

الحمد لله
الرحمن الرحيم



دانشکده علوم ریاضی

رساله دکتری تحقیق در عملیات - کنترل و بهینه‌سازی

حل رده‌ای از مسائل کنترل بهینه کسری مبتنی بر روشی در هوش مصنوعی

نگارنده: صفیه قاسمی صاحبی

استاد راهنما

دکتر علیرضا ناظمی

بهمن ۱۳۹۷



فرم شماره ۱۲: صورت جلسه دفاع از رساله دکتری (Ph.D)

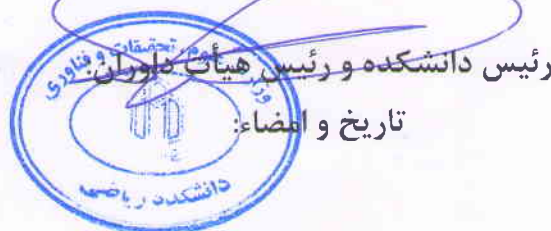
بدینوسیله گواهی می شود آقای/خانم صفیه قاسمی صاحبی دانشجوی دکتری رشته ریاضی کاربردی-کنترل و بهینه سازی به شماره دانشجویی ۹۳۰۱۰۸۵ ورودی مهر سال ۹۳ در تاریخ ۹۷/۱۱/۲۷ از رساله خود با عنوان: حل رده‌ای از مسائل کنترل بهینه کسری مبتنی بر روشی در هوش مصنوعی دفاع و با اخذ نمره ۱۹.۵۲ به درجه: عالی نائل گردید.

☐ الف) درجه عالی: نمره ۱۹-۲۰	☐ ب) درجه بسیار خوب: نمره ۱۸/۹۹ - ۱۷
☐ ج) درجه خوب: نمره ۱۶/۹۹ - ۱۵	☐ د) غیر قابل قبول و نیاز به دفاع مجدد دارد
☐ ه) رساله نیاز به اصلاحات دارد	

ردیف	هیئت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱	دکتر علیرضا ناظمی	استاد/ اساتید راهنما	دانشیار	
۲	دکتر رضا پورقلی	استاد مدعو داخلی	دانشیار	
۳	دکتر حجت احسنی طهرانی	استاد مدعو داخلی	دانشیار	
۴	دکتر محمد هادی نوری	استاد مدعو داخلی	استادیار	
۵	دکتر مهدی قوتمند	سرپرست (نماینده) تحصیلات تکمیلی دانشکده	استادیار	

مدیر محترم تحصیلات تکمیلی دانشگاه:

ضمن تأیید مراتب فوق مقرر فرمائید اقدامات لازم بعمل آید.



تاریخ و امضاء:

تقدیم بہ

یوسف گمشدہ زندیکیم، پدرم، کہ عالمانہ بہ من آموختہ بود تا چگونہ در عرصہ
زندگی، ایستادگی را تجربہ نمایم.

مادرم، دریامی بی کران فداکاری و عشق کہ وجودم برایش ہمہ رنج
است و وجودش برایم ہمہ مهر.

تشکر و قدردانی

سپاس بی‌پایان، پرودگار بی‌همتا را که فرصت علم و دانش را ارزانیم داشت و در تمام مراحل زندگیم یاری نمود. براستی که پیمودن راه دشوار زندگی جزء با اتکا به قدرت لایزال او ممکن نیست. از استاد گرامی جناب دکتر علیرضا ناظمی به خاطر زحمات بی‌دریغ و لطف‌های فراوان سپاس‌گزاری می‌کنم و بسیار قدردان زحمات ایشان هستم. از اساتید داور ارجمند که قبول زحمت مطالعه این پایان‌نامه را پذیرفته‌اند کمال امتنان را دارم. در آخر می‌فشارم دستان تک‌تک اعضای خانواده‌ام را و تشکر فراوان دارم از تمام عزیزانی که مرا در این مسیر یاری‌گر بوده‌اند.

