وزارت علوم، تحقیقات و فناوری
دانشگاه تحصیلات تکمیلی علوم پایه
گاوازنگ - زنجان

حل عددي معادلات ديفرانسيل با مشتقات جزئی با ضرایب نامعین به کمک شبکه عصبی

پایان‌نامهٔ کارشناسی ارشد ریاضیات کاربردی ـ آنالیز عددی

ساجد زرین پور نشرودکلی

استاد راهنما: **دکتر خدیجه ندائی اصل**

استاد مشاور: **دکتر پروین رزاقی**

دی ۱۳۹۹

چکیده

برای مدل‌سازی پدیده‌های واقعی با معادلات دیفرانسیل با مشتقات جزئی که شامل عدم قطعیت است، یکی از مشکلات وجود مجموعه‌ای از پدیده‌ها است که به عنوان مشکلات ابعاد بالا شناخته می‌شوند. خوشبختانه، اغلب تغییرات متغییرهای مدل می‌توانند توسط تعداد کمی خصوصیات دامنه توسط روش‌های کاهش مدل، ثبت شوند. برای مثال، می‌توان با استفاده از روش‌های مبتنی بر شبکه‌های عصبی متغییرهای مورد نظر را به عنوان تابعی از ضرایب ورودی اندازه‌گیری کرد. در این‌صورت، نمایش پذیری متغیرها توسط چنین شبکه‌ای را می‌توان با دید شبکه عصبی به عنوان یک تحول زمانی برای پیداکردن جواب‌های مدل توجیه کرد. در این پایان نامه، ما یک روش میانبر برای پیدا کردن جواب‌های مدل روی دو معادله دیفرانسیل با مشتقات جزئی معروف در فیزیک و مهندسی را بازبینی مینمائیم. همچنین، ما به سراغ بررسی یک روش عددی سنتی از نظر تئوری خواهیم رفت و از این طریق، احتمالات جدیدی برای استفاده از شبکه‌های عصبی در حل معادلات دیفرانسیل را مطرح خواهیم نمود.

واژه‌های کلیدی: *شبکه‌های عصبی، روش تفاصلات متناهی، روش المان‌های متناهی، معادلات دیفرانسیل با مشتقات جزئی، عدم قطعیت.*

فهرست

مقدمه

مقدارسنجی عدم قطعیت (۱۵) در فیزیک و مهندسی اغلب شامل مطالعه معادلات دیفرانسیل با مشتقات جزئی با میدان ضرایب تصادفی است. برای درک رفتار یک سیستم شامل عدم قطعیت، می‌توان کمیت‌های فیزیکی مشتق شده از معادلات دیفرانسیل توصیف کننده آن سیستم را به عنوان توابعی از میدان ضرایب استخراج کرد. اما حتی با گسسته‌سازی مناسب روی دامنه معادله و برد متغیرهای تصادفی، این کار به طور ضمنی به حل عددی معادله دیفرانسیل با مشتق جزئی به تعداد نمایی می‌انجامد.

یکی از روش‌های متداول برای مقدار سنجی عدم قطعیت، روش نمونه برداری مونته کارلو است. گرچه این روش در بسیاری از موارد کاربردی است اما کمیت اندازه‌گیری شده ذاتاً دارای نویز است. به‌علاوه، این روش قادر به پیدا کردن جواب‌های جدید در صورتی که قبلا نمونه‌گیری نشده باشند، نیست. ما به‌دنبال یافت روشی هستیم که نویز داده‌ها در جواب آن تأثیر چندانی نداشته باشند و همچنین، قادر به ارائه جواب برای حالاتی که قبلا نمونه گیری نشده باشند نیز باشد.

روش گالرکین تصادفی با استفاده چند جمله‌ای‌های آشوب یک جواب تصادفی را روی فضای متغیرهای تصادفی بسط می‌دهد و به این طریق مسئله با بعد بالا را به تعدادی معادله دیفرانسیل با مشتقات جزئی معین تبدیل می‌کند. این گونه روش‌ها به دقت زیادی درباره تعیین توزیع عدم قطعیت نیازمند هستند و از آن‌جا که پایه‌های استفاده شده مستقل از مسئله هستند، وقتی بعد متغیرهای تصادفی بالا باشد هزینه محاسباتی بسیار زیاد خواهد شد. ما به‌دنبال یافتن راهی برای حذف نیاز به محاسبه این پایه‌ها به صورت سنتی هستیم. یکی از اهداف این پایان‌نامه مطالعه نقطه ضعف روش‌های عددی سنتی همانند روش گالرکین تصادفی و روش المان‌های متناهی است. برای این منظور، تئوری روش المان‌های متناهی را مورد مطالعه دقیق‌تر قرار می‌دهیم.

هدف کار ما پارامتری کردن جواب یک معادله دیفرانسیل معین به کمک شبکه‌های عصبی (ساخت نمایشی دیگر برای جواب بر پایه ترکیب توابع) و سپس استفاده از روش‌های بهینه‌سازی برای یافتن

جواب معادله است. در این پایان‌نامه تابع مورد نظر برای پارامتری‌سازی روی میدان ضرایب معادله دیفرانسیل با مشتقات جزئی تعریف شده است. در واقع ما به دنبال کاهش بعد مبتنی بر نمایش شبکه عصبی برای حل معادلات دیفرانسیل با مشتقات جزئی همراه با عدم قطیت هستیم.

