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**

شهریور ۱۳۹۹

چکیده

برای مدل‌سازی پدیده‌های واقعی با معادلات دیفرانسیل با مشتقات جزئی که شامل عدم قطعیت است، یکی از مشکلات وجود مجموعه‌ای از پدیده‌هاست 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 باربینی 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

مقدار سنجي عدم قطعيت (۱۵) در فيزيک و مهندسي اغلب شامل مطالعه معادلات ديفرانسيل با مشتقات جزئي با ميدان ضرايب تصادفي است. براي درک رفتار يک سيستم شامل عدم قطعيت، مي‌توان کميت‌هاي فيزيکي مشتق شده از معادلات ديفرانسيل توصيف کننده آن سيستم را به عنوان توابعي از ميدان ضرايب استخراج کرد. اما حتي با گسسته‌سازي مناسب روي دامنه معادله و برد متغييرهاي تصادفي، اين کار به طور ضمني به حل عددي معادله ديفرانسيل با مشتق جزئي به تعداد نمايي مي‌انجامد. يکي از روش‌هاي متداول براي مقدار سنجي عدم قطعيت روش نمونه برداري مونته کارلو است. گرچه اين روش در بسياري از موارد کاربردي است اما کميت اندازه‌گيري شده ذاتاً داراي پراش است. به‌علاوه اين روش قادر به پيدا کردن جواب‌هاي جديد در صورتي که قبلا نمونه‌گيري نشده باشند، نيست. روش گالرکين تصادفي با استفاده چند جمله‌اي‌هاي آشوب يک جواب تصادفي را روي فضاي متغييرهاي تصادفي بسط مي‌دهد و به اين طريق مسئله با بعد بالا را به تعدادي معادله ديفرانسيل با مشتقات جزئي معين تبديل مي‌کند. اين گونه روش‌ها به دقت زيادي درباره تعيين توزيع عدم قطعيت نيازمند هستند و از آن‌جا که پايه‌هاي استفاده شده مستقل از مسئله هستند، وقتي بعد متغييرهاي تصادفي بالا باشد هزينه محاسباتي بسيار زياد خواهد شد. هدف کار ما پارامتري کردن جواب يک معادله ديفرانسيل معيين به کمک شبکه‌هاي عصبي و سپس استفاده از روش‌هاي بهينه‌سازي براي يافتن جواب معادله است. در اين پايان‌نامه تابع مورد نظر براي پارامتري‌سازي روي ميدان ضرايب معادله ديفرانسيل با مشتقات جزئي تعريف شده است. در واقع ما به دنبال کاهش بعد مبتني بر نمايش شبکه عصبي براي حل معادلات ديفرانسيل با مشتقات جزئي همراه با عدم قطيت هستيم.

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 پارامترهایی هستند، معادلات با ضرایب عدم قطعیت گوییم. به طور مثال، در مورد حرکت مایعات، به طور مثال نفت، در سفره‌های زیر زمینی، برای بیان شیوه حرکت مایعات نیاز به دانستن مکان حفره‌ها 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 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 به راحتی تغییر فرم می‌دهد و هرکدام از این روش‌ها نیز به حالت خاصی از قرارگیری بیمار نیاز دارد. به طور مثال، طی MRI بیمار در حالت دمر قرار دارد ولی برای تصویربرداری فراصوت بیمار به پشت می‌خوابد. علاوه‌براین، در روش