صفیه قاسمی صاحبی

بهمن ۱۳۹۷

تعهد نامه

اینجانب صفیه قاسمی صاحبی دانشجوی دکتری رشته ریاضی کاربردی علوم ریاضی دانشگاه شاهرود، نویسنده پایان نامه با عنوان حل رده‌ای از مسائل کنترل بهینه کسری مبتنی بر روشی در هوش مصنوعی، تحت راهنمایی علیرضا ناظمی متعهد می‌شوم:

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های دیگر پژوهش‌گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان نامه، تا کنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ‌جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه صنعتی شاهرود تعلق دارد، و مقالات مستخرج با نام “دانشگاه صنعتی شاهرود” یا “Shahrood University of Technology” به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به‌دست آوردن نتایج اصلی پایان نامه تاثیرگذار بوده‌اند، در مقالات مستخرج از پایان نامه رعایت می‌گردد.
- در تمام مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافت‌های 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.

صفیه قاسمی صاحبی

بهمن ۱۳۹۷

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده) متعلق به دانشگاه صنعتی شاهرود می‌باشد. 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 نمی‌باشد.

چکیده

مسائل کنترل بهینه کسری مسائلی هستند که دستگاههای دینامیکی همراه کننده آنها دارای مشتقات و انتگرالهای مرتبه کسری هستند که این مشتقات و انتگرالها مدل‌های دقیق‌تری برای سیستم‌های موجود در طبیعت فراهم می‌کنند. پس با توجه به ضرورت حل این نوع مسائل و قابلیت بالای مدل‌های شبکه عصبی که یکی از زمینه‌های بسیار فعال و پرکاربرد در حوزه هوش مصنوعی هستند، در این رساله رده‌ای از مسائل کنترل بهینه کسری با محدودیت‌های مساوی و نامساوی را مورد مطالعه قرار می‌دهیم. برای حل این مسائل از روش‌های عددی بر پایه شبکه عصبی پرسپترون و تقریب گرانونالد-لتنیکف، شبکه عصبی بر پایه تابع میتاگ-لفلر، شبکه عصبی بر پایه سری توانی کسری و همچنین شبکه عصبی بر پایه چندجمله‌ای‌های لژاندر استفاده می‌کنیم. از طرفی برخی از مسائل کنترل بهینه کسری شامل مشتق‌های مرتبه کسری وابسته به زمان هستند که در این رساله به حل این نوع مسائل نیز می‌پردازیم.

کلمات کلیدی: مسئله کنترل بهینه کسری با مشتق مرتبه ثابت و متغیر، شرایط بهینگی، شبکه عصبی مصنوعی، بهینه‌سازی نامقید، پایداری و همگرایی.

لیست مقالات مستخرج از پایان نامه

1- Safiye Ghasemi, Alireza Nazemi, Soleiman Hosseinpour (2017), Nonlinear fractional optimal control problems with neural network and dynamic optimization schemes, Non-linear Dyn, DOI 10.1007/s11071-017-3616.

2-Safiye Ghasemi, Alireza Nazemi (2018), "A Neural Network Method Based on Mittag-Leffler Function for Solving a Class of Fractional Optimal Control Problems", Journal of Modeling and Simulation, DOI 10.22060/miscj.2018.14448.5106.

3- Safiye Ghasemi, Alireza Nazemi, A fractional power series neural network-based method for solving a class of fractional optimal control problems with equality and inequality constraints, submitted.

4- Safiye Ghasemi, Alireza Nazemi and Raziye Tajik, On fractional optimal control problems with an application in fractional chaotic systems, submitted.

Conference articles

1-Safiye Ghasemi, Alireza Nazemi (2017), Fractional Optimal Control Problems with a Feed Forward Model, The 10th International Conference of Iranian Operations Research Society, Bobolsar-Iran.

2-Safiye Ghasemi, Alireza Nazemi (2017), A method of artificial intelligence for solving fractional optimal control problems with equality and inequality constraints, 48th Annual Iranian Mathematics Conference, Bu-Alisina University, Hamedan, Iran.

3-Safiye Ghasemi, Alireza Nazemi and Soleiman Hosseinpour (2018), A New Direct Method Based on the Legendre Neural Network for Fractional Optimal Control Problem, The 2th National Seminar on Control and Optimization, Shahrood, Iran.