انگیزه و هدف

مدل‌سازی طبیعت همیشه با پارامترهایی همراه است که مقادیر آنها از کنترل ما خارج است. اما عموما ما درباره محدوده تغییرات این پارامترها اطلاعاتی داریم. به معادلاتی که شامل این‌گونه پارامترها هستند، معادلات با ضرایب عدم قطعیت گوییم. به طور مثال، در مورد حرکت نفت در سفره‌های زیر زمینی؛ برای بیان شیوه حرکت مایعات نیاز به دانستن مکان حفره‌ها در بافت سفره زیر زمینی داریم. این امر را می‌توان به صورت رسانایی مؤثر در حضور ناخالصی نیز در نظر گرفت. به عنوان مثالی دیگر، مسئله‌ای را مطرح می‌کنیم که نقطه شروع این رساله بوده است.

برای تشخیص سرطان پستان روش‌های متعددی موجود است. از جمله آن‌ها می‌توان به تصویر برداری پستان با بازتابش اشعه ایکس (XRM)[۱]، تصویربرداری با استفاده از ارتعاشات مغناطیسی (MRI)[۲]، تصویربرداری فراصوت (US)[۳]، توموسنتز دیجیتال (DBT)[۴]، ماموگرافی انتشار پوزیترون (PET)[۵] و توموگرافی فراصوت (UST)[۶] اشاره کرد. هرکدام از این روش‌ها اطلاعات را به طرق مختلفی نمایش می‌دهند، به این معنا که غده‌ای که در یکی از این روش‌ها غیر قابل تشخیص است در روش دیگر قابل تشخیص است؛ غده‌ای که در یک روش بافت مشکوک معرفی می‌شود، در روش دیگر می‌تواند به عنوان غده‌ای سالم و طبیعی معرفی شود. و این موضوع باعث ایجاد مشکلات بسیاری در روند تشخیص

1 Projection X-ray mammography
2 Magnetic resonance imaging
3 Ultra sound
4 Digital breast tomosynthesis
5 Positron emission mammography
6 Ultra sound tomography

و برنامه ریزی درمان می‌شود. در این مرحله، راه‌حلی که به ذهن می‌رسد، ترکیب نتایج حاصل از این روش‌ها برای بالابردن ضریب دقت است؛ لیکن مشکل دیگری مانع این‌کار می‌شود. بافت پستان بسیار کشسان است و به راحتی تغییر فرم می‌دهد. از طرفی در هرکدام از این روش‌ها بیمار به حالت خاصی قرار می‌گیرد که با روش دیگر متفاوت است. به طور مثال، طی MRI بیمار در حالت دمر قرار دارد ولی برای تصویربرداری فراصوت بیمار به پشت می‌خوابد. علاوه‌براین، در روش بایوپسی راهنمایی شده توسط MRI[1] بافت پستان توسط صفحه‌های سخت و غیرقابل انعطافی بی حرکت می‌شوند که منجر به فشرده شدن بافت نیز می‌شود. بنابراین، شکل، اندازه و مکان غده در این تصاویر متفاوت خواهد بود. این امر مقایسه تصاویر را با سختی بسیار همراه می‌کند. علاوه‌براین، برای برنامه‌ریزی پیش از جراحی، پزشک نیاز به دانستن مکان و اندازه دقیق غده دارد. بنابراین، نیاز به توسعه الگوریتم‌های ثبت غیرسخت[2] احساس می‌شود.

روش‌هائی مبتنی بر روش المان‌های متناهی[3] برای حل این مسئله ارائه شده‌اند. اما مشکل عمده این روش‌ها هزینه محاسباتی بالای آنها است. مطابق آن‌چه در (۷) گفته شده است، به طور متوسط اجرای یک شبیه‌سازی صد و بیست دقیقه به طول می‌انجامد که برای مصارف کلینیکی مقرون به صرفه نیست. مارتینز و همکارانش (۷) در سدد ارائه روشی برای کاهش این هزینه محاسباتی با استفاده از شبکه‌های عصبی بودند. گرچه، مدل ارائه شده توسط آنها زمان محاسبات را به طرز چشمگیری کاهش می‌دهد، اما پارامترهای مدل مورد مطالعه آن‌ها ثابت است. به عبارت دیگر، با توجه به این‌که این پارامترها از بدنی به بدن دیگر متفاوت هستند، برای ارائه یک مدل مختص به بیمار در زمانی قابل قبول، نیاز به توسعه مدل برای فراگیری پارامترهای دارای عدم قطعیت است. به طور مثال، نیازمند در نظر گرفتن ضریب کشسانی بدن بیمار، که یک ضریب عدم قطعیت است، هستیم. از این رو، برآن شدیم که بدنبال حل عددی معادلات دیفرانسیل (بیضوی) به کمک شبکه‌های عصبی باشیم.

