Omega)$ باشد می‌توان خطا را محدود به تقریب f با استفاده از چندجمله‌ای تکه‌ای کرد. مثلث T و جواب گالرکین U_T به T داده شده است. از برآوردگر باقیمانده که در (۲۴.۴) نمایش دادیم، استفاده می‌کنیم. روی هر مثلث $T \in \mathcal{T}$ ، برآوردگر خطا، خطا را از بالا کنترل می‌کند.

$$\|u - U_T\|_{H_0^1(\Omega)} \lesssim \sum_{T \in \mathcal{T}} \eta_T^2$$

برای کنترل خطا از پایین نیازمند به معرفی عبارت نوسان $^{1^\circ}$ (osc) به صورت زیرهستیم.

$$\sum_{T \in \mathcal{T}} \eta_T^2 \lesssim \|u - U_T\|_{H_0^1(\Omega)}^2 + \text{osc}(f, T)^2,$$

که

$$\text{osc}(f, T) \simeq \varepsilon(f, T),$$

بنابراین خواهیم داشت:

$$\|f - \hat{f}\|_{H^{-1}(\Omega)} \leq \text{osc}(f, T) := \left(\sum_{T \in \mathcal{T}} h_T^2 \|f - a_T(f)\|_{L^2(T)}^2 \right)^{\frac{1}{2}}, \quad (34.4)$$

که $a_T(f)$ تصویر متعامد $L^2(T)$ از f به $p_{m_f}(T)$ ، فضای چند جمله‌ای‌ها از درجه کمتر مساوی m_f روی T و $\hat{f}|_T := a_T(f)$ است. سمت راست (۳۴.۴) را نوسان f روی T می‌نامند. برای درک بهتر به مدلی نیاز داریم تا اطلاعات مورد نیاز از f را تعیین کنیم. در این خصوص و برای بررسی بیشتر به [۱۴] مراجعه کنید.

همچنین باید تقریب A در الگوریتم AFEM را داشته باشیم. با داشتن ماتریس انتشار A ، که کراندار و معین مثبت است، افزاز منطبق T و خطای ε ، خواهیم داشت:

$$[\hat{T}, \hat{A}] = \text{COEFF}(A, T, \varepsilon),$$

^۱° oscillation

خروجی‌ها عبارتند از، افراز منطبق $\hat{T}$ که تعریف شده $\mathcal{T}$ است، ماتریس انتشار $\hat{A}$ که دارای مولفه‌های چندجمله‌ای تکه‌ای با حداکثر درجه m_A و وابسته به $\hat{T}$ است. بنابراین با توجه به الگوریتم COEFF رابطه‌ی زیر را خواهیم داشت:

$$\|A - \hat{A}\|_{L_q(\Omega)} \leq \varepsilon. \quad (35.4)$$

رابطه (35.4)، در $q = \frac{2p}{(p-2)} \in [2, \infty]$ صدق می‌کند. بعلاوه، برای تضمین معین مثبت بودن $\hat{A}$ ، باید ثابت C_2 را داشته باشیم که در رابطه‌ی زیر صدق می‌کند:

$$C_2^{-1} \gamma_{\min}(A) \leq \gamma_{\min}(\hat{A}) \leq \gamma_{\max}(\hat{A}) \leq C_2 \gamma_{\max}(A). \quad (36.4)$$

اگر A ماتریس چندجمله‌ای تکه‌ای از درجه کمتر مساوی m_A ، روی افراز $\mathcal{T}_0$ باشد، پس A را به صورت چندجمله‌ای روی هر عنصر T از هر افراز $\mathcal{T}$ معرفی می‌کنیم، در این صورت نیازی به تقریب A نیست. روش استاندارد برای A کلی، این است که تقریب A را در نورم L_∞ ، با چندجمله‌ای‌های تکه‌ای در نظر بگیریم. به این منظور، A تکه‌ای هموار روی افراز اولیه $\mathcal{T}_0$ در نظر گرفته می‌شود، تا میزان خطای L_∞ کمتر باشد. بنابراین $q = \infty$ است. در نظریه اختلال جدید، این فرض برای A را با روش عناصر متناهی تطبیقی استاندارد حل کردیم. فقط کافیت فرض کنید که شرط p برای برخی $p > 2$ برقرار باشد، پس $q < \infty$ خواهد بود.

۳.۳.۴ اندازه‌گیری عملکرد AFEM

هدف نهایی روش عناصر متناهی تطبیقی، ایجاد تقریب مناسب U به u است. در این صورت عملکرد AFEM با $\|u - U\|_{H^1_0(\Omega)}$ اندازه‌گیری می‌شود.

تقریب u برای هر $n \geq 1$ ، $\sum_n^{m_u}$ (بزرگترین درجه چندجمله‌ای) را اجتماعی از همه فضاها عناصر متناهی $V(\mathcal{T}) \subset H^1_0(\Omega)$ با $\mathcal{T} \in \mathcal{F}_n$ (مجموعه افرازشای ناسازگار که با n تعریف از $\mathcal{T}_0$ بدست می‌آید) تعریف می‌کنیم.

با داشتن تابع σ_n ، $v \in H^1_0(\Omega)$ بهترین خطای تقریب v با عناصر $\sum_n^{m_u}$ را نشان می‌دهد.

$$\sigma_n(v)_{H^1_0(\Omega)} := \sigma_n^{m_u}(v)_{H^1_0(\Omega)} := \inf_{V \in \sum_n^{m_u}} \|v - V\|_{H^1_0(\Omega)}, \quad n \geq 1, \quad (37.4)$$

با استفاده از فضای $H^1_0(\Omega)$ را داخل رده‌های تقریب طبقه طبقه می‌کنیم:

برای هر $s > 0$ ، $\mathcal{A}^s := \mathcal{A}^s(\mathcal{T}_0, H^1_0(\Omega))$ را به صورت مجموعه‌ای از همه $v \in H^1_0(\Omega)$ تعریف می‌کنیم:

$$\mathcal{A}^s := \{v \in V : |v|_{\mathcal{A}^s} := \sup_{n > 0} (n^s \inf_{v \in V(\mathcal{T})} \|v - V\|_{H^1_0(\Omega)})\},$$

بنابراین داریم:

$$|v|_{\mathcal{A}^s} := \sup_{n > 0} (n^s \sigma_n(v)_{H^1_0(\Omega)}) < \infty; \quad (38.4)$$

رابطه (38.4) را می‌توان برای دامنه تغییرات s ، یعنی $0 < s < S$ که $S = \frac{m_u}{d}$ کرانی روی مرتبه‌ای از تقریب اعمال شده با درجه چندجمله‌ای m_u است.

با توجه به رابطه $\#\bar{T} - \#\mathcal{T}_\circ \leq C_\circ n(T)$ ، برای هر $v \in \mathcal{A}^s$ ، $S_n \in \sum_n^{m_u}$ ای وجود دارد که وابسته به افراز منطبق است. بنابراین داریم:

$$|v|_{\mathcal{A}^s} := \sup_{n \geq \circ} ((C_\circ n)^s \inf_{S_n \in \sum_n^{m_u}} \|v - S_n\|_{H^1(\Omega)})$$

که

$$\|v - S_n\|_{H^1(\Omega)} \leq C_\circ |v|_{\mathcal{A}^s} n^{-s}, \quad n \geq 1, \quad (39.4)$$

در (39.4) مقدار $C_\circ$ ثابت است. همچنین می‌توانیم رده تقریب مشابه و نیم نرم معادل داشته باشیم، اگر رده تقریب $\mathcal{A}^s$ را با شرط منطبق بودن افرازاها تعریف کنیم. همان طور که قبلاً هم ذکر کردیم، ورودی‌های الگوریتم PDE تقریب $\hat{f}$ و $\hat{A}$ به f و A است. پس الگوریتم PDE، تقریبی برای جواب $\hat{u}$ از (8.4) با ضریب انتشار $\hat{A}$ و سمت راست $\hat{f}$ ایجاد می‌کند.

تعریف ۱.۳.۴. (تقریب ε از مرتبه s). با داشتن $u \in \mathcal{A}^s$ و $\varepsilon > \circ$ ، اگر $\|u - v\|_{H^1(\Omega)} \leq \varepsilon$ باشد، آنگاه تابع v تقریب ε از مرتبه s به u خواهد بود. ثابت C مستقل از ε ، u و v وجود دارد به گونه‌ای که برای همه $\varepsilon \geq \delta$ ، $n \in N$ ای با شرایط زیر خواهیم داشت:

$$\sigma_n(v)_{H^1(\Omega)} \leq \delta, \quad n \leq C |u|_{\mathcal{A}^s}^{\frac{1}{s}} \delta^{-\frac{1}{s}}. \quad (40.4)$$

توجه کنید که اگر $\varepsilon_1 < \varepsilon_2$ ، و v تقریب ε_1 از مرتبه s به u باشد، آنگاه v تقریب ε_2 از مرتبه s به u خواهد بود.

لم ۲.۳.۴. (تقریب ε از مرتبه s). فرض کنید $u \in \mathcal{A}^s(\mathcal{T}_\circ, H^1(\Omega))$ و $v \in H^1(\Omega)$ برای برخی $\varepsilon > \circ$ در نامساوی $\|u - v\|_{H^1(\Omega)} \leq \varepsilon$ صدق کنند. آنگاه v ، تقریب 2ε از مرتبه s به u خواهد بود.

برهان. فرض کنید $\delta \geq 2\varepsilon$ باشد. از نامساوی مثلی استفاده می‌کنیم:

$$\delta \geq 2\varepsilon \rightarrow \varepsilon \leq \frac{\delta}{2},$$

$$\begin{aligned} \sigma_n(v)_{H^1(\Omega)} &\leq \|v - V\|_{H^1(\Omega)} = \|v - u + u - V\|_{H^1(\Omega)} \\ &\leq \|u - v\|_{H^1(\Omega)} + \|u - V\|_{H^1(\Omega)} \\ &\leq \|u - v\|_{H^1(\Omega)} + \sigma_n(u)_{H^1(\Omega)} \\ &\leq \varepsilon + \sigma_n(u)_{H^1(\Omega)}, \end{aligned}$$

با توجه به $\varepsilon \leq \frac{\delta}{2}$ داریم:

$$\sigma_n(v)_{H^1(\Omega)} \leq \frac{\delta}{2} + \sigma_n(u)_{H^1(\Omega)}. \quad (41.4)$$

چون $u \in \mathcal{A}^s(\mathcal{T}_\circ, H^1(\Omega))$ ، لذا:

$$\frac{\delta}{2} + \sigma_n(u)_{H^1(\Omega)} \leq \delta \rightarrow \sigma_n(u)_{H^1(\Omega)} \leq \delta - \frac{\delta}{2} = \frac{\delta}{2},$$

یعنی،

$$n \leq |u|_{\mathcal{A}^s}^{\frac{1}{s}} \left(\frac{\delta}{2}\right)^{-\frac{1}{s}}.$$

بنابراین از تخمین (۴۰.۴) خواهیم داشت: $c = 2^{\frac{1}{s}}$

□

اگر ثابت مطلق C_3 وجود داشته باشد به گونه‌ای که تعداد عناصر $N(\hat{u})$ علامت گذاری شده برای تقریب روی $\mathcal{T}$ ، برای دست یافتن به خطای $\|\hat{u} - \hat{U}\|_{H^1(\Omega)} \leq \varepsilon$ باشد، آنگاه الگوریتم PDE عملکرد بهینه‌ای در $\mathcal{A}^s$ خواهد داشت. به طوری که رابطه زیر را داریم:

$$N(\hat{u}) \leq C_3 |\hat{u}|_{\mathcal{A}^s}^{\frac{1}{s}} \varepsilon^{\frac{-1}{s}}, \quad (42.4)$$

زمانی که جواب $\hat{u}$ در (۱۰.۴) با داده‌های $\hat{A}$ و $\hat{f}$ در $\mathcal{A}^s(\mathcal{T}_0, H^1(\Omega))$ باشند، رابطه (۴۲.۴) صحیح می‌باشد. رابطه (۴۲.۴) زمانی به کار گرفته می‌شود که C_3 را در اختیار داشته باشیم. تقریب f . عملکرد تقریب f را نیز می‌توانیم به همین شیوه اندازه‌گیری کنیم. فرض کنید $\sum_n^{m_f}$ فضای چندجمله‌ای‌های S از درجه بزرگتر مساوی m_f وابسته به افراز $\mathcal{T} \in \mathcal{F}_n$ باشد. بنابراین تعریف می‌کنیم:

$$\sigma_n(f)_{H^{-1}(\Omega)} := \sigma_n^{m_f}(f)_{H^{-1}(\Omega)} := \inf_{S \in \sum_n^{m_f}} \|f - S\|_{H^{-1}(\Omega)}, \quad n \geq 1,$$

در مقایسه با رده $\mathcal{A}^s$ ، فرض می‌کنیم که رده $\mathcal{B}^s := \mathcal{B}^s(\mathcal{T}_0, H^{-1}(\Omega))$ شامل همه توابع $f \in H^{-1}(\Omega)$ به صورت زیر می‌باشد.

$$|f|_{\mathcal{B}^s} := \sup_{n \geq 1} (n^s \sigma_n(f)_{H^{-1}(\Omega)}) < \infty; \quad (43.4)$$

همچنین باید تقریب توابع در سایر نورم‌ها را در نظر بگیریم. اگر $0 < q \leq \infty$ و $g \in L_q(\Omega)$ باشد (در صورتی که $q = \infty$ باشد)، خواهیم داشت:

$$\sigma_n(f)_{L_q(\Omega)} := \sigma_n^{m_f}(f)_{L_q(\Omega)} := \inf_{S \in \sum_n^{m_f}} \|f - S\|_{L_q(\Omega)}, \quad n \geq 1,$$

رده‌های تقریب مشابه $\mathcal{B}^s(L_q(\Omega)) := \mathcal{B}^s(\mathcal{T}_0, L_q(\Omega))$ ، $s \geq 0$ شامل همه توابع $L_q(\Omega)$ می‌باشند که به صورت زیر است:

$$|f|_{\mathcal{B}^s(L_q(\Omega))} := \sup_{n \geq 1} (n^s \sigma_n(f)_{L_q(\Omega)}) < \infty; \quad (44.4)$$

اکنون عملکردهای الگوریتم RHS را بررسی می‌کنیم. اگر $f \in \mathcal{B}^s(\mathcal{T}_0, H^{-1}(\Omega))$ باشد، افرازهای $\mathcal{T}^*$ با $|f|_{\mathcal{B}^s}^{\frac{1}{s}} \varepsilon^{\frac{-1}{s}}$ وجود دارند. چندجمله‌ای تکه‌ای S را می‌توانیم به گونه‌ای که $\|f - S\|_{H^{-1}(\Omega)} \leq \varepsilon$ بیابیم. با داشتن افراز $\mathcal{T}$ بدست آمده از روند تقریب $\mathcal{T}_0$ ، پوشش $\mathcal{T}^* \oplus \mathcal{T}$ از $\mathcal{T}^*$ با $\mathcal{T}$ را به صورت زیر در نظر می‌گیریم:

$$\#(\mathcal{T}^* \oplus \mathcal{T}) - \#\mathcal{T} \leq \#\mathcal{T}^* - \#\mathcal{T}_0 \leq |f|_{\mathcal{B}^s}^{\frac{1}{s}} \varepsilon^{\frac{-1}{s}}, \quad (45.4)$$

اگر $\mathcal{T}_1, \mathcal{T}_2 \in \mathcal{T}$ باشند آنگاه،

$$\#(\mathcal{T}_1 \oplus \mathcal{T}_2) \leq \#\mathcal{T}_1 + \#\mathcal{T}_2 - \#\mathcal{T}_0$$

توجه کنید که در این مرحله ممکن است افرازها منطبق نباشند. بدین ترتیب، اگر تعداد $N(f)$ از عناصر منتخب برای تقریب با الگوریتمی برای رسیدن به خطای ε داشته باشیم، آنگاه الگوریتم RHS رده عملکرد بهینه بر B^s خواهد داشت. رابطه‌ی زیر را با ثابت C_3 خواهیم داشت:

$$N(f) \leq C_3 |f|_{B^s}^{\frac{1}{s}} \varepsilon^{-\frac{1}{s}}, \quad (46.4)$$

برای ساخت چنین الگوریتمی به بخش پنجم مراجعه می‌کنیم. تقریب A . با $\sum_n^{m_A}$ رده ماتریس‌های چندجمله‌ای تکه‌ای از درجه کمتر مساوی m_A ، وابسته به افراز $\mathcal{T} \in \mathcal{F}_n$ را نمایش می‌دهیم. بهترین خطای تقریب از A در $\sum_n^{m_A}$ را با $\sigma_n(A)_{L_q(\Omega)}$ نشان می‌دهیم:

$$\sigma_n(A)_{L_q(\Omega)} := \sigma_n^{m_A}(A)_{L_q(\Omega)} := \inf_{S \in \sum_n^{m_A}} \|A - S\|_{L_q(\Omega)},$$

با $\mathcal{M}^s := \mathcal{M}^s(\mathcal{T}_*, L_q(\Omega))$ رده همه ماتریس‌ها را نشان می‌دهیم. به گونه‌ای که

$$|A|_{\mathcal{M}^s} := \sup_{n \geq 1} (n^s \sigma_n(A)_{L_q(\Omega)}) < \infty; \quad (47.4)$$

این تخمین بر اساس تقریب‌پذیری L_q انجام می‌شود. در عملکرد الگوریتم COEFF، برای اینکه ماتریس $\hat{A}$ را در تخمین (۱۵.۴) بکار ببریم باید این ماتریس به صورت معین مثبت باشد. در بخش پنجم نشان خواهیم داد که اگر کران‌های (۱۱.۴) برای مقادیر ویژه A را بدانیم، با داشتن ثابت C_4 و ماتریس چندجمله‌ای تکه‌ای $\hat{A}$ از درجه کمتر مساوی m_A خواهیم داشت:

$$\begin{aligned} \|A - \hat{A}\|_{L_q(\Omega)} &\leq \|A - S\|_{L_q(\Omega)} + \|\hat{A} - S\|_{L_q(\Omega)} \\ &\leq \|A - S\|_{L_q(\Omega)} + \hat{M} \|A - S\|_{L_q(\Omega)} \\ &\leq (1 + \hat{M}) \|A - S\|_{L_q(\Omega)} \end{aligned}$$

بنابراین داریم،

$$\|A - \hat{A}\|_{L_q(\Omega)} \leq C_4 \sigma_n(A)_{L_q(\Omega)}, \quad (48.4)$$

مقادیر ویژه $\hat{A}$ در (۱۱.۴)، برای برخی $\hat{r}$ و $\hat{M}$ قابل مقایسه با r و M هستند. و در عین حال ثابت C_4 ، مستقل از r ، M و n می‌باشد.