فهرست مطالب

ق	فهرست تصاویر
ث	فهرست جداول
۱	پیش‌نیازها از حسابان کسری و معرفی چندجمله‌ای‌های مرتبه کسری
۱	۱.۱ تاریخچه‌ای از حسابان کسری
۳	۲.۱ حسابان کسری مرتبه متغیر
۴	۳.۱ توابع پرکاربرد در حسابان کسری
۵	۴.۱ انواع مشتق و انتگرال کسری
۵	۱.۴.۱ انتگرال و مشتق ریمن-لیوویل
۶	۲.۴.۱ انتگرال و مشتق لیوویل
۷	۳.۴.۱ مشتق کسری کاپاتو
۷	۴.۴.۱ مشتق و انتگرال ریمن-لیوویل اصلاح‌شده یا جوماری
۸	۵.۴.۱ مشتق و انتگرال کسری گرانوالد-لتنیکف
۹	۶.۴.۱ مشتق کسری همدیس‌پذیر
۱۰	۷.۴.۱ مشتق و انتگرال کاپاتو-فابریسیو
۱۰	۵.۱ خواص انتگرال و مشتقات کسری
۱۲	۱.۵.۱ خواص مشتق کسری کاپاتو
۱۳	۲.۵.۱ خواص مشتق کسری همدیس‌پذیر
۱۴	۶.۱ مشتق‌های کسری مرتبه متغیر
۱۵	۷.۱ چندجمله‌ای‌های مرتبه کسری
۱۵	۱.۷.۱ چندجمله‌ای‌های لاگرانژ مرتبه کسری
۱۶	۲.۷.۱ چندجمله‌ای‌های لژاندر مرتبه کسری
۱۷	۳.۷.۱ چندجمله‌ای‌های چبیشف مرتبه کسری
۱۸	۴.۷.۱ چندجمله‌ای‌های مونتس-لژاندر

۲۱	پیش‌نیازها از شبکه عصبی و معرفی سیستم‌های دینامیکی	۲
۲۱	۱.۲ پیش‌نیازها از شبکه عصبی	
۲۲	۱.۱.۲ شبکه‌های عصبی بیولوژیکی	
۲۳	۲.۱.۲ شبکه‌های عصبی مصنوعی	
۲۷	۳.۱.۲ انواع یادگیری برای شبکه‌های عصبی	
۲۸	۴.۱.۲ انواع شبکه‌های عصبی مصنوعی	
۳۱	۲.۲ سیستم‌های دینامیکی	
۳۷	۳ مسئله کنترل بهینه و شرایط بهینگی	
۳۷	۱.۳ مسئله کنترل بهینه	
۳۸	۱.۱.۳ مسئله کنترل بهینه کلاسیک	
۴۱	۲.۱.۳ مسئله کنترل بهینه کسری و شرایط لازم بهینگی	
۴۵	۳.۱.۳ مسئله کنترل بهینه از مرتبه متغیر و شرایط بهینگی	
۴۵	۲.۳ بررسی روش‌های عددی برای مسائل کنترل بهینه کسری	
۴۹	۴ مسائل کنترل بهینه کسری با شبکه عصبی و طرح‌های بهینه‌سازی دینامیکی	
۴۹	۱.۴ مقدمه	
۵۰	۲.۴ عملگرهای گراناوالد-لتنیکف	
۵۱	۳.۴ مسئله کنترل بهینه کسری و روش حل مسئله	
۵۵	۴.۴ الگوریتم آموزش	
۵۶	۵.۴ پایداری و همگرایی	
۵۹	۶.۴ مثال‌های عددی	
	۵ حل رده‌ای از مسائل کنترل بهینه کسری با محدودیت‌های مساوی و نامساوی با	
۷۱	دو رویکرد متفاوت از شبکه‌های عصبی	
۷۲	۱.۵ بررسی شرایط لازم بهینگی	
۷۳	۲.۵ بهینه‌سازی بر پایه شبکه عصبی سری توانی کسری	
۷۴	۱.۲.۵ ساختار شبکه عصبی سری توانی کسری	
۷۵	۲.۲.۵ بهینه‌سازی و گسسته‌سازی	
۷۸	۳.۵ بهینه‌سازی با کمک شبکه عصبی بر پایه تابع میتاگ-لفلر	
۸۰	۴.۵ مثال‌های عددی	
۹۹	۶ مسئله کنترل بهینه کسری با یک کاربرد در سیستم‌های آشوبناک کسری	
۱۰۰	۱.۶ تقریب مشتق کسری ریمن-لیوویل	
۱۰۱	۲.۶ مدل‌بندی ریاضی مسئله	