در روش‌های عددی متداول، هدف یافتن تقریبی از جواب مدل در یک فضای متناهی یا غیر متناهی

[1] MRI-guided biopsy
[2] Non-rigid registration algorithm
[3] Finite element method

است که به صورت یک ترکیب خطی از پایه های آن فضا در نقاط رأسی بدست می‌آید. دقت این تقریب با تعداد نقاط رأسی به کار رفته رابطه مستقیم دارد. بنابراین، محاسبه تقریبی دقیق برای مسائل پیچیده روی دامنه‌های پیچیده، همچون محاسبه میزان تغییر فرم بافت پستان در اثر نیروهای وارده، به ناچار مستلزم استفاده از تعداد بسیار زیادی نقاط رأسی روی دامنه مدل می‌باشد که به طور مستقیم منجر به افزایش زمان محاسبه جواب مدل می‌شود. از این‌رو با تکیه بر اصول شبکه‌های عصبی بدنبال راهی برای حذف نیاز به محاسبه مستقیم پایه‌های فضای تقریب روی نقاط رأسی با استفاده از روشی مبتنی بر کاهش بعد فضای جواب با حفظ خصوصیات اصلی مورد نیاز خود از جواب هستیم. با توجه به اینکه نمایش جواب بدست آمده از شبکه‌های عصبی به صورت ترکیبی متناهی از توابع است، می‌توانیم با تعویض نمایش جواب از حالت خطی (روش‌های عددی متداول) با حالت غیر خطی (نمایش شبکه‌های عصبی) به این مهم دست یابیم. ایده، استفاده از شبکه‌های عصبی برای یادگیری نگاشتی از دامنه ضرایب عدم قطعیت به فضای جواب بر اساس مجموعه داده از قبل محاسبه شده است. به این ترتیب می‌توان گفت روشی که ما در صدد گزارش آن هستیم، یک روش کاهش بعد برای محاسبه جواب بدون نیاز به حل مستقیم معادله دیفرانسیل است. در این رساله، معادلات دیفرانسیل با مشتقات جزئی بیضوی خطی و غیر خطی ناهمگن را مورد بررسی قرار می‌دهیم. نکته حائز اهمیت این است که از آن‌جا که این روش، یک روش مبتنی بر داده است، نوع مدل (بیضوی، هذلولوی و یا سهموی) یا روش عددی استفاده شده برای نمونه‌گیری جواب و ایجاد پایگاه داده جدید، در آن تغییر چشمگیری ایجاد نمی‌نمایند. بنابراین، این روش برای هر سه نوع معادلات دیفرانسیل با مشتقات جزئی در هردو حالت همگن و ناهمگن قابل استفاده است. از طرف دیگر، از آنجا که ما از محاسبات تنسوری در لایه‌های این شبکه بهره می‌بریم، این مدل با کمترین میزان تغییرات قابل استفاده در ابعاد بالاتر نیز است. برای نمایش این ادعا، معادله اول را در یک بعد و معادله دوم را در دو بعد بررسی خواهیم کرد.

مقدارسنجی عدم قطعیت

مقدارسنجی عدم قطعیت مطالعه اعتمادپذیری نتیجه‌گیری علمی است. به عبارت دیگر به جای مطالعه درستی هر فرضیه به طور جزئی، به دنبال مطالعه ارتباط بین فرضیات هستیم. سؤال اساسی در مسائل مقدار سنجی عدم قطعیت این است که با فرض اینکه اطلاعاتی درمورد خطا در ورودی‌های مدل در دسترس باشد، در مورد تغییرات در خروجی چه می‌توان گفت؟(۱۵) برای مقدار سنجی عدم قطعیت یک کمیت فیزیکی، یکی از راه‌ها پیدا کردن نگاشتی از فضای پارامترهای عدم قطعی در ورودی مدل به فضای خروجی مدل است. خوشبختانه این نگاشت‌ها عموما وابسته به تعداد محدودی 'خصیصه' هستند که می‌توان آنها را با تعداد محدودی نمونه‌برداری و حل مدل دیفرانسیلی به‌دست آورد.

عموما عدم قطیت ناشی از فرضیات ساختار مدل و یا مقدار وردی‌های آن است. در این پایان‌نامه، ما روی عدم قطعیت حاصل از عدم قطعیت روی ورودی‌های مدل تمرکز خواهیم کرد. عدم قطعیت روی داده‌های ورودی به چهاردسته پارامترهای تصادفی، میدان‌های تصادفی، نویز سفید و همبسته طبقه‌بندی می‌شود (۱۶). در مسئله یافتن ضریب رسانایی مؤثر پر محیط ناهمگون عدم قطعیت روی داده های ورودی از نوع پارامترهای تصادفی و در مسئله تعیین سطح انرژی حالت پایه با پتانسیل زمینه ناهمگون از نوع نویز سفید هستند.