در مقایسه برای الگوریتم PDE و RHS، اگر تعداد $N(A)$ از عناصر علامت‌گذاری شده برای تقریب و خطای ε با شروع از افراز $\mathcal{T}$ بدست آمده باشد، آنگاه الگوریتم COEFF دارای رده عملکرد بهینه روی $\mathcal{M}^s$ می‌باشد. از (۴۷.۴) خواهیم داشت:

$$|A|_{\mathcal{M}^s} \geq c^s n^s \|A - S\|_{L_q(\Omega)} \rightarrow n \leq c^{-1} |A|_{\mathcal{M}^s}^{\frac{1}{s}} \varepsilon^{-\frac{1}{s}}$$

بنابراین داریم،

$$N(A) \leq C_3 |A|_{\mathcal{M}^s}^{\frac{1}{s}} \varepsilon^{-\frac{1}{s}}, \quad (49.4)$$

که C_3 ثابت مطلق است. برای ساخت این الگوریتم به بخش پنجم مراجعه می‌کنیم. همانطور که قبلاً در مورد رده تقریب A^s گفتیم، اگر افرازهای منطبق داشته باشیم آنگاه رده‌های تقریب مشابه $B^s = B^s(\mathcal{T}_*, H^{-1}(\Omega))$ و $\mathcal{M}^s = \mathcal{M}^s(\mathcal{T}_*, L_q(\Omega))$ را خواهیم داشت. کاربرد AFEM. رده‌های تقریب A^s ، B^s ، $\mathcal{M}^s$ مفروض هستند. هدف از به کار بردن الگوریتم AFEM

این است که، $f \in B^s$ و $A \in M^s$ باشد. آنگاه AFEM دنباله $\{T_k\}_{k \geq 0}$ از مثلث بندی تودرتو ($T_{k+1} \geq T_k$ for each $k \geq 0$) ایجاد می‌کند، به گونه‌ای که برای $k \geq 1$ داریم:

$$\|u - U_k\|_{H_0^1(\Omega)} \leq C(|u|_{A^s} + |f|_{B^s} + |A|_{M^s})(\#T_k - \#T_0)^{-s}, \quad (50.4)$$

اثبات مشابه اثبات قضیه ۱۰۴.۴ در بخش چهارم است. در (۵۰.۴)، C مقدار ثابتی است که فقط به s بستگی دارد. کران (۵۰.۴) براساس [۴]، [۲۵]، [۱۲] می‌باشد. اخیراً در [۱۴] نشان داده شده که اگر $A = aI$ ، a تابع ثابت تکه‌ای و I ماتریس واحد باشد، $u \in A^s$ بر $f \in B^s$ دلالت خواهد داشت با این شرط که $s < S$ باشد.

۴.۳.۴ تقریب ε و عملکرد بهینه رده

همانطور که در بخش سوم گفتیم، در الگوریتم PDE ضریب انتشار و مقادیر سمت راست در تکرارها تغییر می‌کنند، بنابراین بکارگیری این الگوریتم دشوار است. برای بکارگیری (۴۲.۴) باید اثبات‌های ذکر شده در [۴]، [۲۴]، [۱۲] را مشاهده کنیم.

فرض کنید $\hat{u} \in H_0^1(\Omega)$ جواب (۱۰.۴) با داده‌های $\hat{A}$ و $\hat{f}$ باشد. همچنین $\hat{U} = \text{GAL}(T, \hat{A}, \hat{f}) \in V(T)$ جواب گالرکین وابسته به افراز T است. فرض کنید $\hat{\varepsilon} := \|\hat{u} - \hat{U}\|_{H_0^1(\Omega)}$ خطای بدست آمده با جواب گالرکین روی T است.

کنترل تعداد عناصر منتخب با الگوریتم PDE براساس کوچک‌ترین افراز T^* انجام می‌شود. که دقت آن برای برخی $0 < \mu \leq 1$ برابر با $\mu \hat{\varepsilon}$ است. اگر M مجموعه‌ای از عناصر منتخب برای تقریب باشد، آنگاه می‌توانیم $\#M$ را با $\#(T^* \oplus T) - \#T_0$ مقایسه کنیم. بنابراین داریم:

$$\#T^* - \#T_0 \leq c \|\hat{u} - \hat{U}\|_{H_0^1(\Omega)}^{-\frac{1}{s}} |\hat{u}|_{A^s}^{\frac{1}{s}},$$

با در نظر گرفتن $\mu \hat{\varepsilon} := \|\hat{u} - \hat{U}\|_{H_0^1(\Omega)}$ خواهیم داشت:

$$\#M \leq \tilde{C}_3 |\hat{u}|_{A^s}^{\frac{1}{s}} (\mu \hat{\varepsilon})^{-\frac{1}{s}}, \quad (51.4)$$

رابطه (۵۱.۴) زمانی که $\hat{u} \in A^s$ و $\tilde{C}_3$ ثابت مطلق وابسته به s باشد، برقرار است.

ابتدا، مهم است بدانیم که استدلال منجر به (۵۱.۴) نیازمند نظم کامل $\hat{u} \in A^s$ نیست. اما $\hat{u}$ تقریب $\mu \hat{\varepsilon}$ از مرتبه s به $v \in A^s$ است.

دوم اینکه، چندین تکرار فرعی در $\text{PDE}(T, \hat{A}, \hat{f}, \varepsilon)$ برای رسیدن به خطای ε لازم داریم. در هر تکرار، تعدادی از عناصر را براساس (۵۱.۴)، با ε به جای $\hat{\varepsilon}$ انتخاب می‌کند، و تعداد تکرارهای فرعی براساس نسبت $\frac{\varepsilon}{\hat{\varepsilon}}$ قرار دارد. در الگوریتم جدید AFEM، این نسبت را محدود نگه داشته‌ایم. بنابراین تعداد تکرارهای فرعی در الگوریتم PDE کراندار یکنواخت هستند.

در پایان با ادغام لم (۲۳.۴) با (۵۱.۴) در می‌یابیم که $N = N(\hat{u})$ ، تعداد عناصر علامت‌گذاری شده برای تقریب با الگوریتم $\text{PDE}(T, \hat{A}, \hat{f}, \varepsilon)$ برای رسیدن به خطای ε است. بنابراین رابطه‌ی زیر با این شرط که $\hat{u}$ تقریب $\mu \hat{\varepsilon}$ از مرتبه s به $u \in A^s(T_0, H_0^1(\Omega))$ باشد، برقرار است.

$$N \leq C_3 |u|_{A^s}^{\frac{1}{s}} \varepsilon^{-\frac{1}{s}}, \quad (52.4)$$

همچنین در الگوریتم PDE نسبت $\frac{\varepsilon}{\varepsilon_0}$ بطور یکنواخت کراندار است. از این واقعیت در تحلیل جدید AFEM استفاده خواهیم کرد.

۴.۴ AFEM برای ماتریس‌های انتشار ناپیوسته (DISC)

اکنون در موقعیتی هستیم که الگوریتم جدید AFEM را با نام (DISC)^{۱۱} نمایش می‌دهیم. الگوریتم DISC شامل سه بخش اصلی RHS، COEFF و PDE است. الگوریتم‌های RHS و PDE استاندارد هستند. از الگوریتم COEFF برای یافتن تقریب A در $L_q(\Omega)$ به جای $L_\infty(\Omega)$ برای برخی $q = \frac{2p}{(p-2)}$ استفاده کردیم.

فرض می‌کنیم سه الگوریتم RHS، COEFF و PDE با رده بهینه برای همه $0 < s < S$ با $0 < s < S$ شناخته شده هستند. برای اینکه جواب u از (۱۰.۴) مناسب $\|\nabla u\|_{L_p(\Omega)} \leq C_p \|f\|_{w^{-1}(L_p(\Omega))}$ باشد، الگوریتم DISC را با ورودی‌های افزاز اولیه $\mathcal{T}$ ، خطای اولیه ε_1 ، ماتریس A و سمت راست f به کار می‌بریم.

مبحث شرط p در بخش دوم را بخوانید. اکنون مقادیر ثابت را بصورت $w > 0$ و $\beta < 1$ ثابت می‌کنیم. با توجه به (۱۵.۴) خواهیم داشت:

$$\|u - u_k\|_{H^1(\Omega)} \leq \hat{r}^{-1} \|f - f_k\|_{H^{-1}(\Omega)} + \hat{r}^{-1} \|\nabla u\|_{L_p(\Omega)} \|A - A_k\|_{L_q(\Omega)}$$

بنا به کران (۳۶.۴) برای A_k داریم:

$$\begin{aligned} \|u - u_k\|_{H^1(\Omega)} &\leq C_2 r^{-1} w \varepsilon_k + C_2 r^{-1} C_p \|f\|_{w^{-1}(L_p(\Omega))} w \varepsilon_k \\ &\leq C_2 r^{-1} (1 + C_p \|f\|_{w^{-1}(L_p(\Omega))}) w \varepsilon_k \end{aligned}$$

از طرفی از (۲۵.۴) داریم:

$$C_L \mu \varepsilon_k \leq \|u - u_k\|_{H^1(\Omega)} \leq C_2 r^{-1} (1 + C_p \|f\|_{w^{-1}(L_p(\Omega))}) w \varepsilon_k$$

بنابراین w به صورت زیر است:

$$w \leq \frac{r \mu C_L}{2 C_2 (1 + C_p \|f\|_{w^{-1}(L_p(\Omega))})} \quad (۵۳.۴)$$

که $\mu \leq 1$ همان مقدار ثابت بخش سوم است. C_2 در کران یکنواخت (۳۶.۴) روی مقادیر ویژه به کار گرفته شده و C_L ثابت کران پایین در (۲۵.۴) می‌باشد.

با داشتن افزاز اولیه $\mathcal{T}$ و پارامترهای ε ، w و β ، تکرارها در الگوریتم DISC با $k := 0$ به صورت زیر خواهد بود:

$$\begin{aligned} [\mathcal{T}_k(f), f_k] &= \text{RHS}(\mathcal{T}_k, f, w \varepsilon_k); \\ [\mathcal{T}_k(A), A_k] &= \text{COEFF}(\mathcal{T}_k(f), A, w \varepsilon_k); \\ [\mathcal{T}_{k+1}, U_{k+1}] &= \text{PDE}(\mathcal{T}_k(A), A_k, f_k, \frac{\varepsilon_k}{2}); \\ \varepsilon_{k+1} &= \beta \varepsilon_k; \quad k \leftarrow k + 1. \end{aligned}$$

^{۱۱}Discontinus

قضیه زیر بهینگی الگوریتم DISC را نشان می‌دهد.

قضیه ۱.۴.۴. (بهینگی الگوریتم DISC). فرض کنید سه الگوریتم $COEFF$ ، RHS و PDE برای همه $0 < s \leq S$ و برخی $0 < s_f \leq S$ علاوه فرض کنید f با $0 < s_f \leq S$ در $B^{s_f}(H^{-1}(\Omega))$ قرار دارد. همچنین شرط p برای برخی $p > 2$ برقرار است. ماتریس انتشار A ، در $L_\infty(\Omega)$ و در $\mathcal{M}^{s_A}(L_q(\Omega))$ برای $q = \frac{2p}{(p-2)}$ و $0 < s_A \leq S$ معین مثبت است.

فرض کنید $\mathcal{T}_0$ افراز اولیه و $U_k \in V(\mathcal{T})$ جواب گالرکین بدست آمده بعد از k تکرار از الگوریتم DISC باشد. بنابراین، هر زمان که برای $0 < s_u \leq S$ ، $u \in \mathcal{A}^{s_u}(H_0^1(\Omega))$ ، آنگاه برای $k \geq 1$ داریم:

$$\|u - U_k\|_{H_0^1(\Omega)} \leq C_U \varepsilon_{k-1}, \quad (54.4)$$

که ثابت C_U کران بالا در (۲۵.۴) است. و

$$\#\mathcal{T}_k - \#\mathcal{T}_0 \leq C_{\sharp}(|u|_{\mathcal{A}^s(H_0^1(\Omega))} + |f|_{B^s(H^{-1}(\Omega))} + |A|_{\mathcal{M}^s(L_q(\Omega))}) \varepsilon_{k-1}^{-\frac{1}{s}} \quad (55.4)$$

با

$$C_{\sharp} := \frac{C_0 C_{\sharp} w^{-\frac{1}{s}}}{1 - \beta^{\frac{1}{s}}}, \quad s = \min(s_u, s_A, s_f).$$

برهان. ابتدا (۵۴.۴) را اثبات می‌کنیم. با u_k ، جواب (۱۰.۴) با ضریب انتشار A_k و سمت راست f_k را نشان می‌دهیم. با توجه به (۱۵.۴) و شرط p در بخش دوم خواهیم داشت:

$$\|u - u_k\|_{H_0^1(\Omega)} \leq \hat{r}^{-1} \|f - f_k\|_{H^{-1}(\Omega)} + \hat{r}^{-1} \|\nabla u\|_{L_p(\Omega)} \|A - A_k\|_{L_q(\Omega)}$$

بعلاوه، بنا به محدودیت (۵۳.۴) برای w و کران (۳۶.۴) برای A_k داریم:

$$\begin{aligned} \|u - u_k\|_{H_0^1(\Omega)} &\leq C_{\sharp} r^{-1} w \varepsilon_k + C_{\sharp} r^{-1} C_p \|f\|_{w^{-1}(L_p(\Omega))} w \varepsilon_k \\ &\leq C_{\sharp} r^{-1} (1 + C_p \|f\|_{w^{-1}(L_p(\Omega))}) w \varepsilon_k \\ &\leq C_{\sharp} r^{-1} (1 + C_p \|f\|_{w^{-1}(L_p(\Omega))}) \frac{r \mu C_L}{2 C_{\sharp} (1 + c_p \|f\|_{w^{-1}(L_p(\Omega))})} \varepsilon_k \\ &\leq \frac{\mu \varepsilon_k C_L}{2}, \end{aligned} \quad (56.4)$$

با بررسی (۲۵.۴) و (۲۶.۴) در می‌یابیم که U_{k+1} در رابطه‌ی زیر صدق می‌کند:

$$\begin{aligned} \|u_k - U_{k+1}\|_{H_0^1(\Omega)} &\leq C_U \eta_{k+1}(U_{k+1}, A_k, f_k, \mathcal{T}_{k+1}) \\ &\leq C_U \frac{\varepsilon_k}{2}, \end{aligned} \quad (57.4)$$

از طرفی می‌دانیم که $C_L \leq C_U$ و $\mu \leq 1$ می‌باشد. بنابراین با اضافه و کم کردن u_k به $\|u_k - U_{k+1}\|$ و استفاده از نامساوی مثلثی داریم:

$$\begin{aligned} \|u - U_{k+1}\|_{H_0^1(\Omega)} &\leq \|u - u_k + u_k - U_{k+1}\|_{H_0^1(\Omega)} \\ &\leq \|u - u_k\|_{H_0^1(\Omega)} + \|u_k - U_{k+1}\|_{H_0^1(\Omega)} \\ &\leq \frac{\mu \varepsilon_k C_L}{2} + C_U \frac{\varepsilon_k}{2} \\ &\leq C_U \frac{\varepsilon_k}{2} + C_U \frac{\varepsilon_k}{2} \\ &\leq C_U \varepsilon_k, \end{aligned}$$

حال (۵۵.۴) را اثبات می‌کنیم. در هر مرحله j از الگوریتم DISC، افراز جدید $\mathcal{T}_j$ از $\mathcal{T}_{j-1}$ براساس عناصر منتخب برای تعریف و سایر احتمالات برای منطبق بودن $\mathcal{T}_j$ شکل می‌گیرد. با $N_j(f)$ ، $N_j(A)$ و $N_j(u)$ به ترتیب تعداد عناصر منتخب برای تعریف با الگوریتم‌های RHS، COEFF و PDE را نشان می‌دهیم. $N_j(f)$ و $N_j(A)$ به ترتیب در (۴۶.۴) و (۴۹.۴) آمده است. از کران (۲۲.۴) برای تعاریفات بیشتر و ایجاد زیر تقسیمات منطبق استفاده می‌کنیم. یعنی

$$\#\mathcal{T}_k - \#\mathcal{T}_0 \leq C_0 \sum_{j=0}^{k-1} (N_j(f) + N_j(A) + N_j(u)),$$

با توجه به فرض مساله $s = \min(s_u, s_A, s_f)$ است، یعنی $s \leq s_f, s_A$ می‌باشد. بنابراین فرضیات بهینگی رده RHS و COEFF به طور مستقیم نشان می‌دهد که،

$$N_j(f) \leq C_3 |f|_{B^s(H^{-1}(\Omega))}^{\frac{1}{s}} (w\varepsilon_j)^{-\frac{1}{s}},$$

$$N_j(A) \leq C_3 |A|_{M^s(L_q(\Omega))}^{\frac{1}{s}} (w\varepsilon_j)^{-\frac{1}{s}},$$

از آنجاییکه N_j تحت تاثیر u_j ، جواب (۱۰.۴) با ضرایب انتشار A_j و سمت راست f_j است، نمی‌توانیم به طور مستقیم $N_j = N_j(u)$ را برای کران N_j استفاده کنیم. فرض کنید $U_j(A) = \text{GAL}(\mathcal{T}_j(A), A_j, f_j)$ و $\hat{\varepsilon}_j := \eta_{\mathcal{T}_j(A)}(U_j(A), A_j, f_j, \mathcal{T}_j(A))$ باشد. با توجه به (۲۵.۴) داریم:

$$\|u_j - U_j(A)\|_{H^1(\Omega)} \geq C_L \hat{\varepsilon}_j$$

فرض کنید که $\hat{\varepsilon}_j > \frac{\varepsilon_j}{4}$ باشد. در غیر این صورت، اگر $\hat{\varepsilon}_j < \frac{\varepsilon_j}{4}$ باشد، می‌توان از $\text{PDE}(\mathcal{T}_j(A), A_j, f_j, \frac{\varepsilon_j}{4})$ صرف نظر کرد و $N_j = 0$ خواهد شد.