۳.۶	ساختار شبکه عصبی و بهینه‌سازی	۱۰۲
۴.۶	کنترل بهینه سیستم‌های آشوبناک کسری	۱۰۵
۵.۶	مثال‌های عددی	۱۰۷

فهرست تصاویر

۲۳	 ساختار نرون	۱.۲
۲۶	 مدل ریاضی نرون عصبی مصنوعی	۲.۲
۳۰	 ساختار یک نرون با لایه ورودی، لایه پنهان و لایه خروجی	۳.۲
۳۱	 پایداری لیاپانف.	۴.۲
۳۲	 پایداری مجانبی سراسری.	۵.۲
۵۲	 ساختار شبکه عصبی پرسپترون	۱.۴
۶۰	 مسیر بهینه $x(.)$ برای $\alpha = 1$ در مثال ۱.۶.۴	۲.۴
۶۰	 کنترل بهینه $x(.)$ برای $\alpha = 1$ در مثال ۱.۶.۴	۳.۴
۶۱	 خطای تقریب تابع وضعیت برای $\alpha = 1$ در مثال ۱.۶.۴	۴.۴
۶۱	 خطای تقریب تابع کنترل برای $\alpha = 1$ در مثال ۱.۶.۴	۵.۴
۶۱	 جواب تقریبی مسیر بهینه $x(.)$ برای مقادیر مختلف α در مثال ۱.۶.۴	۶.۴
۶۲	 جواب تقریبی کنترل بهینه $u(.)$ برای مقادیر مختلف α در مثال ۱.۶.۴	۷.۴
	 جواب تقریبی مسیر بهینه $x(.)$ و کنترل بهینه $u(.)$ برای مقادیر α و $\alpha(t)$ در مثال ۱.۶.۴	۸.۴
۶۳	 جواب تقریبی مسیر بهینه $x(.)$ و کنترل بهینه $u(.)$ برای مقادیر α و $\alpha(t)$ در مثال ۱.۶.۴	۹.۴
۶۳	 جواب تقریبی مسیر بهینه $x(.)$ برای مقادیر مختلف α در مثال ۲.۶.۴	۱۰.۴
۶۴	 جواب تقریبی کنترل بهینه $u(.)$ برای مقادیر مختلف α در مثال ۲.۶.۴	۱۱.۴
۶۴	 جواب تقریبی مسیر بهینه $x(.)$ و کنترل بهینه $u(.)$ برای مقادیر مختلف $\alpha(t)$ در مثال ۲.۶.۴	۱۲.۴
۶۵	 مسیر بهینه $x(.)$ برای $\alpha = 1$ در مثال ۳.۶.۴	۱۳.۴
۶۷	 کنترل بهینه $u(.)$ برای $\alpha = 1$ در مثال ۳.۶.۴	۱۴.۴
۶۷	 خطای تقریب تابع وضعیت برای $\alpha = 1$ در مثال ۳.۶.۴	۱۵.۴
۶۸	 خطای تقریب تابع کنترل برای $\alpha = 1$ در مثال ۳.۶.۴	۱۶.۴
۶۹	 مسیر بهینه $x(.)$ برای $\alpha = 1$ در مثال ۴.۶.۴	۱۷.۴