اغلب مقدار خروجی مورد نظر ما مقداری است که از فراوری داده‌های خروجی مدل اولیه حاصل می‌شود. دراین‌جا نیز، هدف ما در واقع کنترل عدم قطعیت خروجی ثانویه مدل است که از فراوری حاصل می‌شود. با فرض اینکه عدم قطعیت ورودی مدل با بردار $\boldsymbol{a} = (a_۱, \ldots, a_N)^T$ پارامتری شده باشد، جواب معادله علاوه بر ابعاد فضایی $\boldsymbol{x}$ و بعد زمان $\boldsymbol{t}$ به بردار پارامترهای رندوم $\boldsymbol{a}$ نیز وابسته است. بنابراین یک نمونه‌گیری از جواب می‌تواند به صورت $u(x, t; a)$ به ازای یک انتخاب مشخص از پارامترهای رندوم بیان شود. گرچه ما علاقه‌مند به بررسی هر کدام از این نمونه‌ها به طور مجزا نیستیم (۱۶). بنابراین ما پایگاه داده خود را بر همین اساس خواهیم ساخت. کنترل عدم قطعیت مقدار خروجی مطلوب معمولا مستلزم تکامل یک انتگرال چندگانه روی بردار پارامتری $\boldsymbol{a}$ است (۱۶).

مدل مطرح شده در این پایان‌نامه، این تکامل و تقریب عددی جواب آن را به طور همزمان بر عهده دارد. بدین منظور، به جای تلاش برای یافت جواب مدل، در سدد یافت نگاشتی از پارامترهای تصادفی ورودی به خروجی پردازش شده مدل هستیم. دقت کنید که به همین دلیل، با اینکه ما روش خود را روی معادلات بیضوی بررسی کردیم، همین روش برای معادلات هذلولوی و سهموی نیز کاراست.

تعریف مسئله

در این پایان نامه، هدف ما بررسی یک مدل میانبر برای حل مسائل معادلات دیفرانسیل با مشتقات جزئی به کمک شبکه‌های عصبی است. به عبارت روشن‌تر یافتن نگاشتی از فضای عدم قطعیت مسئله به فضای جواب. مسائلی که در این رساله برای حل انتخاب شده اند از این جهت حائز اهمیت بوده‌اند که هر دو حالت خطی و غیر خطی، معادلات دیفرانسیل غیر همگن شامل عدم قطعیت را پوشش می‌دهند.

یافتن ضریب رسانایی مؤثر در محیط ناهمگون

معادله اول، رسانایی مؤثر در یک جهت انتخاب شده درون یک محیط غیر همگون را توسط ضریب رسانش توصیف می‌کند. محیط غیر همگون، محیطی است که در آن خصوصیات مورد توجه در تمامی نقاط یکسان نیستد. این امر ممکن است به دلایلی همچون جنس‌های گوناگون مواد تشکیل دهنده یا چگالی های متفاوت مربوط باشد. فرض ما بر آن است که ضریب رسانایی ماده در جهات متفاوت یکسان نباشد و این ضریب را با $\boldsymbol{a}(x)$ نمایش می‌دهیم. این بردار از نمونه‌گیری توابعی به فرم زیر روی نقاط رأسی حاصل می‌شود:

$$\mathscr{A} = \{a \in L^{\infty}([0,1]^d) | \lambda_1 \geq a(x) \geq \lambda_0 > 0\}, \qquad (1)$$

که در آن λ_0 و λ_1 اعداد ثابتی هستند.با فرض انتخاب یک جهت دلخواه ثابت $\xi \in \mathbb{R}^d$، میزان ضریب رسانش در آن جهت مطلوب است. به عبارت دقیق‌تر، جواب معادله زیر مد نظر است

$$A_{\text{eff}}(\boldsymbol{a}) = \min_{u(x)} \int_{[0,1]^d} \boldsymbol{a}(x) ||\nabla u(x) + \xi||_2^2 \mathrm{d}x.$$

این معادله با فرم دیفرانسیلی زیر هم‌ارز است. به عبارت دیگر، جواب مسئله بهنه‌سازی فوق در معادله دیفرانسل زیر سدق میکند و جواب معادله دیفرانسیل زیر نیز تابع هدف مسئله بهینه‌سازی فوق را بهینه می‌کند.

$$-\nabla \cdot (a(x)(\nabla u(x) + \xi)) = 0 \tag{2}$$

ازاین‌رو، با جای حل مسئله مینیم‌سازی، فرم مذکور را به کمک گسسته‌سازی زیر با گام $h = \frac{1}{n}$ حل می‌نماییم.