با توجه به مباحثی که در بخش سوم داشتیم، به محض اینکه اثبات کنیم نسبت $\frac{\hat{\varepsilon}_j}{\varepsilon_j}$ با توجه به تکرار شمارش j و اینکه u_j تقریب $\mu C_L \varepsilon_j$ از مرتبه s به u است، به طور یکنواخت کراندار باشد، تخمین (۵۲.۴) معتبر خواهد بود. این مورد آخر را بر اساس تخمین $\|u - u_j\|_{H^1(\Omega)} \leq \mu \varepsilon_j^{\frac{C_L}{4}}$ در (۵۳.۴) و لم (۲۳.۴) در می‌یابیم. اکنون فقط اثبات کران یکنواخت $\frac{\hat{\varepsilon}_j}{\varepsilon_j}$ باقی مانده است. با یادآوری ویژگی تصویر گالرکین داریم:

$$\|A_j^{\frac{1}{r}} \nabla(u_j - U_j(A))\|_{L_r(\Omega)} \leq \|A_j^{\frac{1}{r}} \nabla(u_j - V)\|_{L_r(\Omega)}$$

این ویژگی برای هر $v \in V(\mathcal{T}_j(A))$ به ویژه $v \in V(\mathcal{T}_j(A)) \subset V(\mathcal{T}_j)$ صدق می‌کند. چون $\mathcal{T}_j(A)$ تعریف شده $\mathcal{T}_j$ می‌باشد. اگر $r_j = \gamma_{\min}(A_j)$ و $M_j = \gamma_{\max}(A_j)$ به ترتیب نشانگر حداقل و حداکثر مقدار ویژه A_j باشند. پس با توجه به خاصیت تصویر گالرکین و کران پایین در (۲۵.۴) و (۳۶.۴) داریم:

$$\begin{aligned} \hat{\varepsilon}_j &\leq C_L^{-1} \|u_j - U_j(A)\|_{H^1(\Omega)} \\ &\leq C_L^{-1} \left(\frac{M_j}{r_j}\right)^{\frac{1}{r}} \|u_j - U_j\|_{H^1(\Omega)} \\ &\leq C_L^{-1} C_2 \left(\frac{M}{r}\right)^{\frac{1}{r}} (\|u - U_j\|_{H^1(\Omega)} + \|u - u_j\|_{H^1(\Omega)}) \end{aligned}$$

با ادغام (۵۴.۴) با $\varepsilon_j = \beta \varepsilon_{j-1}$ و (۵۶.۴)، کران زیر را بدست می‌آوریم:

$$\begin{aligned} \hat{\varepsilon}_j &\leq C_L^{-1} C_{\mathcal{T}} \left(\frac{M}{r}\right)^{\frac{1}{s}} (C_U \varepsilon_{j-1} + \mu \varepsilon_j \frac{C_L}{\mathcal{T}}) \\ &\leq C_L^{-1} C_{\mathcal{T}} \left(\frac{M}{r}\right)^{\frac{1}{s}} (C_U \frac{\varepsilon_j}{\beta} + \mu \varepsilon_j \frac{C_L}{\mathcal{T}}) \\ &\leq \varepsilon_j C_L^{-1} C_{\mathcal{T}} \left(\frac{M}{r}\right)^{\frac{1}{s}} \left(\frac{C_U}{\beta} + \frac{\mu C_L}{\mathcal{T}}\right) \end{aligned}$$

کران یکنواختی برای $\frac{\hat{\varepsilon}_j}{\varepsilon_j}$ به صورت زیر حاصل می‌شود:

$$\frac{\hat{\varepsilon}_j}{\varepsilon_j} \leq C_L^{-1} C_{\mathcal{T}} \left(\frac{M}{r}\right)^{\frac{1}{s}} \left(\frac{C_U}{\beta} + \frac{\mu C_L}{\mathcal{T}}\right), \quad (58.4)$$

مبحث ارایه شده در بخش سوم، کران (۵۲.۴) را تضمین می‌کند، یعنی

$$N_j \leq C_{\mathcal{T}} |u|_{A^{s_u}(H^1(\Omega))}^{\frac{1}{s}} \varepsilon_j^{-\frac{1}{s}} \quad (59.4)$$

با جمع بندی کران‌های $N_j(u)$ ، $N_j(A)$ و $N_j(f)$ و استفاده از $w < 1$ خواهیم داشت:

$$\begin{aligned} \#\mathcal{T}_k - \#\mathcal{T}_0 &\leq C_0 \sum_{j=0}^{k-1} (N_j(f) + N_j(A) + N_j(u)) \\ &\leq C_0 \sum_{j=0}^{k-1} (C_{\mathcal{T}} |f|_{B^s(H^{-1}(\Omega))} (w \varepsilon_j)^{-\frac{1}{s}} \\ &\quad + C_{\mathcal{T}} |A|_{M^s(L_q(\Omega))} (w \varepsilon_j)^{-\frac{1}{s}} \\ &\quad + C_{\mathcal{T}} |u|_{A^s(H^1(\Omega))} (\varepsilon_j)^{-\frac{1}{s}}) \\ &\leq C_0 C_{\mathcal{T}} w^{-\frac{1}{s}} (|A|_{M^s(L_q(\Omega))} + |f|_{B^s(H^{-1}(\Omega))} + |u|_{A^s(H^1(\Omega))}) \sum_{j=0}^{k-1} \varepsilon_j^{-\frac{1}{s}} \end{aligned} \quad (60.4)$$

تخمین مورد نظر (۵۵.۴) پس از نوشتن $\varepsilon_j = \beta^{k-j} \varepsilon_k$ و با بکارگیری $0 < \beta < 1$ بدست می‌آید، یعنی $\square$

$$\sum_{j=0}^{k-1} \beta^{\frac{j}{s}} \leq (1 - \beta^{\frac{1}{s}})^{-1}$$

توجه کنید که نتیجه قضیه (۵۳.۴) را نیز می‌توان بصورت زیر بیان کرد:

$$\|u - U_k\|_{H^1(\Omega)} \leq C(|u|_{A^s} + |f|_{B^s} + |A|_{M^s})(\#\mathcal{T}_k - \#\mathcal{T}_0)^{-s}, \quad (61.4)$$

در این رابطه C مستقل از k است. الگوریتم DISC عملکرد بهینه براساس (۵۰.۴) خواهد داشت.

تذکر ۲.۴.۴. (حالت $s < s_u$). به طور مختصر توضیح می‌دهیم که چرا نسبت تجزیه $s = \min(s_u, s_f, s_A)$

نمی‌تواند در (۵۵.۴) برای s_u (نسبت بهینه برای تقریب $u \in A^{s_u}$) با هر الگوریتمی که تقریب $\hat{A}$ از A و $\hat{f}$ از f را استفاده می‌کند، اصلاح شود. بر تاثیر ضریب انتشار A تاکید می‌کنیم و فرض می‌کنیم که f به طور دقیق با $\mathcal{T}_0$ ، مثلث بندی اولیه $\mathcal{T}_0$ ، تصرف می‌شود. چون بدین ترتیب بحث ساده برای تقریب f داریم.

تقریب u با $\hat{U}$ ، جواب گالرکین با ضریب انتشار $\hat{A}$ نمی‌تواند از تقریب A با $\hat{A}$ بهتر باشد. در واقع، دو ثابت c و C وجود دارند به گونه ای که برای هر $\delta > 0$ داریم:

$$c\delta \leq \sup_{A_1, A_2 \in B(A, \delta)} \|u_{A_1} - u_{A_2}\|_{H^1(\Omega)} \leq C\delta, \quad (62.4)$$

در (۶۲.۴)، $i = 1, 2$ و $u_{A_i} \in H^1_0(\Omega)$ جواب‌های ضعیف از $-\text{div}(A_i \nabla u_{A_i}) = f$ می‌باشد. برای $\delta > 0$ داریم:

$$B(A, \delta) := \{d \times d \text{ ماتریس معین مثبت } B \mid \|A - B\|_{L_q(\Omega)} \leq \delta\}.$$

نامساوی راست در (۶۲.۴)، پیامد مستقیم از قضیه اختلال است، نامساوی چپ با انتخاب خاص $A_1, A_2 \in B(A, \delta)$ بدست می‌آید. در واقع این انتخاب نشان می‌دهد که $A_1, A_2 \in B(A, \delta)$ و $u_{A_2} = (1 + \frac{\delta}{\|A\|_{L_q(\Omega)}})u_{A_1}$ هستند. بنابراین داریم:

$$\begin{aligned} \|u_{A_1} - u_{A_2}\|_{H^1_0(\Omega)} &= \|u_{A_1} - (1 + \frac{\delta}{\|A\|_{L_q(\Omega)}})u_{A_1}\|_{H^1_0(\Omega)} \\ &= \|u_{A_1} (1 - 1 - \frac{\delta}{\|A\|_{L_q(\Omega)}})\|_{H^1_0(\Omega)} \\ &= \|u_{A_1}\|_{H^1_0(\Omega)} \frac{\delta}{\|A\|_{L_q(\Omega)}} \quad (۶۳.۴) \\ &\geq \frac{\delta}{M^{\frac{1}{r}} \|A\|_{L_q(\Omega)}} \|A^{\frac{1}{r}} \nabla u_{A_1}\|_{L_r(\Omega)} \\ &\geq \frac{r^{\frac{1}{r}} \|f\|_{H^{-1}(\Omega)}}{M^{\frac{1}{r}} \|A\|_{L_q(\Omega)}} \delta \end{aligned}$$

که $0 < r < M < \infty$ کران‌های بالا و پایین برای مقادیر ویژه A هستند.

۵.۴ الگوریتم‌های RHS و COEFF

بهینگی الگوریتم DISC را در قضیه ۱.۴.۴ با این شرط که الگوریتم‌های RHS و COEFF خود بهینه باشند، ثابت کردیم. در این بخش ساختار الگوریتم‌های بهینه برای RHS و COEFF را بررسی می‌کنیم. همانطور که گفته بودیم الگوریتم RHS، تقریبی از f در H^{-1} و الگوریتم COEFF تقریبی از A در L_q ایجاد می‌کند.

ساخت این چنین الگوریتم‌هایی بنا به دو سطح انجام می‌شود. سطح اولی را با فرض اینکه دانش کاملی از f یا A داریم و هرچه نیاز داریم بدون هزینه محاسبه می‌شود، سطح تئوری می‌نامیم. برای مثال، این سطح زمانی بکار می‌رود که f و A تابع‌های هموار تکه‌ای باشند که روی افراز شناخته شده ثابتی قرار دارند (این افراز می‌تواند با افراز اولیه مرتبط نباشد).

سطح دوم را سطح عملی می‌نامیم. در این سطح بیشتر کاربردهای AFEM شناخته شده هستند. به طور دقیق f یا A را نمی‌شناسیم، اما می‌توانیم آن را محاسبه کنیم. اصولاً ساخت الگوریتم‌های تئوری بسیار ساده است، بنابراین ابتدا به این موضوع می‌پردازیم.

۱.۵.۴ الگوریتم‌های بهینه برای تقریب تطبیقی تابع

مطالعه و ساخت الگوریتم‌هایی مثل RHS، موضوعی است که نه تنها برای روش‌های عناصر متناهی تطبیقی بلکه برای نظریه تقریب نیز به کار می‌رود.

این الگوریتم‌ها در همه روش‌های عناصر متناهی تطبیقی مورد نیاز هستند. دو الگوریتم تطبیقی برای توابع تقریب در بخش چهارم معرفی شدند. بهینه بودن آنها را در چندین مرحله ثابت کردیم. این الگوریتم‌ها بر اساس تخمین‌گرهای خطای محلی ساخته شده‌اند. با داشتن تابع $g \in L_q(\Omega)$ ، خطای L_q را در عنصر چند ضلعی T به صورت زیر نشان می‌دهیم:

$$E(T) := E(g, T)_{L_q(T)} := \inf_{P \in P_m(T)} \|g - P\|_{L_q(T)}, \quad (۶۴.۴)$$

فضای چندجمله‌ای‌ها از درجه کمتر مساوی m روی T است. با داشتن T ، بهترین تقریب برای g با چندجمله‌ای تکه‌ای از درجه کوچکتر مساوی m بدست می‌آید. برای این کار باید بهترین تقریب چندجمله‌ای P_T به g را روی T برای هر $T \in \mathcal{T}$ بدست آورد. پس داریم:

$$S_{\mathcal{T}} := \sum_{T \in \mathcal{T}} P_T \chi_T, \quad (۶۵.۴)$$

χ_T تابع مشخصه T است. خطای کلی آن برابر است با:

$$\varepsilon(g, \mathcal{T}) := \|g - S_{\mathcal{T}}\|_{L_q(\Omega)} = \left(\sum_{T \in \mathcal{T}} E(g, T)_{L_q(T)}^q \right)^{\frac{1}{q}}. \quad (۶۶.۴)$$

الگوریتمی که مشاهده می‌کنید، افرازا را با بررسی خطای محلی $E(g, T)_{L_q(\Omega)}$ برای T در همان افراز می‌سازد. نتیجه‌ای که می‌گیریم این است که، $0 < q < \infty$ است. این دو الگوریتم از $\mathcal{T}$ آغاز شده و افرازهای $\mathcal{T}_n$ ، $n = 1, 2, \dots$ را می‌سازند به گونه‌ای که، $\mathcal{T}_n \in \mathcal{F}_{c_n}$ است.

$$\varepsilon(g, \mathcal{T}_n) \leq C \sigma_n(g)_{L_q(\Omega)}, \quad n = 1, 2, \dots \quad (۶۷.۴)$$

که $c > 1$ و $C > 1$ ثابت‌های پایدار هستند. با مرور (۴۵.۴) متوجه می‌شویم که این الگوریتم‌ها هر دو برای تقریب L_q با $1 \leq q < \infty$ برای همه مقادیر $S > 0$ بهینه هستند. در حالیکه الگوریتم‌های بالا برای تقریب L_q بهینه هستند اما بیشتر با راه‌حل‌های ساده‌تر نشانه‌گذاری جایگزین می‌شوند، و فقط عناصری با خطای محلی بیشتر را تظریف می‌کنند.