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

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[۲] احساس می‌شود. روش‌هائی مبتنی بر روش المان‌های متناهی[۳] برای حل این مسئله ارائه شده‌اند. اما مشکل عمده این روش‌ها هزینه محاسباتی بالای آنهاست. به طور متوسط اجرای یک شبیه سازی صد و بیست دقیقه به طول می‌انجامد که مقرون به صرفه نیست. ما به دنبال ارائه روشی برای کاهش این هزینه محاسباتی با استفاده از شبکه‌های عصبی و ارائه یک مدل مختص به بیمار در زمانی قابل قبول بودیم. این امر مستلزم در نظر گرفتن ضریب کشسانی بدن بیمار، که یک ضریب عدم قطعیت است، می‌باشد. از این رو، برآن شدیم که بدنبال حل عددی معادلات دیفرانسیل (بیضوی) به کمک شبکه‌های عصبی باشیم. در این رساله، بدنبال حل عددی معادلات دیفرانسیل با مشتقات جزئی بیضوی خطی و غیر خطی ناهمگن هستیم. روشی که ما در سدد معرفی آن هستیم، یک روش کاهش بعد برای محاسبه جواب بدون نیاز به حل مستقیم معادله دیفرانسیل است. ایده، استفاده از شبکه‌های عصبی برای یادگیری نگاشتی از دامنه ضرایب عدم قطعیت به فضای جواب بر اساس مجموعه داده‌ای از قبل محاسبه شده است.

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

تعریف مسئله

در این پایان نامه، هدف ما بررسی یک مدل میانبر برای حل مسائل معادلات دیفرانسیل با مشتقات جزئی به کمک شبکه‌های عصبی بوده است. به عبارت روشن‌تر یافت 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.

یافت ضریب رسانایی مؤثر در محیط ناهمگون

معادله اول، رسانایی مؤثر در یک جهت انتخاب شده درون یک ماده غیر همگون را توسط ضریب رسانش توصیف می‌کند. ماده غیر همگون، ماده‌ای 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 $a(x)$ نمایش می‌دهیم. 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:

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

معادله دوم، معادله غیرخطی شرودینگر دو بعدی است. هدف از این معادله، یافت میزان انرژی حالت پایه الکترون با پتانسیل اولیه همراه با عدم قطعیت است. این معادله به صورت یک مسئله مقدار ویژه به صورت زیر تعریف می‌شود، که هدف ما در حل این مسئله یافتن کوچکترین مقدار ویژه آن است:

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

شبکه‌های عصبی

شبکه عصبی از تعدادی واحد متصل به هم نام نورون تشکیل می‌شود. هر نورون دارای یک وضعیت داخلی است که در ترکیب با داده ورودی تغییر می‌کند و خروجی نورون را به حالت روشن یا خاموش تغیر می‌دهد. به عبارت ریاضی، هر نورون دارای ضرایب داخلی به نام وزن و بایاس است که به ترتیب با W و b نمایش داده می‌شوند. و خروجی نورون در این صورت با فرض اینکه X ورودی نورون باشد عبارت خواهد بود از $\phi(WX + b)$. در این صورت، می‌توان روند یادگیری یک شبکه عصبی را معادل با یک مسئله کمینه‌سازی در نظر گرفت. به شبکه رابطه‌ای روی داده های خروجی میدهیم تا آن را کمینه کند و ابزار شبکه برای کمینه سازی آن رابطه، تغییر وزن‌ها و بایاس‌های نورون‌های خود است. قضیه زیر که به قضیه تقریب جهانی مشهور است، این را بیان می‌کند که می‌توان تحت شرایطی از شبکه‌های عصبی برای تقریب جواب مسئله استفاده کرد:

قضیه. (قضیه تقریب جهانی)

۱. **(حالت نامتناهی)** فرض کنید $\varphi : \mathbb{R} \to \mathbb{R}$ یک تابع غیرثابت پیوسته بی کران باشد که آن را تابع فعال سازی می‌نامیم. فرض کنید I_m بیانگر ابر مکعب mـ بعدی $[۰,۱]^m$ باشد، و فضای توابع پیوسته حقیقی مقدار روی I_m با $C(I_m)$ نمایش داده شود. در این‌صورت، به ازای هر $\varepsilon > ۰$ **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 $i = ۱, \ldots, N$ وجود دارند، به طوری‌که می‌توانیم

$$F(x) = \sum_{i=۱}^{N} v_i \varphi\left(w_i^T x + b_i\right)$$

را به عنوان تقریبی از f ارائه دهیم که عبارت است از:

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

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 محموعه فشرده دیگری از $\mathbb{R}^m$ به جای I_m برقرار است.