$$\begin{aligned}
&-\sum_{k=1}^{d} \frac{a_{i+\frac{1}{2}e_k} - a_{i-\frac{1}{2}e_k}}{h} \cdot \frac{u_{i+e_k} - u_i}{h} - \sum_{k=1}^{d} a_{i-\frac{1}{2}e_k} \frac{u_{i+e_k} - 2u_i + u_{i-e_k}}{h^2} - \sum_{k=1}^{d} \xi_k \frac{a_{i+\frac{1}{2}e_k} - a_{i-\frac{1}{2}e_k}}{h} \\
&= -\sum_{k=1}^{d} \frac{a_{i+\frac{1}{2}e_k} u_{i+e_k} - a_{i+\frac{1}{2}e_k} u_i - a_{i-\frac{1}{2}e_k} u_{i+e_k} + a_{i-\frac{1}{2}e_k} u_i + a_{i-\frac{1}{2}e_k} u_{i+e_k} - 2a_{i-\frac{1}{2}e_k} u_i + a_{i-\frac{1}{2}e_k} u_{i-e_k}}{h^2} \\
&\qquad - \sum_{k=1}^{d} \xi_k \frac{a_{i+\frac{1}{2}e_k} - a_{i-\frac{1}{2}e_k}}{h} \\
&= -\sum_{k=1}^{d} \frac{a_{i+\frac{1}{2}e_k} [u_{i+e_k} - u_i] - a_{i-\frac{1}{2}e_k} [u_i - u_{i-e_k}]}{h^2} - \sum_{k=1}^{d} \xi_k \frac{a_{i+\frac{1}{2}e_k} - a_{i-\frac{1}{2}e_k}}{h} \\
&= \sum_{k=1}^{d} \frac{-a_{i+\frac{1}{2}e_k} u_{i+e_k} + [a_{i+\frac{1}{2}e_k} + a_{i-\frac{1}{2}e_k}] u_i - a_{i-\frac{1}{2}e_k} u_{i-e_k}}{h^2} - \sum_{k=1}^{d} \xi_k \frac{a_{i+\frac{1}{2}e_k} - a_{i-\frac{1}{2}e_k}}{h} = 0.
\end{aligned}$$

که می‌تواند به صورت $(L_aU)_i = (b_a)_i$ نمایش داده شود که در آن

$$(L_a u)_i := \sum_{k=1}^{d} \frac{-a_{i+\frac{1}{2}e_k} u_{i+e_k} + (a_{i-\frac{1}{2}e_k} + a_{i+\frac{1}{2}e_k}) u_i - a_{i-\frac{1}{2}e_k} u_{i-e_k}}{h^2} \tag{3}$$

$$(b_a)_i := \sum_{k=1}^{d} \frac{\xi_k (a_{i+\frac{1}{2}e_k} - a_{i-\frac{1}{2}e_k})}{h}. \tag{4}$$

همان‌طور که اشاره کردیم، قصد ما حل این مسئله در حالت یک بعدی است. بنابراین در روابط بالا $d = 1$ قرار می‌دهیم. با درنظر گرفتن شرط مرزی $u_{|\partial\Omega} = 0$، در مرحله بعد اقدام به نوشتن فرم ماتریسی آن می‌نماییم.

$$L_a = \frac{1}{h^2} \begin{bmatrix} a_{1-\frac{1}{2}e_1} + a_{1+\frac{1}{2}e_1} & -a_{1+\frac{1}{2}e_1} & & & \\ -a_{2-\frac{1}{2}e_1} & a_{2-\frac{1}{2}e_1} + a_{2+\frac{1}{2}e_1} & -a_{2+\frac{1}{2}e_1} & & \\ & \ddots & \ddots & \ddots & \\ & & -a_{n-1-\frac{1}{2}e_1} & a_{n-1-\frac{1}{2}e_1} + a_{n-1+\frac{1}{2}e_1} & -a_{n-1+\frac{1}{2}e_1} \\ & & & -a_{n-\frac{1}{2}e_1} & a_{n-\frac{1}{2}e_1} + a_{n+\frac{1}{2}e_1} \end{bmatrix}_{n\times n},$$

$$b_a = \frac{1}{h} \begin{bmatrix} \xi_1 (a_{1+\frac{1}{2}e_1} - a_{1-\frac{1}{2}e_1}) \\ \xi_1 (a_{2+\frac{1}{2}e_1} - a_{2-\frac{1}{2}e_1}) \\ \vdots \\ \xi_1 (a_{n+\frac{1}{2}e_1} - a_{n-\frac{1}{2}e_1}) \end{bmatrix}_{n\times 1}, U = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}_{n\times 1}.$$

با تعاریف بالا جواب را به صورت $L_aU = b_a$ می‌توان نوشت. اما برای رسیدن به گسسته‌سازی نهایی، هنوز نیاز داریم که مسئبه را با در نظر گرفتن شرط مرزی دوره‌ای بازنویسی نماییم.

$$L_a^{\text{patched}} = \frac{1}{h^2}\begin{bmatrix} a_{1-\frac{1}{2}e_1} + a_{1+\frac{1}{2}e_1} & -a_{1+\frac{1}{2}e_1} & & & & -a_{1-\frac{1}{2}e_1} \\ & & \ddots & \ddots & \ddots & \\ -a_{n+\frac{1}{2}e_1} & & & & -a_{n-\frac{1}{2}e_1} & a_{n-\frac{1}{2}e_1} + a_{n+\frac{1}{2}e_1} \end{bmatrix}_{n\times n},$$

$$U = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}_{n\times 1},$$

که در آن $a_{1-\frac{1}{2}e_1} = \frac{a_n+a_1}{2}$ و $a_{n+\frac{1}{2}e_1} = \frac{a_1+a_n}{2}$. نهایتا پس از حل مسئله توسط یک روش تکراری با گسسته سازی مطرح شده، جواب بدست آمده را در رابطه زیر قرار می‌دهیم تا به رسانایی مؤثر دست یابیم

$$A_{\text{eff}}(a) = h^d(u_a^T L_a^{\text{patched}} u_a - 2u_a^T b_a + a^T 1). \tag{5}$$