یکی از این راه‌حل‌ها به نام الگوریتم حریص^{۱۲} را بررسی می‌کنیم. با داشتن هر $\mathcal{T}$ تظریف شده از افراز اولیه $\mathcal{T}$ ، فرایند $\mathcal{T}' = \mathcal{T}'(\varepsilon) = \text{GREEDY}(\mathcal{T}, g, \varepsilon)$ را از $\mathcal{T}$ می‌سازد به گونه‌ای که $\varepsilon(g, \mathcal{T}') \leq \varepsilon$ باشد. برای توصیف این الگوریتم ابتدا قانون دو بخشی از بخش سوم را یادآور می‌شویم. با توجه به این قانون جنگل $\mathcal{F}$ از ریشه $\mathcal{T}$ به وجود می‌آید. الگوریتم GREEDY را به صورت

^{۱۲}GREEDY

زیر نشان می‌دهیم:

$$\begin{aligned}
 \mathcal{T}(\varepsilon) &= \text{GREEDY}(\mathcal{T}, g, \varepsilon) \\
 \mathcal{T}' &= \mathcal{T}; \\
 &\text{while } \varepsilon(g, \mathcal{T}') > \varepsilon, \\
 T &:= \arg \max\{E(g, T) : T \in \mathcal{T}'\}; \\
 \mathcal{T}' &:= \text{REFINE}(\mathcal{T}', T); \\
 &\text{end while} \\
 \mathcal{T}(\varepsilon) &= \text{CONF}(\mathcal{T}')
 \end{aligned}$$

روند $\mathcal{T}' = \text{REFINE}(\mathcal{T}, T)$ دو فرزندش را جایگزین T می‌کند. T به گونه‌ای انتخاب می‌شود تا منجر به شکستن حلقه شود. مشاهده می‌کنیم که الگوریتم GREEDY، عنصر $T \in \mathcal{T}$ با بزرگترین خطای $E(g, T)_{L_q(\Omega)}$ را انتخاب کرده و دو فرزندش را جایگزین T می‌کند تا نظریات بعدی $\mathcal{T}'$ را انجام دهد. این کار تا زمانی که خطای $\varepsilon(g, \mathcal{T}')$ به پایین‌تر از خطای تعیین شده یعنی ε برسد، ادامه دارد. تا زمانی که این چرخه ادامه دارد، نظریات‌های دیگری با CONF روی $\mathcal{T}'$ صورت می‌گیرد تا کوچکترین افزایش سازگار $\mathcal{T}(\varepsilon)$ که شامل $\mathcal{T}'$ باشد بدست آید. ویژگی مهم خطای محلی $E(T)$ یکنواخت بودن آن است.

$$E(T_1)^q + E(T_2)^q \leq E(T)^q \quad \forall T \in \mathcal{T}, \quad (۶۸.۴)$$

که T_1 و T_2 دو فرزند T هستند. این مساله به خاصیت یکنواختی کلی منجر می‌شود، یعنی

$$\varepsilon(g, \mathcal{T}') \leq \varepsilon(g, \mathcal{T}), \quad (۶۹.۴)$$

رابطه (۶۹.۴) برای همه نظریات $\mathcal{T}'$ از $\mathcal{T}$ چه منطبق باشد و چه نباشد برقرار است.

۲.۵.۴ الگوریتم‌های بهینه برای COEFF

با داشتن عدد صحیح m ، فضای $\sum_n^m := \sum_n^m$ از ماتریس توابع چندجمله‌ای تکه‌ای از درجه کوچکتر مساوی m و خطای $\sigma_n(A)_{L_q(\Omega)}$ و رده‌های تقریب $\mathcal{M}^s(L_q(\Omega))$ که در بخش سوم معرفی شده‌اند را یادآوری می‌کنیم. فرض کنید که $A \in L_\infty(\Omega)$ ماتریس معین مثبتی است که مقادیر ویژه آن در رابطه زیر صدق می‌کند.

$$r \leq \lambda_{\min}(A) \leq \lambda_{\max}(A) \leq M, \quad (۷۰.۴)$$

رابطه (۷۰.۴) معادل است با:

$$r \leq \sum_{i,j=1}^d a_{ij}(x) Z_i Z_j = Z^t A(x) Z \leq M, \quad |Z| = 1, \quad (۷۱.۴)$$

ساخت الگوریتم COEFF برای تقریب A با عناصر $B \in \sum_n^m$ دارای دو مولفه است. اولی یافتن تقریب‌های $B \in L_q(\Omega)$ برای $q < \infty$ است. دومی اطمینان از معین مثبت بودن B است.

اجرای معین مثبت بودن

تقریب‌های مرتبه بالای $B \in \sum_n^m$ از A ، یعنی $m > 0$ ، ممکن است معین مثبت نباشد. حال نشان می‌دهیم که چطور B را برای معین مثبت بودن تنظیم کنیم، بدون اینکه تقریبش به A کاسته شود.

گزاره ۱.۵.۴. (اجرای معین مثبت بودن). فرض کنید A تابع ماتریسی معین مثبت متقارن در $L_q(\Omega)$ باشد، که مقادیر ویژه آن در $[r, M]$ با $M \geq 1$ قرار دارد. فرض کنید $T \subset \Omega$ عنصر چندضلعی باشد که از دو بخشی کردن جدیدترین راس $\mathcal{T}$ بدست آمده است. اگر تابع ماتریس مقداری B را داشته باشیم که چندجمله‌ای از درجه کوچکتر مساوی m روی $\mathcal{T}$ است، آنگاه خواهیم داشت:

$$\|A - B\|_{L_q(T)} \leq \varepsilon,$$

آنگاه ماتریس معین مثبت $\tilde{B}$ وجود دارد که ورودی‌های آن چندجمله‌ای از درجه کوچکتر مساوی m می‌باشد. به گونه‌ای که مقادیر ویژه $\tilde{B}$ در رابطه‌ی زیر صدق می‌کند:

$$\frac{r}{4} \leq \lambda_{\min}(\tilde{B}) \leq \Lambda_{\max}(\tilde{B}) \leq CM, \quad (72.4)$$

و

$$\|A - \tilde{B}\|_{L_q(T)} \leq \tilde{C}\varepsilon,$$

که C وابسته به d و m و افزایش اولیه $\mathcal{T}$ است، و $\tilde{C}$ وابسته به $\frac{M}{r}$ است.

برهان. برهان را با عکس نامساوی شروع می‌کنیم، ثابت C_1 وابسته به m, d و $\mathcal{T}$ وجود دارد، به گونه‌ای که برای هر چند جمله‌ای P از درجه کوچکتر مساوی m و هر چندضلعی T که از قانون دو بخشی جدیدترین راس ناشی شده است داریم:

$$\|\nabla P\|_{L_\infty(T)} \leq C_1 \|P\|_{L_\infty(T)} |T|^{-\frac{1}{d}} \quad (73.4)$$

اکنون با داشتن A و تقریب B ، M_0 را به صورت زیر فرض می‌کنیم:

$$M_0 := \sup_{x \in T} \sup_{|y|=1} y^t B(x) y,$$

ابتدا نشان می‌دهیم که این ادعا برای $M_0 > CM$ با ثابت C صحیح می‌باشد. در اینجا مقدار $1 \leq C$ و محدودیت‌های C را حین پیشرفت کار استخراج می‌کنیم.

اگر $M_0 > CM$ باشد، سپس y با $|y| = 1$ و $x_0 \in T$ وجود دارد به گونه‌ای که $y^t B(x_0) y = M_0 > CM$ باشد. این y را ثابت نگه داشته و تابع $a(x) := y^t A(x) y$ و چندجمله‌ای $P(x) := y^t B(x) y$ از درجه m را در نظر می‌گیریم. توجه کنید که $\|P\|_{L_\infty(T)} = M_0$ است و داریم:

$$\|a - P\|_{L_q(T)} \leq \|A - B\|_{L_q(T)} \leq \varepsilon \quad (74.4)$$

با در نظر گرفتن (۷۳.۴) خواهیم داشت:

$$|P(x) - P(x_0)| \leq C_1 M_0 |T|^{-\frac{1}{d}} |x - x_0| \leq \frac{M_0}{4}, \quad |x - x_0| \leq \frac{|T|^{\frac{1}{d}}}{(2C_1)}.$$

فرض کنید T_0 مجموعه‌ای از $x \in T$ ، به گونه‌ای که $|x - x_0| \leq \frac{|T|^{\frac{1}{d}}}{(2C_1)}$ باشد. آنگاه $P(x) \geq \frac{M_0}{4}$ روی T_0 برقرار است. از آنجاییکه $0 < a(x) < M < \frac{M_0}{C}$ است. بنابراین با این شرط که $C \geq 4$ باشد،

دارد، رابطه زیر را خواهیم داشت:

$$M_*(c|T|)^{\frac{1}{q}} \leq \|a - P\|_{L_q(T)} \leq 4\varepsilon. \quad (۷۵.۴)$$

اگر $\tilde{B} := rI$ را روی T ، با I همانی تعریف کنیم، سپس با استفاده از (۷۵.۴)، با این شرط که C به اندازه کافی بزرگ انتخاب شود به گونه‌ای که $\frac{1}{\lambda} \leq C^{-1}c^{\frac{-1}{q}} \leq \frac{1}{\lambda}$ باشد. رابطه‌ی زیر را بدست می‌آوریم:

$$\|A - \tilde{B}\|_{L_q(T)} \leq (M + r)|T|^{\frac{1}{q}}$$

با توجه به روابط $0 < r \leq M < \infty$ و $M < \frac{M_*}{C}$ داریم:

$$\|A - \tilde{B}\|_{L_q(T)} \leq 2M|T|^{\frac{1}{q}} \leq (2\frac{M_*}{C})|T|^{\frac{1}{q}} \leq (\frac{M_*}{4})(c|T|)^{\frac{1}{q}} \leq \varepsilon$$

قضیه را با فرض $M_* > CM$ حل کردیم.

اکنون می‌خواهیم حالت $M_* \leq CM$ را بررسی کنیم. μ را به صورت زیر در نظر بگیرید:

$$\mu := \inf_{x \in T} \inf_{|y|=1} y^t B(x) y \quad (۷۶.۴)$$

اگر $\mu \geq \frac{r}{4}$ باشد، نیازی به اثبات نداریم. بنابراین فرض می‌کنیم که $\mu < \frac{r}{4}$ و $|y_*| = 1$ باشد، به گونه‌ای که $P(x) := y_*^t B(x) y_*$ است، با این فرض که مقدار μ در x_* است. با استفاده از (۷۳.۴) خواهیم داشت:

$$|p(x) - p(x_*)| \leq C_* M_* |T|^{\frac{-1}{d}} |x - x_*| \leq C_* CM |T|^{\frac{-1}{d}} |x - x_*|, \quad x \in T, \quad (۷۷.۴)$$

فرض کنید T_* مجموعه‌ای از $x \in T$ باشد، به گونه‌ای که $|x - x_*| \leq r|T|^{\frac{1}{d}} / (4C_* CM)$ است. توجه کنید که $|T_*| \geq cr^d C_*^{-d} C^{-d} M^{-d} |T|$ برای برخی ثابت c و وابسته به d و T_* است. همچنین $|P(x) - P(x_*)| \leq \frac{r}{4}$ است.

نتیجه می‌گیریم که $P(x) \leq \mu + \frac{r}{4}$ روی زیر مجموعه T_* قرار دارد. بنابراین برای $a(x) := y_*^t A(x) y_*$ با $x \in T_*$ خواهیم داشت:

$$a(x) - P(x) \geq \frac{3r}{4} - \mu$$

نتیجه می‌گیریم که:

$$|T_*|^{\frac{1}{q}} (\frac{3r}{4} - \mu) \leq \|a - P\|_{L_q(T)} \leq \varepsilon \quad (۷۸.۴)$$

حال $\tilde{B} = B + (\frac{3r}{4} - \mu)I$ را تعریف می‌کنیم. به گونه‌ای که

$$\bar{y}^t \tilde{B}(x) y \geq \frac{3r}{4}, \quad |y| = 1, \quad x \in T,$$

$$\|A - \tilde{B}\|_{L_q(T)} \leq \|A - B\|_{L_q(T)} + |T|^{\frac{1}{q}} (\frac{3r}{4} - \mu)$$

$$\leq \varepsilon + |T_*|^{\frac{1}{q}} (\frac{3r}{4} - \mu) (|T| |T_*|^{-1})^{\frac{1}{q}}$$

$$\leq \tilde{C}\varepsilon,$$

□

از (۷۸.۴) و مقدار اندازه‌گیری شده T_* استفاده کردیم.

تذکر ۲.۵.۴. (شکل $\tilde{B}$). در گزاره قبلی ماتریس $\tilde{B}$ به یکی از سه شکل زیر است:

$$\tilde{B} = B \quad (\text{i})$$

$$\tilde{B} = rI \quad (\text{ii})$$

$$\tilde{B} = B + \alpha I \quad (\text{iii}) \quad \alpha > 0 \text{ برای مقادیر}$$

برای تغییر این دستورالعمل به روند عددی، ابتدا باید تقریب B از A را بدست آوریم. اکنون قضیه زیر را ثابت می‌کنیم که نشان می‌دهد اگر تقریب B به A معین مثبت یکنواخت باشد، از صحت کلی آن کاسته نمی‌شود.

قضیه ۳.۵.۴. (اجرای معین مثبت بودن کلی). فرض کنید A تابع ماتریسی معین مثبت در $L_q(\Omega)$ باشد که مقادیر ویژه آن همگی در $[r, M]$ ، برای هر $x \in \Omega$ با $1 \leq M$ است. اگر $B \in \sum_n^m$ باشد، خواهیم داشت:

$$\|A - B\|_{L_q(\Omega)} \leq \varepsilon,$$

آنگاه $\tilde{B} \in \sum_n^m$ ای وجود دارد که مقادیر ویژه آن در $[\frac{r}{C}, CM]$ برای همه $x \in \Omega$ قرار دارند.

$$\|A - \tilde{B}\|_{L_q(\Omega)} \leq \tilde{C}\varepsilon, \quad (79.4)$$

که در آن C و $\tilde{C}$ مقادیر ثابت گزاره ۱.۵.۴ هستند.

برهان. فرض کنید $B = \sum_{T \in \mathcal{T}} B_T \chi_T$ باشد، هر کدام از ماتریس‌های B_T ورودی‌های چندجمله‌ای از درجه کمتر مساوی m داشته و χ_T تابع مشخصه T است. اگر $\tilde{B}_T$ ماتریس گزاره ۱.۵.۴، برای A و $\tilde{B}_T$ روی T به کار گرفته شده باشد، آنگاه $\tilde{B}(x) := \sum_{T \in \mathcal{T}} \tilde{B}_T(x) \chi_T(x)$ دارای مقادیر ویژه در $[\frac{r}{C}, CM]$ برای $x \in \Omega$ است. تخمین خطای (79.4) از $\|A - \tilde{B}\|_{L_q(\Omega)} \leq \tilde{C}\|A - B\|_{L_q(T)}$ برای هر $T \in \mathcal{T}$ نتیجه می‌شود. $\square$

اگر $\tilde{\sigma}_n(A)_{L_q(\Omega)}$ را بصورت σ_n تعریف کنیم، به غیر از اینکه ماتریس‌های تقریب معین مثبت هستند، سپس قضیه ۳.۵.۴ به معنای $\tilde{\sigma}_n(A)_{L_q(\Omega)} \leq \tilde{C}\sigma_n(A)_{L_q(\Omega)}$ برای همه $n \geq 1$ است.

الگوریتم‌ها برای تقریب A .

با توجه به قضیه ۳.۵.۴، اکنون بدون در نظر گرفتن معین مثبت بودن A ، روی تقریب A در $L_q(\Omega)$ متمرکز می‌شویم. در ابتدا مشاهده می‌کنیم که تقریب A با عناصری از $\sum_n^m$ ، بزرگتر از تقریب ورودی‌های آن است، برای هر ماتریس $d \times d$ ، $B = (b_{i,j})$ داریم:

$$\max_{i,j} \|b_{i,j}\|_{L_q(\Omega)} \leq \|B\|_{L_q(\Omega)} \leq \sum_{i,j} \|b_{i,j}\|_{L_q(\Omega)} \leq d^2 \max_{i,j} \|b_{i,j}\|_{L_q(\Omega)}, \quad (80.4)$$

نتیجه می‌گیریم که تقریب A در $L_q(\Omega)$ با عناصری از $\sum_n^m$ هم ارز با تقریب ورودی‌هایش یعنی $a_{i,j}$ در $L_q(\Omega)$ با چندجمله‌ای‌های تکه‌ای از درجه کمتر مساوی m است. بنابراین توجه کنید که

$A \in \mathcal{M}^s(\mathcal{T}_\circ, L_q(\Omega))$ هم ارز با هر ورودی $a_{i,j}$ موجود در $\mathcal{B}^s(\mathcal{T}_\circ, L_q(\Omega))$ است. تحلیل فوق به این معنی است که ساخت الگوریتم‌های بهینه برای COEFF از ساخت الگوریتم‌های بهینه برای توابعی در L_q نتیجه می‌شود. اگر بتوانیم خطای محلی $E(a_{i,j}, T)_{L_q}$ را برای هر ضریب $a_{i,j}$ و هر عنصر T محاسبه کنیم، آنگاه می‌توانیم الگوریتم بهینه COEFF را برای $\mathcal{M}^s(\mathcal{T}_\circ, L_q(\Omega))$ بسازیم. بعلاوه، الگوریتم COEFF تضمین می‌کند که تقریب‌ها معین مثبت بوده و دارای شرایط (۳۶.۴) می‌باشند.

فصل ۵

آزمایشات عددی

در این فصل با پیاده‌سازی روش توصیف شده در فصل قبل در قالب دو مثال به بررسی قابلیت‌ها و محدودیت‌های آن می‌پردازیم. در این مثال‌ها مش اولیه $\mathcal{T}_0$ از Ω ، به صورت اجتماعی از مثلث‌های مساوی می‌باشد و الگوریتم مورد استفاده برای تظریف آن همان الگوریتم دو بخشی جدیدترین راس است.