۶۹	۱۸.۴ کنترل بهینه $u(.)$ برای $\alpha = 1$ در مثال ۴.۶.۴
۶۹	۱۹.۴ خطای تقریب تابع وضعیت برای $\alpha = 1$ در مثال ۴.۶.۴
۷۰	۲۰.۴ خطای تقریب تابع کنترل برای $\alpha = 1$ در مثال ۴.۶.۴
۷۰	۲۱.۴ مسیر بهینه جواب تقریبی $x(.)$ با مقادیر مختلف α در مثال ۴.۶.۴
۷۰	۲۲.۴ مسیر کنترل جواب تقریبی $u(.)$ با مقادیر مختلف α در مثال ۴.۶.۴
۷۴	۱.۵ نمودار بلوکی ساختار عصبی ارائه شده
۷۸	۲.۵ نمودار بلوکی ساختار شبکه عصبی ارائه شده.
۸۱	۳.۵ مسیر بهینه $x(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$
۸۱	۴.۵ کنترل بهینه $u(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$
	۵.۵ مسیر بهینه $x(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8, 0.7, 0.5$ و خطای مطلق برای $\alpha = 1$
۸۲	۶.۵ کنترل بهینه $u(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8, 0.7, 0.5$ و خطای مطلق برای $\alpha = 1$
۸۳	۷.۵ مسیر بهینه $x(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$
۸۵	۸.۵ کنترل بهینه $u(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$
	۹.۵ مسیر بهینه $x(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8, 0.7, 0.5$ و خطای مطلق برای $\alpha = 1$
۸۶	۱۰.۵ کنترل بهینه $u(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8, 0.7, 0.5$ و خطای مطلق برای $\alpha = 1$
۸۶	۱۱.۵ مسیر بهینه $x(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$
۸۷	۱۲.۵ کنترل بهینه $u(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$
۸۸	۱۳.۵ مسیر بهینه $x_1(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$
۸۹	۱۴.۵ مسیر بهینه $x_2(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$
۹۰	۱۵.۵ کنترل بهینه $u(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$
۹۰	۱۶.۵ مسیر بهینه $x(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$
۹۱	۱۷.۵ کنترل بهینه $u(.)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$
۹۲	۱۸.۵ مسیر بهینه $x(.)$ خطای مطلق برای $\alpha = 0.5$
۹۳	۱۹.۵ کنترل بهینه $u(.)$ و خطای مطلق برای $\alpha = 0.5$
۹۳	۲۰.۵ مسیر بهینه $x(.)$ و خطای مطلق برای $\alpha = 0.5$
۹۵	۲۱.۵ کنترل بهینه $u(.)$ و خطای مطلق برای $\alpha = 0.5$
۹۶	۲۲.۵ مسیر جواب‌های عددی متغیر کنترل $u(t)$ برای چند $\alpha(t)$
۹۶	۲۳.۵ مسیر جواب‌های عددی متغیر کنترل $u(t)$ برای چند $\alpha(t)$
۱۰۲	۱.۶ نمودار بلوکی شبکه عصبی لژاندر.

۲.۶	چپ (مسیر بهینه $x(t)$ برای $\alpha = 0.5$)، راست (خطای مطلق $x(t)$ برای $\alpha = 0.5$)	۱۰۸
۳.۶	چپ (کنترل بهینه $u(t)$ برای $\alpha = 0.5$)، راست (خطای مطلق $u(t)$ برای $\alpha = 0.5$)	۱۰۹
۴.۶	مسیر بهینه $x(\cdot)$ و کنترل بهینه $u(\cdot)$ با مقادیر مختلف α	۱۰۹
۵.۶	چپ (مسیر بهینه $x(t)$ برای $\alpha = 0.2, 0.3, 0.5, 0.7$)، راست (خطای مطلق $x(t)$ برای $\alpha = 0.2, 0.3, 0.5, 0.7$)	۱۱۰
۶.۶	چپ (کنترل بهینه $u(t)$ برای $\alpha = 0.2, 0.3, 0.5, 0.7$)، راست (خطای مطلق $u(t)$ برای $\alpha = 0.2, 0.3, 0.5, 0.7$)	۱۱۱
۷.۶	رفتار پایدار تابع کنترل و وضعیت برای $\alpha = 0.9$	۱۱۴
۸.۶	رفتار پایدار تابع کنترل و وضعیت برای $\alpha = 0.8$	۱۱۴
۹.۶	رفتار پایدار تابع کنترل و وضعیت برای $\alpha = 0.5$	۱۱۵