تعیین سطح انرژی حالت پایه با پتانسیل زمینه ناهمگون

معادله دوم، معادله غیرخطی شرودینگر دو بعدی است. هدف از این معادله، یافتن میزان انرژی حالت پایه الکترون با پتانسیل اولیه همراه با عدم قطعیت است. حالت پایه سطح انرژی‌ای است که الکترون مایل به اخذ آن در دمای صفر مطلق می‌باشد. این معادله به صورت یک مسئله مقدار ویژه به فرم زیر

تعریف می‌شود. هدف در حل این مسئله یافتن کوچکترین مقدار ویژه آن (حالت پایه) است

$$-\Delta u(x) + a(x)u(x) + \sigma u(x)^3 = E_0 u(x)$$
$$x \in [0,1]^d, s.t. \int_{[0,1]^d} u(x)^2 dx = 1.$$

مشابه آنچه در قسمت قبل آمد معادله فوق را گسسته می‌کنیم

$$(Lu)_i + a_i u_i + \sigma u_i^3 = E_0 u_i, \sum_{i=1}^{n^d} u_i^2 h^d = 1 \qquad (6)$$

که در آن

$$(lu)_i := \sum_{k=1}^{d} \frac{-u_{i+e_k} + 2u_i - u_{i-e_k}}{h^2}.$$

شبکه‌های عصبی

شبکه عصبی[1] از تعدادی واحد متصل به هم نام نورون [2] تشکیل می‌شود. هر نورون دارای یک وضعیت داخلی است که در ترکیب با داده ورودی تغییر می‌کند و توسط تابع فعال‌سازی[3] خروجی نورون را به حالت روشن یا خاموش تغیر می‌دهد. به عبارت ریاضی، هر نورون دارای اسکالرهای داخلی به نام وزن[4] و بایاس[5] و تابع فعالسازی‌ای است که به ترتیب با W و b و ϕ نمایش داده می‌شوند. و خروجی نورون در این صورت با فرض اینکه X ورودی نورون باشد به صورت $\phi(WX+b)$ خواهد بود. برای این ساختار تابعی به عنوان تابع انحراف[6] به عنوان تابع هدف تعریف می‌شود که در حالت یادگیری تحت

1 Neural network
2 Neuron
3 Activation function
4 Weight
5 Bias
6 Loss function

نظارت[1] به عنوان معیاری برای تعیین انحراف جواب‌های شبکه از جواب‌های واقعی معرفی شده به شبکه است. در این صورت، می‌توان روند یادگیری یک شبکه عصبی را معادل با یک مسئله کمینه‌سازی برای کمینه کردن میزان این تابع هدف در نظر گرفت. ابزار شبکه برای کمینه سازی این تابع هدف، تغییر وزن‌ها و بایاس‌های نورون‌های خود است. قضیه زیر که به قضیه تقریب جهانی مشهور است، شرایط استفاده از شبکه‌های عصبی برای تقریب جواب مسئله را بیان می‌کند.

قضیه. (قضیه تقریب جهانی)

۱. **(حالت نامتناهی)** فرض کنید $\varphi : \mathbb{R} \to \mathbb{R}$ یک تابع غیرثابت پیوسته بی‌کران باشد که آن را تابع فعال‌سازی می‌نامیم. فرض کنید I_m بیانگر ابر مکعب mـ بعدی $[0,1]^m$ باشد، و فضای توابع پیوسته حقیقی مقدار روی I_m با $C(I_m)$ نمایش داده شود. در این‌صورت، به ازای هر $\varepsilon > 0$ دلخواه و هرتابع $f \in C(I_m)$، ثوابت حقیقی مانند $v_i, b_i \in \mathbb{R}$ و بردارهای $w_i \in \mathbb{R}^m$ برای $i = 1, \ldots, N$ وجود دارند، به طوری‌که می‌توانیم

$$F(x) = \sum_{i=1}^{N} v_i \varphi\left(w_i^T x + b_i\right)$$

را به عنوان تقریبی از f ارائه دهیم:

$$|F(x) - f(x)| < \varepsilon$$

که در آن $x \in I_m$ است. به عبارت دیگر، توابع به شکل $F(x)$ در $C(I_m)$ چگال‌اند. این نتیجه به ازای هر زیر مجموعه فشرده دیگری از $\mathbb{R}^m$ به جای I_m نیز برقرار است.

۲. **(حالت کران‌دار)** در شبکه‌های کران‌دار، برای هرتابع انتگرال‌پذیر لبگ مانند $f : \mathbb{R}^n \to \mathbb{R}$ و

[1] Supervised learning

هر $\epsilon > 0$ یک شبکه ReLu[1] کامل $\mathcal{A}$ با عرض $d_m \leq n + 4$، به گونه ای موجود است که $F_{\mathcal{A}}$ نمایش داده شده با این شبکه در رابطه

$$\int_{\mathbb{R}^n} |f(x) - F_{\mathcal{A}}(x)| \, \mathrm{d}x < \epsilon$$

صدق نماید.