قبل از شروع آزمایشات باید در خصوص انتخاب توان لِبِگ، $q = \frac{2p}{p-2}$ که برای تقریب A لازم است، تصمیم‌گیری کنیم. از آنجا که بطور کلی p را نمی‌شناسیم، بنابراین پرسش اصلی نحوه تعیین q در عمل است. از این حقیقت که A و تقریب چندجمله‌ای تکه‌ای آن یعنی $\hat{A}$ بطور یکنواخت در $L_\infty(\Omega)$ کران دار هستند (برای مثال با ثابت C_A) استفاده کرده و از $\hat{A}$ به عنوان بهترین تقریب $L_2(\Omega)$ که عنصر به عنصر محاسبه شده بهره می‌بریم و از تخمین درونیابی

$$\|A - \hat{A}\|_{L_q(\Omega)} \leq (2C_A)^{\frac{1}{p}} \|A - \hat{A}\|_{L_2(\Omega)}^{\frac{1}{q}} \quad 2 \leq q \leq \infty \quad (1.5)$$

استفاده می‌کنیم. بنابراین تخمین اختلال (۱.۲.۴) به صورت زیر تقلیل می‌یابد:

$$\|u - \hat{u}\|_{H^1(\Omega)} \leq \hat{r}^{-1} \|f - \hat{f}\|_{H^{-1}(\Omega)} + \frac{(2C_A)^{\frac{1}{p}}}{\hat{r}} \|\nabla u\|_{L_p(\Omega)} \|A - \hat{A}\|_{L_2(\Omega)}^{\frac{1}{q}}.$$

بدین ترتیب یک انتخاب کلی از $\hat{A}$ ، بدون در نظر گرفتن ارزش‌های p و q خواهیم داشت. اما توجه داشته باشید که نیاز به شرط کافی $A \in \mathcal{M}^{\frac{q^s A}{2}}(L_2(\Omega))$ و همچنین $2 \leq q < \infty$ داریم؛ بنابراین این انتخاب ممکن است همواره نرخ همگرایی $\|A - \hat{A}\|_{L_q(\Omega)}$ را حفظ نکند. در عمل یک استثنای مهم زمانی روی می‌دهد که A ثابت تکه‌ای با تعداد متناهی از تکه‌ها باشد و پرش‌های آن روی یک منحنی لپ شیتزی باشد، بنابراین

$$\begin{aligned} \|A - \hat{A}\|_{L_q(\Omega)} &\approx \left| \{x \in \Omega : A(x) \neq \hat{A}(x)\} \right|^{\frac{1}{q}} \\ &\approx \|A - \hat{A}\|_{L_2(\Omega)}^{\frac{1}{q}} \end{aligned}$$

بدین ترتیب در نرخ همگرایی (تقارب) A خللی وارد نمی‌گردد. در آزمایشات عددی بعدی، A ثابت تکه‌ای است. به همین دلیل می‌توانیم از تقریب ثابت تکه‌ای برای ماتریس انتشار A و f سمت راست

بر اساس مقدار میانگین آنها یعنی $\bar{A}$ و $\bar{f}$ استفاده کنیم. خاطر نشان می‌کنیم که $\bar{A}$ دارای کران‌های طیفی یکسان هستند.

۱.۵ نتایج عددی

۱.۱.۵ مثال ۱: دامنه L شکل.

ابتدا DISC را با بهترین انتخاب ممکن $q = 2$ بررسی می‌کنیم. برای این منظور دامنه L شکل $\Omega = [-5, 5] \times [-5, 5] \setminus [0, 5] \times [0, -5]$ را در نظر می‌گیریم. از مختصات قطبی (ρ, δ) به جای مختصات دکارتی به مرکز $(0, 0)$ استفاده می‌کنیم. ماتریس انتشار را $A = aI$ در نظر می‌گیریم که I برابر با ماتریس واحد است،

$$a(\rho) = \begin{cases} 1 & \text{if } \rho \leq \rho_0 \\ \mu & \text{otherwise,} \end{cases}$$

که در آن $\rho_0 = 2\sqrt{2}$ و $\mu = 5$ است. با تعریف $v(\rho, \delta) := \rho^{\frac{1}{3}} \sin(\frac{\sqrt{2}\delta}{3})$ آن را به عنوان جواب استاندارد روی دامنه L شکل در نظر بگیرید. برای نشان دادن عملکرد DISC جواب استاندارد را به ترتیب زیر دستکاری می‌کنیم:

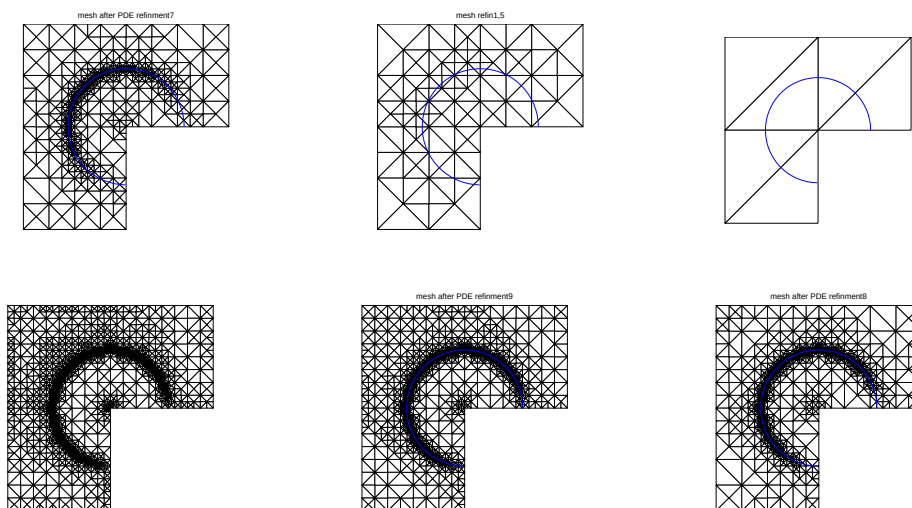
$$u(\rho, \delta) = \begin{cases} v(\rho, \delta) & \text{if } \rho \leq \rho_0 \\ v(\rho_0, \delta) + \frac{\sqrt{2}}{3\mu} \rho_0^{\frac{1}{3}} \sin(\frac{\sqrt{2}\delta}{3})(\rho - \rho_0) & \text{otherwise,} \end{cases}$$

دقت کنید که $f := -\operatorname{div}(A\nabla u) \in L^2(\Omega)$ و کران‌های طیفی $\bar{A}$ و A بصورت $M = 5$ و $r = 1$ است.

تاکید می‌کنیم که ناپیوستگی A هرگز با افرازاها هماهنگ نمی‌باشد. شکل (۱.۵) توالی افرازهای شکل گرفته بر اساس الگوریتم کامل DISC را نشان می‌دهد.

پارامترهای انتخابی $\omega = 0.8$ ، $\beta = 0.7$ و $\varepsilon_0 = 2$ هستند. حلقه استاندارد AFEM در PDE تحت تاثیر تخمین‌گر باقیمانده خطا و استراتژی نشانه گذاری با پارامتر $\theta = 0.3$ به پیش می‌رود که نسبتاً انتخابی محافظه کارانه است.

اکنون درباره انتخاب p به طوری که به ازای آن شرط p معتبر باشد بحث می‌کنیم. فرض کنید $\Omega = \Omega_1 \cup \Omega_2$ ، که در آن $\Omega_1 = \{(\rho, \delta) \in \Omega : \rho \leq \frac{\rho_0}{2}\}$ و $\Omega_2 = \Omega \setminus \Omega_1$ باشد. $u \in W_{L^p(\Omega_1)}^1$ برای هر $p < 6$ در Ω_1 صادق است و u لیپشیتزی در Ω_2 است و این در حالیکه شرط p برای $p < 6$ در Ω_1 معتبر است، یعنی هر $q = \frac{2p}{(p-2)} > 3$ و $p = \infty$ در Ω_2 ، یعنی هر $q \geq 2$. از آنجاییکه ضریب انتشار در Ω_1 ثابت است خطای تقریب A صفر بوده و فقط باید پرش از منحنی $\{\rho = \rho_0\}$ بر روی Ω_2 را در نظر بگیریم.



شکل ۱۰۵: مثال یک (دامنه L شکل). مش اولیه (ردیف بالا سمت راست) برای مسئله دامنه L -شکل با ۶ المان که از مثلث‌های یک اندازه بدون گره‌های آویز ساخته شده است. دایره ناقص آبی رنگ نشان دهنده نقاط ناپیوستگی u است که انتظار داریم تظریف مش در اطراف آن بیشتر باشد. توالی افرازاها (خلاف عقربه‌های ساعت) بر اساس الگوریتم DISC با شاخص $q = 2$ برای A و همچنین مقدار پارامترها $\beta = 0.7$ و $\omega = 0.8$ و $\varepsilon_0 = 2$ انتخاب می‌شود. الگوریتم DISC فقط در مراحل اولیه برای پیدا کردن پرش‌ها در ماتریس انتشار اجرا می‌شود. تظریف‌های حول مبدا به دلیل رفتار منفرد جواب است که در مرحله اصلی فرایند تطبیقی به وجود می‌آید.

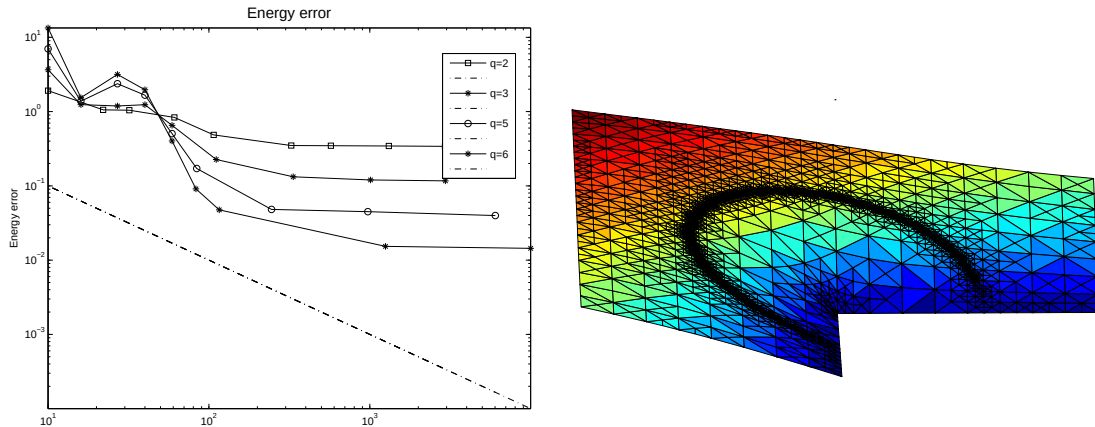
چنین پرشی هرگز توسط افرازاها تطبیق داده نمی‌شود، بنابراین A به صورت تکه‌ای هموار روی افراز $\mathcal{T}_0$ در نظر گرفته نمی‌شود. همین مساله باعث می‌شود که نتوانیم از AFEM استاندارد استفاده کنیم. به آسانی می‌توان بررسی نمود که برای هر $1 \leq q < \infty$ ماتریس A در $\mathcal{M}_q^1(\mathcal{T}_0, L_q(\Omega))$ قرار دارد. از آنجا که عملکرد DISC برای q بزرگتر، کاهش می‌یابد، باید کوچک ترین مقدار $q = \frac{2p}{(p-2)}$ که با $u \in w^1(L_\infty(\Omega_2))$ سازگاری بیشتری داشته باشند را انتخاب کنیم. به عبارتی دیگر برای $p = \infty$ باید $q = 2$ را قرار دهیم. در اینجا f سمت راست در $B^1_2(\mathcal{T}_0, H^{-1}(\Omega)) \subset B^1_2(\mathcal{T}_0, L_2(\Omega))$ صدق می‌کند در حالیکه جواب $u \in A^1_2(\mathcal{T}_0, H^1_0(\Omega))$ است. زیرا $u \in A^1_2(\mathcal{T}_0, H^1(\Omega_i))$ و $i = 1, 2$ ، پرش‌های گرادیان u روی منحنی لیپ شیتز اتفاق می‌افتد.

برای تست کردن توضیحات بالا، چهار مقدار متفاوت p و در نتیجه q مرتبط به آنها را در آزمایشات عددی خود در نظر گرفتیم. برای هر یک از این انتخاب‌های متفاوت، شکل (۲.۵) (سمت چپ) نزول خطای انرژی در برابر تعداد درجات آزادی در مقیاس $\log - \log$ را نشان می‌دهد. ترتیب‌های همگرایی آزمایشی عبارتند از،

$$-0.48 \text{ برای } q = 2, -0.35 \text{ برای } q = 3, -0.23 \text{ برای } q = 5, \text{ و } -0.19 \text{ برای } q = 6.$$

این موارد با تقریب پذیری A که در قبل ذکر کردیم مطابقت دارند. این نرخ‌های محاسباتی به مقدارهای قابل انتظار $\frac{1}{q}$ بسیار نزدیک هستند و اهمیت تقریب و ارزیابی A در دامنه‌های فرعی و با کوچک ترین توان لیگ ممکن برای q را نشان می‌دهند. در این مثال انتخاب $q = 2$ نرخ مطلوب همگرایی برای

عناصر دوخطی تکه‌ای می‌باشد. شکل (۲.۵) (سمت راست) جواب گالرکین پس از ده تکرار DISC برای $q = 2$ را نشان می‌دهد.



شکل ۲.۵: مثال یک (دامنه L شکل): (سمت چپ) خطای انرژی در برابر تعداد درجات آزادی برای مقادیر $q = 2, 3, 5, 6$. بالاترین میزان همگرایی برای $q = 2$ بدست می‌آید. (سمت راست) جواب گالرکین به همراه افراز مربوطه پس از ده تکرار الگوریتم DISC با $q = 2$ است. ناپیوستگی A هرگز در افرازاها تطبیق داده نشده است و افراز حول تکینگی‌های A و ∇u نظریف می‌گردد.

۲.۱.۵ مثال ۲: صفحه چکرز

اکنون DISC را برای مثالی بررسی می‌کنیم که در آن $q = 2$ مجاز نیست. در این مثال که ابتدا در [۹] پیشنهاد شده بود، خط ناپیوستگی ماتریس انتشار A با تکینگی جواب u یکسان است. فرض کنید $A = aI$ و $\Omega = (-1, 1)^2$ باشد، که در آن I ماتریس واحد 2×2 است:

$$a(x, y) = \begin{cases} b & (x - \frac{\sqrt{2}}{4})(y - \frac{\sqrt{2}}{4}) \geq 0, \\ 1 & \text{otherwise,} \end{cases}$$

که $b > 0$ معلوم می‌باشد. سمت راست را به صورت $f \equiv 0$ در نظر می‌گیریم، بنابراین با تعیین شرایط مرزی مناسب، جواب u در مختصات (ρ, δ) با مرکز واقع شده در $(\frac{\sqrt{2}}{4}, \frac{\sqrt{2}}{4})$ به صورت زیر است.

$$u(\rho, \delta) = \rho^\alpha \mu(\delta),$$

که در آن $0 < \alpha < 2$ و $\mu(\delta)$ به صورت زیر است.

$$\mu(\delta) = \begin{cases} \cos((\frac{\pi}{4} - \sigma)\alpha) \cos((\delta - \frac{\pi}{4})\alpha) & 0 \leq \delta < \frac{\pi}{4}, \\ \cos(\frac{\pi}{4}\alpha) \cos((\delta - \pi + \sigma)\alpha) & \frac{\pi}{4} \leq \delta < \pi, \\ \cos(\alpha\sigma) \cos((\delta - \frac{5\pi}{4})\alpha) & \pi \leq \delta < \frac{3\pi}{4}, \\ \cos(\frac{\pi}{4}\alpha) \cos(\delta - \frac{3\pi}{4} - \sigma)\alpha & \frac{3\pi}{4} \leq \delta < 2\pi. \end{cases}$$

در رابطه بالا پارامترهای b و a و σ در روابط غیر خطی زیر

$$b = -\tan\left(\left(\frac{\pi}{4} - \sigma\right)\alpha\right) \cot\left(\frac{\pi}{4}\alpha\right),$$

$$\frac{1}{b} = -\tan\left(\frac{\pi}{4}\alpha\right) \cot(\sigma\alpha),$$

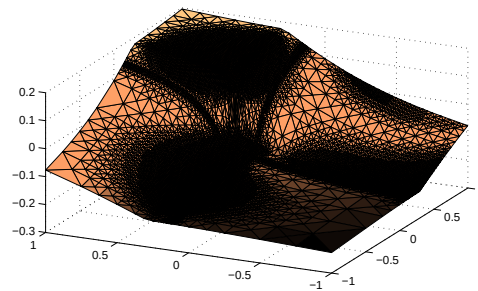
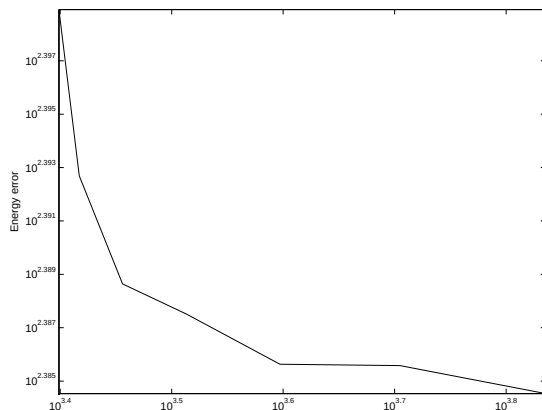
$$b = -\tan(\sigma\alpha) \cot\left(\frac{\pi}{4}\alpha\right),$$

همراه با قیدهای

$$\max(0, \pi(\alpha - 1)) < \frac{\pi}{4}\alpha < \min(\pi\alpha, \pi),$$

$$\max(0, \pi(1 - \alpha)) < -2\alpha\sigma < \min(\pi, \pi(2 - \alpha)).$$

صدق می‌کنند.



شکل ۳.۵: مثال دوم (صفحه چکرز): پارامترها بر اساس جواب $u \in H^{1+s}$ ، $s < 0.25$ انتخاب شده‌اند. (سمت چپ) خطای انرژی در برابر تعداد درجات آزادی. سطح مطلوب همگرایی تقریباً $0.2 -$ پوشش داده شده است. (سمت راست) جواب گالرکین به همراه افرازهای مربوطه پس از شش تکرار الگوریتم DISC. ناپیوستگی A هرگز با افرازها جبران نمی‌شود و تکینگی‌های A و ∇u هر دو با هم باعث تظریف می‌شوند.