۲. **(حالت کران‌دار)** در شبکه‌های کران‌دار، برای هرتابع انتگرال‌پذیر لبگ مانند $f : \mathbb{R}^n \to \mathbb{R}$ و هر $\epsilon > ۰$ یک شبکه ReLu کامل $\mathcal{A}$ با عرض $d_m \leq n + ۴$، 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. ابتدا معادلات فوق را با روش تفاضلات متناهی (و یا هر روش عددی دیگری) حل نموده و یک پایگاه داده می‌سازیم. برای اینکار، معادله تعیین ضرایب رسانش مؤثر روی یک شبکه نه نقطه‌ای متساوی‌الفاصله، با مقادیر ضرایب عدم قطعیت با توزیع نرمال $\mathscr{U}[0.3, 1.5]$، 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 $\mathscr{U}[1, 16]$ 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.

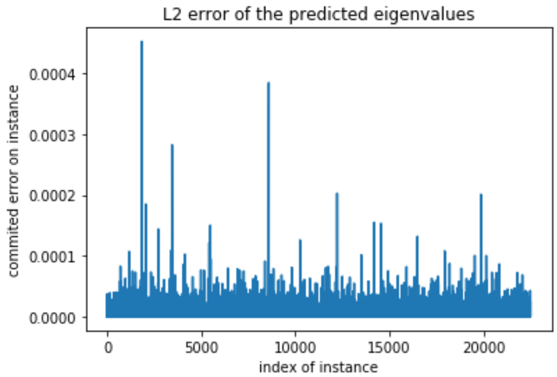
شبکه عصبی متشکل از سه بخش است. بخش اول و سوم قرینه یکدیگر و متشکل از لایه‌های پیچشی [1] اند که به واسطه بخش دوم که یک استخر مجموع [2] است به هم متصل شده‌اند. ورودی این شبکه برای رسانایی مؤثر و یک ماتریس برای معادله شرودینگر است. دقت شود که در حالت دو بعدی، قبل از لایه‌های پیچشی، ابعاد داده ورودی گسترش می‌یابد. این امر با توجه به اینکه شرط مرزی مسئله دوره‌ای است، به صورت کاشی کاری دوره‌ای انجام می‌شود. خروجی شبکه در هر دو حالت یک اسکالر است. که در مورد ضزیب رسانائی مؤثر، این اسکالر برابر ضریب رسانائی مؤثر در جهت ثابت ξ و در مورد معادله شرودینگر، برابر با سطح انرژی پایه است.

شبکه پس از چندین بار مرور داده ها در انتها ضرایب خود را به گونه ای تنظیم میکند که تابع هدفی که به آن معرفی کرده ایم 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 یافت جواب استفاده نماییم.

[1] convolutional layers
[2] sum-pooling

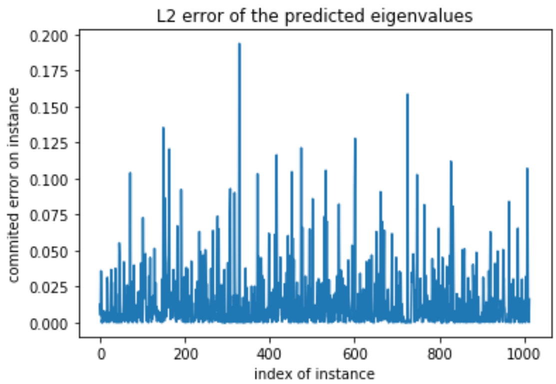
نتایج

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 است:



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

نتیجه‌گیری و کارهای پیش‌رو

همان‌گونه که از نتایج مشهود است، شبکه‌های عصبی توانایی بالایی در تقریب روابط پنان 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 ورودی‌هاست: 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 یک روش نرمال‌سازی برای نرمال‌سازی استفاده نمودیم. چه روش‌های دیگری برای حل این مسئله موجود است و آیا این نرمال سازی خود خطایی به مدل تحمیل می‌کند؟ کران این خطای تحمیلی چیست؟

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