روش پیشنهادی

معادله تعیین ضریب رسانش مؤثر تنها در یک بعد، و معادله شرودینگر در دو بعد حل خواهند شد. هدف این انتخاب تأکید بر این نکته بوده است که با اندک تغییری در ساختار ورودی‌های شبکه، می‌توان از این مدل برای حل معادلات دیفرانسیل در هر بعد دلخواهی استفاده کرد. دلیل این امر نیز استفاده از محاسبات تنسوری در ساختار شبکه عصبی می‌باشد.

توجه خواننده را به این نکته جلب می‌نمائیم که روش عددی‌ای که برای ساخت پایگاه داده اولیه مورد نیاز برای شبکه عصبی استفاده می‌شود، از جهتی حائز اهمیت است؛ زیرا به میزانی که خطای ما روی داده اولیه کمتر باشد اطمینان ما از جواب شبکه عصبی بیشتر است. از طرفی، بررسی اینکه این خطا در داده‌های اولیه چگونه در شبکه عصبی منتشر می‌شود و شبکه تا چه میزان به خطا در داده های ورودی خود حساس است، از حیطه این پایان نامه خارج است. بنابراین، از آنجایی که شبکه عصبی صرفا از یک جدول داده بهره می‌برد و نسبت به اینکه این جدول داده‌ای از چه راهی بدست آمده اطلاع قبلی ندارد، در این پایان نامه فرض بر این است که روش عددی استفاده شده برای ایجاد پایگاه داده با شبکه عصبی غیر مرتبط است و بنابراین از هر روش عددی‌ای می‌توان استفاده نمود. به همین منظور از روش تفاضلات متناهی برای تولید پایگاه داده در هر دو مسئله استفاده نموده‌ایم.

[1] Rectified linear unit

برای حل معادلات با روش تفاضلات متناهی، معادله تعیین ضریب رسانش مؤثر روی یک شبکه نه نقطه‌ای متساوی‌الفاصله، با مقادیر ضرایب عدم قطعیت با توزیع نرمال $\mathcal{U}[۰/۳, ۱/۵]$، و معادله شرودینگر غیر خطی روی یک شبکه هشتاد و یک نقطه‌ای متساوی‌الفاصله (گسسته سازی نه نقطه‌ای هر کدام از ابعاد)با مقادیر ضرایب عدم قطعیت با توزیع نرمال $\mathcal{U}[۱, ۱۶]$ به تعداد نمونه‌های مورد نیاز حل می‌شوند. سپس درصدی از تکرارها (در اینجا هفتاد و پنج درصد) به عنوان داده برای مرحله آموزش و الباقی برای مرحله آزمون کنار گذاشته می‌شوند.

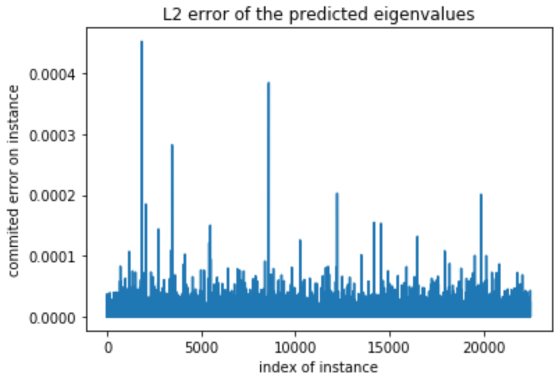
شبکه عصبی متشکل از سه بخش است. بخش اول و سوم قرینه یکدیگر و متشکل از لایه‌های پیچشی [۱] هستند که به واسطه بخش دوم که یک استخر مجموع [۲] است به هم متصل شده‌اند. ورودی این شبکه برای رسانایی مؤثر یک بردار و برای معادله شرودینگر یک ماتریس است. دقت شود که در حالت دو بعدی، قبل از لایه‌های پیچشی، ابعاد داده ورودی گسترش می‌یابد. این امر با توجه به اینکه شرط مرزی مسئله دوره‌ای است، به صورت گسترش دوره‌ای انجام می‌شود. خروجی شبکه در هر دو حالت یک اسکالر است. که در مورد ضریب رسانائی مؤثر، این اسکالر برابر ضریب رسانائی مؤثر در جهت ثابت ξ و در مورد معادله شرودینگر، برابر با سطح انرژی حالت پایه است.

شبکه پس از چندین بار مرور داده‌ها در انتها ضرایب خود را به گونه ای تنظیم می‌کند که تابع هدفی که به آن معرفی کرده‌ایم را کمینه نماید. وقتی تابع مذکور به میزان کمینه خود برسد می‌گوییم آموزش شبکه به اتمام رسیده است. از این پس می‌توانیم با خوراندن ورودی جدید به شبکه از آن برای یافتن جواب استفاده نماییم.