تاکید می‌کنیم که برای α ، $s < \alpha$ ، $u \in H^{1+s}(\Omega)$ را داریم. اما با این حال $u \in A^\dagger(H_0^1(\Omega))$ است. اما ناپیوستگی A با تکینگی‌های u مطابقت داشته است. در این صورت خواهیم داشت:

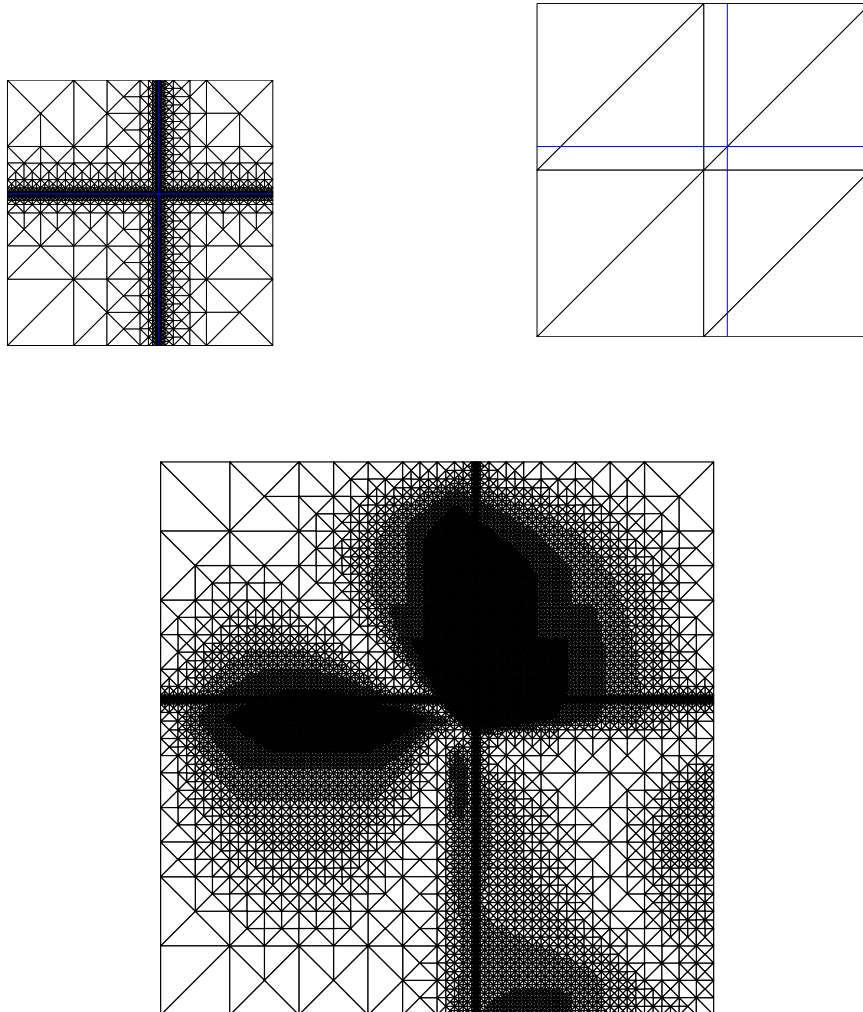
$$p < \frac{2}{(1-\alpha)}, \quad s_A = \frac{1}{q} = \frac{(p-2)}{2p} < \frac{\alpha}{2}.$$

حال الگوریتم DISC را با پارامترهای تقریب به چالش می‌کشیم.

$$\begin{aligned} \alpha &= 0.25, & b &\approx 25.27414236908818, \\ \sigma &\approx -5.49778714378214, \end{aligned} \quad (2.5)$$

که روابط بالا متناظر با $p < \frac{4}{3}$ و $s_A < \frac{1}{8}$ است. از رابطه (۱.۵) استفاده کرده و $\bar{A}$ را به عنوان مقدار میانگین A بصورت عنصر به عنصر در نظر می‌گیریم. مرتبه همگرایی تجربی (EOC) خطای انرژی در برابر تعداد درجات آزادی در شکل ۳.۵ به همراه جواب در مرحله پایانی رسم شده است. مقدار EOC

مجانبی (میانگین ۳ نقطه آخر) برابر با $0.51 -$ است که در حد مطلوب بوده و بسیار بهتر از مقدار مورد انتظار $s_A \approx -0.125$ می باشد. از سوی دیگر، مقدار مجانبی پیشینی EOC (بدون ۳ نقطه آخر) حدود $0.2 -$ است. اکنون شکل ۴.۵ را که نشانگر افزایشهای مثلی در مرحله $k = 0, 5, 19$ می باشد، را نتیجه می گیریم.



شکل ۴.۵: مثال دوم (صفحه چکرز). توالی افزایشها (از راست به چپ) که براساس DISC با $\omega = 0.8$ شکل گرفته است. افزایش اولیه (طرح اول) از مثلثها تشکیل شده است. الگوریتم فقط در مراحل اولیه تعریف می شود تا ناپیوستگی ضریب انتشار (دوم) را جبران کند. سپس تکنیکیهای u وارد عمل می شود و به همراه تکنیکیهای A ، باعث تعریف می گردد (سطح سوم). تقسیم بندی متناظر شامل سه میلیون درجه از آزادی هستند. کوچکترین سلول دارای قطر 2^{-6} می باشد. که نشانگر توانایی زیاد DISC در ساخت شبکه های درجه بندی شده می باشد.

۲.۵ نتیجه گیری

نشان دادیم که محدودیت معین مثبت $\hat{A}$ بر مرتبه تقریب L_q تاثیری ندارد. استخراج الگوریتم‌های قابل اجرای عددی که تضمین کننده معین مثبت بودن، بوده و در شرایط $L_q(\Omega)$ عملکرد بهتری داشته باشند یک مساله فرعی است. هر استفاده از COEFF با خطاهای هدف کوچکتر و $q = \infty$ همگرا نخواهد بود. اصولاً تعریف بر اساس PDE با دقت بهتری انجام می‌گیرد. در نهایت توانستیم PDE استاندارد را بر اساس تقریب سمت راست $\hat{f}$ و ضریب انتشار $\hat{A}$ که در RHS و COEFF ارایه شدند، انجام دهیم.

پیوست کدهای متلب

۱.آ کد متلب مثال ۱

در اینجا کد متلب مثال ۱، قرار داده شده است.

```
1 function mesh=Lshape(mesh_0,q)
2 close all
3 % Parameters
4 MaxIt =10; % MaxIt: maximum iteration number
5 N = zeros(MaxIt,1);
6 energy = zeros(MaxIt,1);
7 theta = 0.3;
8 epsilon=2;
9 omega=0.8;
10 beta=0.7;
11 %-----
12 % PreStep 0: generate initial mesh
13 %-----
14 for iter=1:MaxIt
15     %Start RHS
16     fAprox=RHS(mesh);
17     fexact=ff(mesh.node);
18     max(fAprox-fexact)
19     eta_f = estimate_f(fAprox,mesh,q);
20     E=sum(eta_f)^(1/q);
21     while E>epsilon*omega
22         [mesh,~] = bisection(mesh,eta_f,theta);
23         fAprox=RHS(mesh);
```

```

24     eta_f=estimate_f(fAprox,mesh,q);
25     E=sum(eta_f.^(q)).^(1/q);
26     end
27     %End RHS
28     %Start PDE
29     % Step 1: Solve
30     mesh.f=RHS(mesh);
31     [mesh.solu, energy(iter)] = PDE(mesh,@g_D);
32     % Step 2: Estimate
33     eta_u = estimate(mesh).^2;
34     % record number of elements for comparison
35     N(iter) = size(mesh.node,1);
36     % Step 3: Refine
37     [mesh,~] = bisection(mesh,eta_u,theta);
38     epsilon=beta*epsilon;
39     end

```

```

1  function [mesh,eta] = bisection(mesh,eta,theta)
2  %-----
3  % Construct data structure
4  %-----
5  edge = [mesh.elem(:,[1,2]); mesh.elem(:,[1,3]); mesh.elem(:,[2,3])];
6  edge = unique(sort(edge,2),'rows');
7
8  N = size(mesh.node,1);
9  NT = size(mesh.elem,1);
10 NE = size(edge,1);
11
12 dualEdge = ...
13     sparse(mesh.elem(:,[1,2,3]),mesh.elem(:,[2,3,1]),[1:NT,1:NT,1:NT]);
14 d2p = sparse(edge(:,[1,2]),edge(:,[2,1]),[1:NE,1:NE]);
15 %-----
16 % Memory management for node array
17 %-----
18 recycle = find(mesh.type==0);
19 last = length(recycle);
20 %-----
21 % Mark triangles according to the error indicator
22 %-----
23 total = sum(eta);
24 current = 0;

```

```

25 [temp,ix] = sort(-eta); % sort in descent order
26 marker = zeros(NE,1);
27
28 % initialize for possible new nodes
29 mesh.node = [mesh.node; zeros(NE,2)];
30 mesh.type = [mesh.type; uint8(5*ones(NE,1))];
31 mesh.solu = [mesh.solu; zeros(NE,1)];
32 mesh.f = [mesh.f; zeros(NE,1)];
33 for t = 1:NT
34
35     if (current > theta*total), break; end % est on marked elem ...
        big enough
36
37     index = 1;
38     ct = ix(t);
39
40     while (index==1)
41
42         base = d2p(mesh.elem(ct,2),mesh.elem(ct,3));
43         if (marker(base)>0) % base is already marked
44             index = 0;
45         else
46             current = current + eta(ct);
47             if (last==0)
48                 newNode = N + 1;
49                 N = N+1;
50             end
51             if (last>0)
52                 newNode = recycle(last);
53                 last = last -1;
54             end
55             marker( d2p(mesh.elem(ct,2),mesh.elem(ct,3)) ) = ...
                newNode;
56
57             % A new node is added to the mesh. Numerical ...
                solution at this
58             % new added node is approximated by linear ...
                interpolation. Type
59             % of this node is 2 (generated by newest vertex ...
                bisection)
60             mesh.node(newNode,:) = ( ...
                mesh.node(mesh.elem(ct,2),:) + ...
61                                     mesh.node(mesh.elem(ct,3),:) ...
                                        )/2;

```

```

62         mesh.solu(newNode) = ( mesh.solu(mesh.elem(ct,2)) + ...
63                               mesh.solu(mesh.elem(ct,3)) )/2;
64         mesh.type(newNode) = 2;
65
66         % Find the element which shares the base edge of the ...
67         % current
68         % element. If it is 0, it means the base of the ...
69         % current element
70         % is on the boundary.
71         ct = dualEdge( mesh.elem(ct,3) , mesh.elem(ct,2) );
72         if ( ct == 0 ), index = 0; end
73         % the while will ended if
74         % 1. ct==0 means we are on the boundary
75         % 2. base(ct) is already marked
76
77     end
78     end % end while for recursive marking
79 end % end of for loop on all elements
80 % Detailed explanation of the algorithm can be found at
81 % Manual --> Algorithms --> Bisection
82
83 % delete possible empty entries
84 ix = (mesh.type == 5);
85 mesh.node(ix,:) = [];
86 mesh.type(ix) = [];
87 mesh.solu(ix) = [];
88 mesh.f(ix) = [];
89 %-----
90 % Refine marked edges for each triangle
91 %-----
92 numnew = 2*sum(marker~=0); % number of new elements need to be added
93 mesh.elem = [mesh.elem; zeros(numnew,3)];
94 eta = [eta; zeros(numnew,1)];
95
96 inew = NT + 1; % index for current new added right child
97 for t = 1:NT
98     base = d2p(mesh.elem(t,2),mesh.elem(t,3));
99     if (marker(base)>0)
100         p = [mesh.elem(t,:) , marker(base)];
101
102         % Case 1: divide the current marked triangle
103         mesh.elem(t,:) = [p(4),p(1),p(2)]; % t is always a left ...
104         % child
105         mesh.elem(inew,:) = [p(4),p(3),p(1)]; % new is a right child

```

```

103     eta(t) = eta(t)/2; % update error indicators
104     eta(inew) = eta(t);
105     inew = inew + 1;
106
107     % Case 2: divide the right child, different, careful!!!
108     right = d2p(p(3),p(1));
109     if (marker(right)>0)
110         mesh.elem(inew-1,:) = [marker(right),p(4),p(3)];
111         mesh.elem(inew,:) = [marker(right),p(1),p(4)];
112         eta(inew-1) = eta(inew-1)/2;
113         eta(inew) = eta(inew-1);
114         inew = inew + 1;
115     end
116
117     % Case 3: divide the left child, similar to the case 1.
118     left = d2p(p(1),p(2));
119     if (marker(left)>0)
120         mesh.elem(t,:) = [marker(left),p(4),p(1)];
121         mesh.elem(inew,:) = [marker(left),p(2),p(4)];
122         eta(t) = eta(t)/2;
123         eta(inew) = eta(t);
124         inew = inew + 1;
125     end
126
127     end % end of refinement of one element
128 end % end of for loop on all elements
129
130 % delete possible empty entries
131 mesh.elem = mesh.elem(1:inew-1,:);
132 eta = eta(1:inew-1,:);
133
134 %-----
135 % Update boundary edges
136 %-----
137 mesh.Dirichlet = updatebd(mesh.Dirichlet,marker,d2p);
138 %-----
139 % End of function BISECTION
140 %-----
141
142 function bdEdge = updatebd(bdEdge,marker,d2p)
143 % UPDATEDBD refine the boundary edges
144 % INPUT
145 %   bdEdge: set of boundary edges
146 %   marker: new node index for marked edge

```

```

147 %      d2p:  index mapping from dual edge to primary edge
148
149 NB = size (bdEdge,1);
150 for k = 1:NB
151     i = bdEdge(k,1);
152     j = bdEdge(k,2);
153     if marker(d2p(i,j)) >0
154         bdEdge(k,:) = [i,marker(d2p(i,j))];
155         bdEdge(size(bdEdge,1)+1,:) = [marker(d2p(i,j)),j];
156     end
157 end

```

```

1 function [mesh,eta] = bisection_f(mesh,eta,theta)
2 % BISECTION refines the triangulation using newest vertex bisection
3 %
4 % USAGE
5 %      [mesh,eta] = bisection(mesh,eta,theta)
6 %
7 % INPUT
8 %      mesh:  current mesh
9 %      eta:   error indicator for each triangle
10 %      theta: parameter in (0,1).
11 %            We mark minimal number of triangles M such that
12 %             $\sum_{T \in M} \eta_T > \theta * \sum \eta_T$ 
13 %
14 % OUTPUT
15 %      mesh:  new mesh after refinement
16 %      eta:   new error indicator for each triangle
17 %-----
18 % Construct data structure
19 %-----
20 edge = [mesh.elem(:,[1,2]); mesh.elem(:,[1,3]); mesh.elem(:,[2,3])];
21 edge = unique(sort(edge,2),'rows');
22
23 N = size(mesh.node,1);
24 NT = size(mesh.elem,1);
25 NE = size(edge,1);
26
27 dualEdge = ...
28     sparse(mesh.elem(:,[1,2,3]),mesh.elem(:,[2,3,1]),[1:NT,1:NT,1:NT]);
29 d2p = sparse(edge(:,[1,2]),edge(:,[2,1]),[1:NE,1:NE]);
30 %-----

```

```

30 % Memory management for node array
31 %-----
32 recycle = find(mesh.types==0);
33 last = length(recycle);
34 % Coarsening can create empty spaces in the node array. We ...
    collect those
35 % scattered spaces in recycle array and 'last' will point out ...
    the last
36 % empty node index.
37 % mesh.type array is used to distinguish the type of nodes:
38 %   0: empty nodes (deleted by coarsening);
39 %   1: nodes in the initial triangulation or nodes from regular ...
    refinement;
40 %   2: new added nodes due to refinement;
41 %   5: temporary node (will be deleted when bisection finished).
42
43 %-----
44 % Mark triangles according to the error indicator
45 %-----
46 total = sum(eta);
47 current = 0;
48
49 [temp,ix] = sort(-eta); % sort in descent order
50 marker = zeros(NE,1);
51
52 % initialize for possible new nodes
53 mesh.node = [mesh.node; zeros(NE,2)];
54 mesh.type = [mesh.type; uint8(5*ones(NE,1))];
55 mesh.solu = [mesh.solu; zeros(NE,1)];
56 mesh.f=[mesh.f; zeros(NE,1)];
57 for t = 1:NT
58
59     if (current > theta*total), break; end % est on marked elem ...
        big enough
60
61     index = 1;
62     ct = ix(t);
63
64     while (index==1)
65
66         base = d2p(mesh.elem(ct,2),mesh.elem(ct,3));
67         if (marker(base)>0) % base is already marked
68             index = 0;
69         else