[1] convolutional layers
[2] sum-pooling

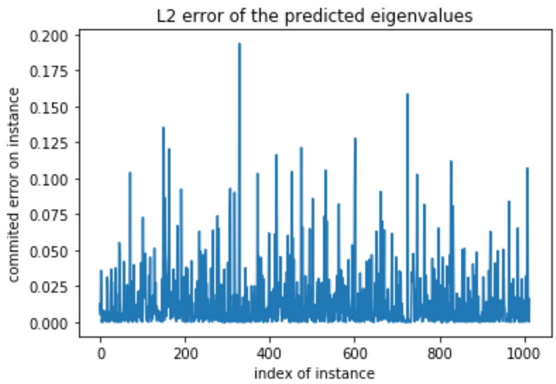
نتایج

مقادیر ضریب رسانایی مؤثر و انرژی حالت پایه به ترتیب ۰٫۷۶۸۰۰۶۵۰ و ۱۰٫۱۷۴۷۴۵۵۶ بدست آمده اند که خطای $L^۲$ به ترتیب، عبارت اند از $۱٫۰۲۱۰۰ \times ۱۰^{-۳}$ و $۷٫۲۳۵ \times ۱۰^{-۵}$. نمودار توزیع خطا بر حسب نمونه برای ضریب رسانایی مؤثر به شرح زیر است.



شکل ۱: خطای مرتکب شده روی مجموعه آزمون به تفکیک نمونه

همچنین، نمودار مشابه برای انرژی حالت پایه نیز به شرح زیر است. همان‌گونه که مشاهده می‌شود،



شکل ۲: خطای مرتکب شده روی مجموعه آزمون به تفکیک نمونه

تعداد نمونه‌های مورد استفاده قرا گرفته برای آموزش ماشین در معادله شرودینگر بسیار کمتر از معادله ضریب رسانایی مؤثر است. دلیل این امر هزینه بالای محاسباتی برای محاسبات این معادله نسبت به معادله اول بوده است. دقت کنید که در این معادله، به دلیل وجود جمله غیر خطی، مجبور به استفاده

از یک روش تکراری برای حل معادله هستیم. روش حل ما بر مبنای روش هموتوپی نیوتنی بوده است. همچنین دقت شبکه عصبی در معادله شرودینگر به مراتب کمتر از معادله ضریب رسانایی مؤثر است. دلیل این امر را می‌توان به دوعامل نسبت داد. عامل اول تعداد کمتر نمونه‌هاست که به آن اشاره شد و دلیل دوم، بزرگ بودن برچسب ها و کوچک بودن مقیاس فواصل آن‌هانسبت به اندازه خود برچسب‌ها. بهر روی، برای تعدیل این مشکل و بالا بردن دقت داده‌ها توسط روش نرمال‌سازی Z-Score نرمال شده‌اند.

نتیجه‌گیری و کارهای پیش‌رو

همان‌گونه که از نتایج مشهود است، شبکه‌های عصبی توانایی بالایی در تقریب روابط پنهان مابین داده‌ها دارند. همچنین سادگی روش، آن را به یک روش در دسترس تبدیل می‌کند. ضمن اینکه پس از طی مرحله آموزش، شبکه عصبی قادر است جواب مسئله را تقریبا به طور آنی ارائه کند. یکی از محدودیت‌های شبکه‌های عصبی در مورد اندازه مقیاس ورودی‌ها است. به این معنی که در صورتی که برچسب‌ها بسیار بزرگ باشند یا با فاصله بسیار از هم، روند یادگیری با مشکل مواجه می‌شود. همان‌گونه که مشاهده می‌شود خطا در معادله شرودینگر به علت بزرگ بودن برچسب ها در مقایسه با ضرایب عدم قطعیت بیشتر است. ما از روش نرمال سازی Z-Score برای نرمال‌سازی استفاده نمودیم.
در ادامه این پایان نامه، سؤالات زیر می‌تواند مورد بررسی بیشتر قرار گیرد

- چه روش‌های دیگری برای حل این مسئله موجود است؟
- آیا این نرمال سازی خود خطایی به مدل تحمیل می‌کند؟ کران این خطای تحمیلی چیست؟
- میزان حساسیت شبکه عصبی به خطا در داده‌های ورودی و همچنین الگوی انتشار خطا در آن به چه صورت است؟
- آیا این انتشار خطا با افزایش عمق شبکه و یا استفاده از ساختار های متفاوت اعم از توابع متفاوت برای توابع فعالسازی و همچنین نوع آرایش نورون ها ارتباطی دارد؟

واژه‌نامه فارسی به انگلیسی

شبکه عصبی پیچشی	convolutional neural-network
یادگیری عمیق	deep learning
روش تفاضلات متناهی	finite difference method
روش عناصر متناهی	finite element method
محیط ناهمگن	inhomogeneous media
شبکه عصبی	neural-network
بعد فضائی	spatial dimension
عدم قطعیت	uncertainty
مقدار سنجی عدم قطعیت	uncertainty quantification
غیر خطی	nonlinear
مشتق‌گیری اتوماتیک	automatic differentiation
هسته پردازنده گرافیکی	graphic processor unit (GPU)
شبکه رو به جلو	feedforward neural network
شبکه عصبی مصنوعی	artificial neural network
بایاس	bias
وزن	weight
تابع فعال‌سازی	activation function
تابع هزینه	loss function
نورون	neuron

پایگاه داده مستخرج از یک منبع .... dataset
پایگاه داده مستخرج از منابع متفاوت .... data set
بیش همسان سازی .... overfitting
کم همسان سازی .... underfitting
گشت در داده ها .... epoch
دسته .... batch
روش کاهش آشوبناک گرادیان .... stochastic gradient decent (SGD)