```

```

70         current = current + eta(ct);
71         if (last==0)
72             newNode = N + 1;
73             N = N+1;
74         end
75         if (last > 0)
76             newNode = recycle(last);
77             last = last - 1;
78         end
79         marker( d2p(mesh.elem(ct,2),mesh.elem(ct,3)) ) = ...
            newNode;
80
81         % A new node is added to the mesh. Numerical ...
            solution at this
82         % new added node is approximated by linear ...
            interpolation. Type
83         % of this node is 2 (generated by newest vertex ...
            bisection)
84         mesh.node(newNode,:) = ( ...
            mesh.node(mesh.elem(ct,2),:) + ...
85                                     mesh.node(mesh.elem(ct,3),:) ...
                                     )/2;
86         mesh.solu(newNode) = ( mesh.solu(mesh.elem(ct,2)) + ...
            mesh.solu(mesh.elem(ct,3)) )/2;
87
88
89         mesh.type(newNode) = 2;
90         mesh.f(newNode) = ( mesh.f(mesh.elem(ct,2)) + ...
            mesh.f(mesh.elem(ct,3)) )/2;
91
92         % Find the element which shares the base edge of the ...
            current
93         % element. If it is 0, it means the base of the ...
            current element
94         % is on the boundary.
95         ct = dualEdge( mesh.elem(ct,3), mesh.elem(ct,2) );
96         if ( ct == 0 ), index = 0; end
97         % the while will ended if
98         % 1. ct==0 means we are on the boundary
99         % 2. base(ct) is already marked
100
101     end
102     end % end while for recursive marking
103 end % end of for loop on all elements
104 % Detailed explanation of the algorithm can be found at
105 % Manual --> Algorithms --> Bisection

```

```

106
107 % delete possible empty entries
108 ix = (mesh.type == 5);
109 mesh.node(ix,:) = [];
110 mesh.type(ix) = [];
111 mesh.solu(ix) = [];
112 mesh.f(ix) = [];
113
114 %-----
115 % Refine marked edges for each triangle
116 %-----
117 numnew = 2*sum(marker~=0); %number of new elements need to be added
118 mesh.elem = [mesh.elem; zeros(numnew,3)];
119 eta = [eta; zeros(numnew,1)];
120
121 inew = NT + 1; % index for current new added right child
122 for t = 1:NT
123     base = d2p(mesh.elem(t,2), mesh.elem(t,3));
124     if (marker(base)>0)
125         p = [mesh.elem(t,:), marker(base)];
126
127         % Case 1: divide the current marked triangle
128         mesh.elem(t,:) = [p(4),p(1),p(2)]; % t is always a left ...
            child
129         mesh.elem(inew,:) = [p(4),p(3),p(1)]; % new is a right child
130         eta(t) = eta(t)/2; % update error indicators
131         eta(inew) = eta(t);
132         inew = inew + 1;
133
134         % Case 2: divide the right child, different, careful!!!
135         right = d2p(p(3),p(1));
136         if (marker(right)>0)
137             mesh.elem(inew-1,:) = [marker(right),p(4),p(3)];
138             mesh.elem(inew,:) = [marker(right),p(1),p(4)];
139             eta(inew-1) = eta(inew-1)/2;
140             eta(inew) = eta(inew-1);
141             inew = inew + 1;
142         end
143
144         % Case 3: divide the left child, similar to the case 1.
145         left = d2p(p(1),p(2));
146         if (marker(left)>0)
147             mesh.elem(t,:) = [marker(left),p(4),p(1)];
148             mesh.elem(inew,:) = [marker(left),p(2),p(4)];

```

```

149         eta(t) = eta(t)/2;
150         eta(inew) = eta(t);
151         inew = inew + 1;
152     end
153
154     end % end of refinement of one element
155 end % end of for loop on all elements
156
157 % delete possible empty entries
158 mesh.elem = mesh.elem(1:inew-1,:);
159 eta = eta(1:inew-1,:);
160
161 %-----
162 % Update boundary edges
163 %-----
164 mesh.Dirichlet = updatebd(mesh.Dirichlet,marker,d2p);
165
166
167 %-----
168 % End of function BISECTION
169 %-----
170
171 %-----
172 % Sub functions called by BISECTION
173 %-----
174
175 function bdEdge = updatebd(bdEdge,marker,d2p)
176 % UPDATEDBD refine the boundary edges
177 %
178 % USAGE
179 %     bdEdge = updatebd(bdEdge,marker,d2p)
180 %
181 % INPUT
182 %     bdEdge: set of boundary edges
183 %     marker: new node index for marked edge
184 %     d2p: index mapping from dual edge to primary edge
185 %
186 % OUTPUT
187 %     bdEdge: set of refined boundary edges
188 %
189
190 NB = size(bdEdge,1);
191 for k = 1:NB
192     i = bdEdge(k,1);

```

```

193     j = bdEdge(k,2);
194     if marker(d2p(i,j)) >0
195         bdEdge(k,:) = [i,marker(d2p(i,j))];
196         bdEdge(size(bdEdge,1)+1,:) = [marker(d2p(i,j)),j];
197     end
198 end
199 %-----
200 % End of function UPDATEDBD
201 %-----

```

```

1 function eta = estimate(mesh)
2 % INPUT
3 %     mesh:    current mesh
4 %     u:       finite element solution
5 %
6 % OUTPUT
7 %     eta:     error estimator on each element
8
9
10 [dudx,dudy] = ZZ(mesh);
11 [dudxx,dudxy] = sgrad(mesh,dudx);
12 [dudyx,dudyy] = sgrad(mesh,dudy);
13 eta = abs(dudxx) + abs(dudxy) + abs(dudyx) + abs(dudyy);
14
15 function [dudx,dudy] = ZZ(mesh)
16 % ZZ is the Z-Z recovery error estimator
17 N = size(mesh.node,1);
18 NT = size(mesh.elem,1);
19 activeNode = find(mesh.type~=0);
20 nonActive = setdiff(1:N,activeNode);
21
22 [sdudx,sdudy,area] = sgrad(mesh,mesh.solu);
23 patcharea = sparse(mesh.elem,ones(NT,3),area*[1,1,1],N,1);
24 dudx = accumarray(mesh.elem(:),[sdudx;sdudx;sdudx],[N 1]);
25 dudy = accumarray(mesh.elem(:),[sdudy;sdudy;sdudy],[N 1]);
26 dudx(activeNode) = dudx(activeNode)./patcharea(activeNode);
27 dudy(activeNode) = dudy(activeNode)./patcharea(activeNode);
28 dudx(nonActive) = 0;
29 dudy(nonActive) = 0;
30 function [sdudx,sdudy,area] = sgrad(mesh,u)
31 % SGRAD compute the scaled grad of w on the mesh
32 % it is elementwise

```

```

33 ve(:, :, 1) = mesh.node(mesh.elem(:, 3), :) - mesh.node(mesh.elem(:, 2), :);
34 ve(:, :, 2) = mesh.node(mesh.elem(:, 1), :) - mesh.node(mesh.elem(:, 3), :);
35 ve(:, :, 3) = mesh.node(mesh.elem(:, 2), :) - mesh.node(mesh.elem(:, 1), :);
36 area = 0.5 * abs(-ve(:, 1, 3) .* ve(:, 2, 2) + ve(:, 2, 3) .* ve(:, 1, 2));
37 sdudx = -0.5 * ( u(mesh.elem(:, 1)) .* ve(:, 2, 1) ...
38                 + u(mesh.elem(:, 2)) .* ve(:, 2, 2) ...
39                 + u(mesh.elem(:, 3)) .* ve(:, 2, 3) );
40 sdudy = 0.5 * ( u(mesh.elem(:, 1)) .* ve(:, 1, 1) ...
41                + u(mesh.elem(:, 2)) .* ve(:, 1, 2) ...
42                + u(mesh.elem(:, 3)) .* ve(:, 1, 3) );

```

```

1 function eta_f = estimate_f(faprox, mesh, q)
2 % Estimate ||f-f_hat||
3 ve(:, :, 1) = mesh.node(mesh.elem(:, 3), :) - mesh.node(mesh.elem(:, 2), :);
4 ve(:, :, 2) = mesh.node(mesh.elem(:, 1), :) - mesh.node(mesh.elem(:, 3), :);
5 ve(:, :, 3) = mesh.node(mesh.elem(:, 2), :) - mesh.node(mesh.elem(:, 1), :);
6 area = 0.5 * abs(-ve(:, 1, 3) .* ve(:, 2, 2) + ve(:, 2, 3) .* ve(:, 1, 2));
7 fexact = ff(mesh.node);
8 temp1 = abs(faprox(mesh.elem(:, 1)) - fexact(mesh.elem(:, 1))) .^ q;
9 temp2 = abs(faprox(mesh.elem(:, 2)) - fexact(mesh.elem(:, 2))) .^ q;
10 temp3 = abs(faprox(mesh.elem(:, 3)) - fexact(mesh.elem(:, 3))) .^ q;
11 T = (1/3) * (temp1 + temp2 + temp3) .* area;
12 eta_f = T .^(1/q);

```

```

1 function [l2error, h1error] = femerror(mesh, u, u_x, u_y)
2 %
3 % FEMERROR calculates the error of the piecewise linear finite ...
4 % approximation to the exact solution of a PDE in the L2- and ...
5 % INPUT
6 % mesh : current mesh
7 % u : exact solution, given as a function
8 % u_x, u_y : partial derivatives of the exact solution, given ...
9 % as fct's
10 % OUTPUT
11 % l2error : error measured in the L2-norm
12 % h1error : error measured in the H1-norm
13 %-----
14 % Check H1-norm needed or not

```

```

15 %-----
16 if (nargin < 4 || nargsout < 2)
17     isH1 = 0; h1error = 0;
18 else
19     isH1 = 1;
20 end
21
22 %-----
23 % Evaluate u, u_x, u_y at the midpoint of each element
24 %-----
25 center = ( mesh.node(mesh.elem(:,1),:) ...
26             + mesh.node(mesh.elem(:,2),:) ...
27             + mesh.node(mesh.elem(:,3),:) )/3;
28 u_m = feval(u, center); % exact solution at the midpoints
29
30 if (isH1 == 1)
31     u_x_m = feval(u_x, center); % x-derivative at the midpoints
32     u_y_m = feval(u_y, center); % y-derivative at the midpoints
33 end
34
35 %-----
36 % Evaluate uh, uh_x, uh_y at the midpoint of each element
37 %-----
38 edge(:, :, 1) = ...
39     mesh.node(mesh.elem(:,3),:) - mesh.node(mesh.elem(:,2),:);
40 edge(:, :, 2) = ...
41     mesh.node(mesh.elem(:,1),:) - mesh.node(mesh.elem(:,3),:);
42 edge(:, :, 3) = ...
43     mesh.node(mesh.elem(:,2),:) - mesh.node(mesh.elem(:,1),:);
44 area = 0.5*abs(-edge(:,1,3).*edge(:,2,2)+edge(:,2,3).*edge(:,1,2));
45
46 uh_m = ( mesh.solu(mesh.elem(:,1)) + mesh.solu(mesh.elem(:,2)) ...
47           + mesh.solu(mesh.elem(:,3)) )/3; % compute uh at the ...
48           midpoint
49
50 if (isH1 == 1) % compute gradient of uh if needed
51     uh_x = -( mesh.solu(mesh.elem(:,1)).*edge(:,2,1) ...
52               + mesh.solu(mesh.elem(:,2)).*edge(:,2,2) ...
53               + mesh.solu(mesh.elem(:,3)).*edge(:,2,3) )./(2*area);
54     uh_y = ( mesh.solu(mesh.elem(:,1)).*edge(:,1,1) ...
55               + mesh.solu(mesh.elem(:,2)).*edge(:,1,2) ...
56               + mesh.solu(mesh.elem(:,3)).*edge(:,1,3) )./(2*area);
57 end
58
59

```

```

55 %-----
56 % Midpoint rule for numerical integration
57 %-----
58 l2error2 = sum(area.*((u_m - uh_m).^2));
59 l2error = sqrt(l2error2);
60
61 if (isH1 == 1) % compute H1 error if needed
62     h1error2 = sum(area.*((u_x_m - uh_x).^2 + (u_y_m - uh_y).^2));
63     h1error = sqrt(l2error2 + h1error2);
64 end
65
66 %-----
67 % End of function FEMERROR
68 %-----

```

```

1 function z = ff(p)
2 % input
3 %     p coordinate;
4 %     output the value of right hand side at p
5 d=size(p);
6 r0=2*sqrt(2);
7 vone=(r0)*ones(d(1),1);
8 r = sqrt(sum(p.^2,2));
9 t = atan2(p(:,2),p(:,1));
10 t = (t>=0).*t + (t<0).*(t+2*pi);
11 S2=sin((2*t)/3);
12 z1=(S2./r).*(vone/15+(4/9)*(1./r).*(2*ones(d(1),1)+r0/15*(r-vone)));
13 z1(isnan(z1))=10^3;
14 z=(r<=vone)*0+(r>vone).*z1;

```

```

1 function elem = label(node,elem)
2 % LABEL label the longest edge of each triangle as the base
3 %
4 % USAGE
5 %     elem = label(node,elem)
6 %
7 % Input:
8 %     node: coordinates of nodes
9 %     elem: element array
10 % Compute length of each edge
11 %-----

```

```

12     edgelenlength = zeros(size(elem));
13     % initialize to accelerate access
14
15     edgelenlength(:,1) = (node(elem(:,3),1)-node(elem(:,2),1)).^2 ...
16                         + (node(elem(:,3),2)-node(elem(:,2),2)).^2;
17     edgelenlength(:,2) = (node(elem(:,1),1)-node(elem(:,3),1)).^2 ...
18                         + (node(elem(:,1),2)-node(elem(:,3),2)).^2;
19     edgelenlength(:,3) = (node(elem(:,1),1)-node(elem(:,2),1)).^2 ...
20                         + (node(elem(:,1),2)-node(elem(:,2),2)).^2;
21
22     % Re-labelling according the longest edge
23     [temp,I] = max(edgelenlength,[],2);
24     elem((I==2),[1 2 3]) = elem((I==2), [2 3 1]);
25     elem((I==3),[1 2 3]) = elem((I==3), [3 1 2]);
26
27     % End of function LABEL

```

```

1  function out=RHS(mesh)
2  % compute approximation of RHS on each node
3  N = size(mesh.node,1);
4  ve(:, :, 1) = mesh.node(mesh.elem(:,3), :) - mesh.node(mesh.elem(:,2), :);
5  ve(:, :, 2) = mesh.node(mesh.elem(:,1), :) - mesh.node(mesh.elem(:,3), :);
6  ve(:, :, 3) = mesh.node(mesh.elem(:,2), :) - mesh.node(mesh.elem(:,1), :);
7  area = 0.5*abs(-ve(:,1,3).*ve(:,2,2)+ve(:,2,3).*ve(:,1,2));
8  % Assemble Mass matrix
9  M = sparse(mesh.elem(:,[1,1,1,2,2,2,3,3,3]), ...
10            mesh.elem(:,[1,2,3,1,2,3,1,2,3]), ...
11            area*[2,1,1,1,2,1,1,1,2]/12, N, N);
12
13  c1=mesh.node(mesh.elem(:,1),:);
14  d21=mesh.node(mesh.elem(:,2),:);
15  d31=mesh.node(mesh.elem(:,3),:);
16  Ast=(ff(c1)+ff(d21)+ff(d31))/3;
17  nC=size(mesh.node,1);
18  b = accumarray(mesh.elem(:), repmat(area.*Ast/3,3,1),[nC 1]);
19  out=M\b;

```

```

1  function [u, energy] = PDE(mesh,f, g_D)
2  % INPUT
3  %     mesh:    current mesh
4  %     f:       right side or data

```

```

5 % g_D: Dirichelet condition
6 % OUTPUT
7 % u: solution on the current mesh
8 % energy: energy of the discrete solution u
9 %-----
10 % Initialize the data
11 %-----
12 N = size(mesh.node,1);
13 A = sparse(N,N);
14 u = zeros(N,1);
15
16 %-----
17 % Compute vedge: edge as a vector and area of each element
18 %-----
19 ve(:, :, 1) = mesh.node(mesh.elem(:, 3), :) - mesh.node(mesh.elem(:, 2), :);
20 ve(:, :, 2) = mesh.node(mesh.elem(:, 1), :) - mesh.node(mesh.elem(:, 3), :);
21 ve(:, :, 3) = mesh.node(mesh.elem(:, 2), :) - mesh.node(mesh.elem(:, 1), :);
22 area = 0.5*abs(-ve(:, 1, 3).*ve(:, 2, 2)+ve(:, 2, 3).*ve(:, 1, 2));
23
24 %-----
25 % Assemble Stiffness matrix
26 %-----
27 for i = 1:3
28     for j = 1:3
29         % Aij = \int_T grad u_i * grad u_j
30         Aij = (ve(:, 1, i).*ve(:, 1, j)+ve(:, 2, i).*ve(:, 2, j))./(4*area);
31         A = A + sparse(mesh.elem(:, i), mesh.elem(:, j), Aij, N, N);
32     end
33 end
34 %-----
35 % Assemble Mass matrix
36 %-----
37 M = sparse(mesh.elem(:, [1, 1, 1, 2, 2, 2, 3, 3, 3]), ...
38             mesh.elem(:, [1, 2, 3, 1, 2, 3, 1, 2, 3]), ...
39             area*[2, 1, 1, 1, 2, 1, 1, 1, 2]/12, N, N);
40 %-----
41 % Assemble right-hand-side
42 %-----
43 b = M*mesh.f;
44 %-----
45 bdNode = unique(mesh.Dirichlet);
46 u(bdNode) = feval(g_D, mesh.node(bdNode, :));
47 b = b - A * u; % adjust the right hand side
48 %-----

```

```

49 % Solve the linear system A * U = B for the free nodes
50 %-----
51 freeNode = find(mesh.type>0);
52 freeNode = setdiff(freeNode, bdNode);
53 if size(freeNode)>0
54     u(freeNode) = A(freeNode, freeNode) \ b(freeNode);
55 end
56
57 %-----
58 % Compute the energy of u
59 %-----
60 if (nargout > 1)
61     energy = full(0.5*u'*(A*u) - u'*M*mesh.f);
62 end % otherwise do not need to compute the discrete energy
63 %-----
64 % end of function POISSON
65 %-----

```

```

1 function z = u(p)
2 % exact solution of the test problem
3 rho_0=2*sqrt(2);
4 mu=5;
5 d=size(p);
6 vone=(rho_0)*ones(d(1),1);
7 r = sqrt(sum(p.^2,2));
8 theta = atan2(p(:,2),p(:,1));
9 theta = (theta>=0).*theta + (theta<0).*(theta+2*pi);
10 S=sin(2*theta/3);
11 z1 =(r.^(2/3)).*S;
12 z2=(vone.^(2/3)).*S+(2/(3*mu)).*(vone.^(1/3)).*S.*(r-vone);
13 z=(r<=vone).*z1+(r>vone).*z2;

```

```

1 function [lqnorm,E] = Uerror(mesh,q)
2 %This Function Returns ||u-u_hat||_L^q and Eta(current mesh)
3 ve(:, :, 1) = mesh.node(mesh.elem(:,3),:) - mesh.node(mesh.elem(:,2),:);
4 ve(:, :, 2) = mesh.node(mesh.elem(:,1),:) - mesh.node(mesh.elem(:,3),:);
5 ve(:, :, 3) = mesh.node(mesh.elem(:,2),:) - mesh.node(mesh.elem(:,1),:);
6 area_u = 0.5*abs(-ve(:,1,3).*ve(:,2,2)+ve(:,2,3).*ve(:,1,2));
7 Uapprox=mesh.solu;
8 U_exact=u(mesh.node);
9 T1=abs(Uapprox(mesh.elem(:,1))-U_exact(mesh.elem(:,1))).^q;

```

```

10 T2=abs(Uapprox(mesh.elem(:,2))-U_exact(mesh.elem(:,2))).^q;
11 T3=abs(Uapprox(mesh.elem(:,3))-U_exact(mesh.elem(:,3))).^q;
12 E=((T1+T2+T3)).*area_u/3).^ (1/q);
13 lqnorm=sum(E.^q)^(1/q);

```

۲.آ کد متلب مثال ۲

کد متلب مثال دوم همان کد متلب‌های مثال ۱ می‌باشد با این تفاوت که در مثال ۱ $f = 0$ است و بنابراین نیازی به اجرای الگوریتم RHS نمی‌باشد.

- [1] Mohammad Asadzadeh, *An introduction to the finite element method(fem) for differential equations*, (2012).
- [2] Ivo Babuška and Werner C Rheinboldt, *A-posteriori error estimates for the finite element method*, **12** (1978), 1597–1615.
- [3] Wolfgang Bangerth, Ralf Hartmann, and Guido Kanschat, *deal.ii general-purpose object-oriented finite element library*, **33** (2007), 24.
- [4] Peter Binev, Wolfgang Dahmen, and Ron DeVore, *Adaptive finite element methods with convergence rates*, **97** (2004), 219–268.
- [5] Peter Binev, Wolfgang Dahmen, Ronald DeVore, and Pencho Petrushev, *Approximation classes for adaptive methods*, **28** (2002), 391p–416p.
- [6] Peter Binev and Ronald DeVore, *Fast computation in adaptive tree approximation*, **97** (2004), 193–217.
- [7] Mikhail Shlemovich Birman and Mikhail Zakharovich Solomyak, *Piecewise-polynomial approximations of functions of the classes w_p^α* , **115** (1967), 331–355.
- [8] Andrea Bonito, J Manuel Cascón, Pedro Morin, and Ricardo H Nochetto, *Afem for geometric pde*, Springer, 2013, pp. 257–306.
- [9] Andrea Bonito and Ricardo H Nochetto, *Quasi-optimal convergence rate of an adaptive discontinuous galerkin method*, **48** (2010), 734–771.
- [10] Susanne C Brenner and Ridgway Scott, *The mathematical theory of finite element methods*, vol. 15, Springer, 2008.
- [11] R Bruce kellogg, *On the poisson equation with intersecting interfaces*, **4** (1974), 101–129.

- [12] J Manuel Cascon, Christian Kreuzer, Ricardo H Nochetto, and Kunibert G Siebert, *Quasi-optimal convergence rate for an adaptive finite element method*, **46** (2008), 2524–2550.
- [13] Albert Cohen, Wolfgang Dahmen, Ingrid Daubechies, and Ronald DeVore, *Tree approximation and optimal encoding*, **11** (2001), 192–226.
- [14] Albert Cohen, Ronald DeVore, and Ricardo H Nochetto, *Convergence rates of afem with h^{-1} data*, **12** (2012), 671–718.
- [15] Stephan Dahlke and Ronald A DeVore, *Besov regularity for elliptic boundary value problems*, **22** (1997), 1–16.
- [16] Michel C Delfour and J-P Zolésio, *Shapes and geometries: metrics, analysis, differential calculus, and optimization*, vol. 22, Siam, 2011.
- [17] Stig Larsson and Vidar Thomée, *Partial differential equations with numerical methods*, vol. 45, Springer Science & Business, 2008.
- [18] Khamron Mekchay and Ricardo H Nochetto, *Convergence of adaptive finite element methods for general second order linear elliptic pdes*, **43** (2005), 1803–1827.
- [19] N.G. Meyers, *An l_p -estimate for the gradient of solutions of second order elliptic divergence equations*, (1963), 189–206.
- [20] Ricardo H Nochetto, Kunibert G Siebert, and Andreas Veiser, *Theory of adaptive finite element methods: an introduction*, Springer, 2009, pp. 409–542.
- [21] Junuthula Narasimha Reddy, *An introduction to the finite element method*, vol. 2, McGraw-Hill New York, 1993.
- [22] Michael Renardy and Robert C. Rogers, *An introduction to partial differential equation*, **4** (2004).
- [23] V.Thomee S.Larsson, *Partial differential equations*, (2009).
- [24] Stevenson and Rob, *The completion of locally refined simplicial partitions created by bisection*, **77** (2008), 227–241.

-
- [25] Rob Stevenson, *Optimality of a standard adaptive finite element method*, **7** (2007), 245–269.
 - [26] DJ Turcke and GM McNeice, *Guidelines for selecting finite element grids based on an optimization study*, **4** (1974), 499–519.
 - [27] Vitaly A Volpert, *Elliptic partial differential equations*, Springer, 2011.
 - [28] Olgierd C Zienkiewicz, *The background of error estimation and adaptivity in finite element computations*, **195** (2006), 207–213.

واژه‌نامه فارسی به انگلیسی

Diffusion	انتشار
Initiall	اولیه
Electro-Magnatic	الکترومغناطیس
Perturbation	اختلال
Finite Element	اجزای محدود
Euclidean	اقلیدس
Measure	اندازه
Algoritm	الگوریتم
Quadrature	انتگرال گیری عددی
Scalar	اسکالر
Standard	استاندارد
Partition	افراز
Strategy	استراتژی
Elliptic	بیضوی
Optimal	بهینه
Estimate	برآورد
Vector	بردار
Ellipticity	بیضیت
Proof	برهان
Label	برچسب
Continuous	پیوسته
Basic	پایه
Jump	پرش
Adaptive	تطبیق
Approximation	تقریب

Adaptivity	تطبیقی
Function	تابع
Piecewise	تکه‌ای
Theoretical	تئوری
Local perturbation estimates	تخمین‌های اختلال محلی
Refinement	تظریف
Galerkin approximation	تقریب گالرکین
Orthogonal projection	تصویر متعامد
Combine	ترکیب
Constant	ثابت
Partial	جزیی
Forest	جنگل
Manifolds	چندشاخه‌ای
Polyhedral	چندوجهی
Polynomial	چندجمله‌ای
A posteriori error	خطای پسین
A priori error	خطای پیشین
Tolerance	خطا
Output	خروجی
Linear	خطی
Domain	دامنه
Accuracy	دقت
Degree	درجه
Bisection	دوبخشی
Interpolation	درون‌یابی
Sequence	دنباله
Media	رسانه
Method	روش
Vertex	راس
Solution	راه‌حل
Galerkin method	روش گالرکین
Subroutines	زیرروال

Subdomain	زیردامنه
Cell	سلول
Surface	سطوح
Right hand side	سمت راست
Level	سطح
Compatible	سازگار
Coefficient	ضریب
Weak	ضعیف
Spectral	طیفی
Numerical	عددی
General	عمومی
Performance	عملکرد
Operator	عملگر
Membership	عضویت
Element	عنصر
Space	فضا
Sobolev space	فضای سوبولوف
Quadratic form	فرم دوخطی
Test space	فضای آزمون
Trial space	فضای آزمایش
Inner product space	فضای ضرب داخلی
Hilbert space	فضای هیلبرت
Green formula	فرمول گرین
Theory	قضیه
Lax-Milgram theory	قضیه لکس میلگرم
Rule	قانون
Perturbation theory	قضیه اختلال
Boundary	کرانه
Partial	کاربردی
Lipschitz	لیپ شیتز
Field	میدان
Triangulation	مثلث‌سازی

Curve.....	منحنی
Partial differential equation.....	معادله دیفرانسیل جزئی
Positive.....	مثبت
Matrix.....	ماتریس
Value.....	مقدار
Set.....	مجموعه
Symmetric.....	متقارن
Eigenvalue.....	مقدار ویژه
Direct.....	مستقیم
Mesh.....	مش
Mean value.....	مقدار میانی
Laplace equation.....	معادله لاپلاس
Poisson problem.....	مساله پواسون
Invertible.....	معکوس شدنی
Triangle.....	مثلث
Opposite.....	متضاد
Absolute.....	مطلق
Admissible.....	مجاز
Positive definite.....	معین مثبت
Discontinuous.....	ناپیوستگی
Heterogeneous.....	ناهمگن
Norm.....	نورم
Result.....	نتیجه
Rate.....	نسبت
Holder inequality.....	نامساوی هولدر
Map.....	نگاشت
Energy norm.....	نورم انرژی
Oscillation.....	نوسان
Regularity.....	نظم
Real.....	واقعی
Input.....	ورودی
Convergence.....	همگرا

Equivalent.....	هم‌ارز
Smooth.....	هموار
Uniquenss.....	یکتایی
Uniformly.....	یکنواخت

واژه‌نامه انگلیسی به فارسی

Adaptive	تطبیق
Adaptivity	تطبیقی
Approximation	تقریب
Algorithm	الگوریتم
Accuracy	دقت
Absolute	مطلق
Admissible	مجاز
A posteriori error	خطای پسین
A priori error	خطای پیشین
Boundary	کرانه
Basic	پایه
Coefficient	ضریب
Cell	سلول
Curve	منحنی
Convergence	همگرا
Constant	ثابت
Continuous	پیوسته
Compatible	سازگار
Conforming	یکنواخت
Combin	ترکیب
Discontinuous	ناپیوستگی
Diffusion	انتشار
Domain	دامنه
Definite	معین
Direct	مستقیم

Degree	درجه
Elliptic	بیضوی
Electro-magnetic	الکترومغناطیس
Estimate	برآورد
Euclidean	اقلیدس
Ellipticity	بیضیت
Eigenvalue	مقدار ویژه
Equivalent	هم ارز
Energy norm	نورم انرژی
Element	عنصر
Field	میدان
Finite element	اجزای محدود
Function	تابع
Forest	جنگل
General	عمومی
Heterogeneous	ناهمگن
Hilbert space	فضای هیلبرت
Holder inequality	نامساوی هولدر
Initial	اولیه
Interpolation	درون‌یابی
Invertible	معکوس شدنی
Input	ورودی
Inner product space	فضای ضرب داخلی
Indicator	شاخص
Jump	پرش
Lax-milgram theory	قضیه لکس میلگرم
Lipchitz	لیپشیتز
Level	سطح
Local perturbation estimate	تخمین‌های اختلال محلی
Laplace's equation	معادله لاپلاس
Label	برچسب
Linear	خطی

Media	رسانه
Manifolds	چندشاخه‌ای
Method	روش
Matrix	ماتریس
Measure	اندازه
Mesh	مش
Mean value	مقدار میانی
Map	نگاشت
Membership	عضویت
Norm	نورم
Optimal	بهینه
Operator	عملگر
Opposite	متضاد
Output	خروجی
Orthogonal projection	تصویر متعامد
Oscillation	نوسان
Partial	جزیی
Porousmedia	رسانه متخلخل
Perrturbation	اختلال
Problem	مسئله
Polyhedral	چندوجهی
Practical	کاربردی
Perturbation theory	قضیه اختلال
Piecewise	تکه‌ای
Performance	عملکرد
Perturbed problem	مسئله اختلال
Proof	برهان
Poisson problem	مسئله پواسون
Partition	افراز
Polynomial	چندجمله‌ای
Positive definite	معین مثبت
Quadratic form	فرم دوخطی

Quadrature.....	انتگرال‌گیری عددی
Result	نتیجه
Rate	نسبت
Real.....	واقعی
Rule	قانون
Right hand side.....	سمت راست
Refinement	تظریف
Regularity	نظم
Surface.....	سطوح
Set	مجموعه
Spectral.....	طیفی
Sobolev space.....	فضای سوبولوف
Symmetric	متقارن
Scalar.....	اسکالر
Smooth.....	هموار
Subroutines.....	زیرروال
Standard	استاندارد
Subdomain	زیردامنه
Solution	راه‌حل
Strategy.....	استراتژی
Space.....	فضا
Sequence	دامنه
Triangulation	مثلثی‌سازی
Theory.....	قضیه
Theoretical.....	تئوری
Triangle.....	مثلث
Tolerance	خطا
Test space	فضای آزمون
Trial space.....	فضای آزمایش
Uniformly	یکنواخت
Uniqueness	یکتایی
Vector.....	بردار

Vertex.....	راس
Value.....	مقدار
Weak.....	ضعیف

نمایه

۱	اتحادهای گرین، ۶
ت	تابع خطی، ۲
	تصویر متعامد، ۲۵
چ	چند اندیشه، ۷
د	دامنه لیپ شیتز، ۸
	دنباله‌ی کوشی، ۳
	دنباله‌ی همگرا، ۳
	دیورژانس، ۵
ر	روش عناصر متناهی، ۱۲
س	سوپریمم اساسی، ۶
ض	ضرایب ناپیوسته، ۱
	ضرب داخلی، ۲
ع	عملگرهای بیضوی، ۱۹
	عناصر متناهی، ۱۲
ف	فرم دوخطی متقارن، ۴
۶	فضای L_P ، ۶
۱۲	فضای آزمایش، ۱۲
۱۲	فضای آزمون، ۱۲
۴	فضای باناخ، ۴
۲	فضای برداری، ۲
۲	فضای خطی، ۲
۴	فضای خطی نورم‌دار، ۴
۴	فضای دوگان، ۴
۷	فضای سوبولوف، ۷
۲	فضای ضرب داخلی، ۲
۳	فضای کامل، ۳
۳	فضای هیلبرت، ۳
ق	قضیه دیورژانس، ۵
۶	قضیه گرین، ۶
۴	قضیه لکس میلگرم، ۴
ک	کوارسیو، ۴
گ	گرادیان، ۵
ل	لم پوانکاره، ۸
م	مقارن، ۲
۲	مثبت معین، ۲

مساله پایدار، ۷

مساله خوش وضع، ۱۰

معادله دیفرانسیل با مشتقات جزئی، ۹

میدان برداری، ۵

ن

نامساوی فریدریچز، ۸

نامساوی کوشی- شوارتز، ۳

نامساوی مثلثی، ۳

نامساوی هولدر، ۷

نورم، ۴

نورم اقلیدسی، ۷

نیم نورم، ۴

Aabstract

Elliptic partial differential equations (PDEs) with discontinuous diffusion coefficients occur in application domains such as diffusions through porous media, electromagnetic field propagation on heterogeneous media, and diffusion processes on rough surfaces. The standard approach to numerically treating such problems using finite element methods is to assume that the discontinuities lie on the boundaries of the cells in the initial triangulation. However, this does not match applications where discontinuities occur on curves, surfaces, or manifolds, and could even be unknown beforehand. One of the obstacles to treating such discontinuity problems is that the usual perturbation theory for elliptic PDEs assumes bounds for the distortion of the coefficients in the L^1 norm and this in turn requires that the discontinuities are matched exactly when the coefficients are approximated. We present a new approach based on distortion of the coefficients in an L^q norm with $q < 1$ which therefore does not require the exact matching of the discontinuities. We then use this new distortion theory to formulate new adaptive finite element methods (AFEMs) for such discontinuity problems. We show that such AFEMs are optimal in the sense of distortion versus number of computations, and report insightful numerical results supporting our analysis.

keywords: Elliptic Problem, Discontinuous Coefficients, Perturbation Estimates, Adaptive Finite Element Methods, Optimal Rates of Convergence.



Shahrood University
Faculty Of Mathematical Sciences

Dissertation Submitted in Partial
Fulfillment of The Requirements For The
Degree of Master of Science in
Applied Mathematics

Adaptive Finite Element Methods For Elliptic Problems With Discontinuous Coefficients

Supervisor
Dr.Ali Mesforush

Advisor
Dr.Mehdi Ghovvatmand Jazi

by
Solmaz Enkari

December 2014