فهرست جداول

۸۳	۱.۵	مقایسه بین جواب تحلیلی و چند روش تقریبی برای مسیر $x(t)$
۸۴	۲.۵	مقایسه بین جواب تحلیلی و چند روش تقریبی برای کنترل $u(t)$
۸۶	۳.۵	مقدار تابع خطا $E(y)$ برای چند مقدار $۰ < \alpha < ۱$
۸۸	۴.۵	مقدار تابع خطا $E(y)$ برای چند مقدار $۰ < \alpha < ۱$
۸۹	۵.۵	مقدار تابع خطای $E(y)$ برای چند مقدار $۰ < \alpha < ۱$
۹۱	۶.۵	مقدار تابع خطا $E(y)$ برای چند مقدار $۰ < \alpha < ۱$
	۷.۵	مقایسه بین روش ارائه شده در [۱۰۶] با $N = ۱۵$ و روش اول ارائه شده در
۹۷		این فصل با $N = ۸$
	۸.۵	مقایسه بین روش ارائه شده در [۱۰۶] با $N = ۱۵$ و روش اول ارائه شده در
۹۷		این فصل با $N = ۸$
۱۰۸	۱.۶	مقدار تابع خطای $E(y)$ برای چند مقدار $۰ < \alpha < ۱$

فصل ۱

پیش‌نیازها از حسابان کسری و معرفی چند جمله‌ای‌های مرتبه کسری

حساب کسری یک نظریه ریاضی کلاسیک و تعمیمی از مشتق‌ها و انتگرال‌های مرتبه صحیح به مرتبه غیرصحیح است. چون مدل‌های مرتبه کسری اغلب مناسب‌تر از مدل‌های مرتبه صحیح برای توصیف سیستم‌های فیزیکی هستند، یک رشد قابل توجهی در این حوزه طی دهه‌های اخیر ایجاد شده است [۳]–[۵].

۱.۱ تاریخچه‌ای از حسابان کسری

تاریخچه نظریه مشتقات از مرتبه غیرصحیح و یا کسری به یادداشت لایبنیتز^۱ برای هوپیتال^۲ در ۳۰ سپتامبر ۱۶۹۵ باز می‌گردد. در این یادداشت معنای مشتق مرتبه $\frac{1}{p}$ مورد بحث قرار گرفت. در سال ۱۷۷۲ لاگرانژ^۳ قاعده‌ی زیر را برای عملگر دیفرانسیلی از مرتبه صحیح توسعه داد

$$\frac{d^m}{dx^m} \frac{d^n}{dx^n} = \frac{d^{m+n}}{dx^{m+n}}.$$

¹Leibniz

²Hopital

³Lagrange

اویلر^۴ در سال ۱۷۸۳ روی تصاعد اعداد کار کرد و برای اولین بار تابع گاما را به عنوان تعمیم فاکتوریل معرفی کرد. در سال ۱۸۱۲، لاپلاس^۵ یک مشتق کسری توسط مفهوم یک انتگرال تعریف کرد و در سال ۱۸۱۹ اولین نماد مشتق از مرتبه دلخواه معرفی شد. در این سال ریاضی‌دانی به نام لاکرویکس^۶ با شروع از تابع $y = x^m$ ، $m \in \mathbb{N}$ ، مشتقات کسری با مرتبه دلخواه را با تعمیم فرمول مشتق معمولی زیر مورد توجه قرار داد

$$\frac{d^n x^m}{dx^n} = \frac{m!}{(m-n)!} x^{m-n}, \quad m \geq n, \quad m, n \in \mathbb{N},$$

که تعمیم آن برای هر m و ν به صورت زیر است

$$\frac{d^\nu x^m}{dx^\nu} = \frac{\Gamma(m+1)}{\Gamma(m-\nu+1)} x^{m-\nu}.$$

فوریه در ۱۸۲۲ عملگرهای کسری را با نمایش تابع انتگرالی تابع $f(x)$ به صورت زیر معرفی کرد

$$f(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(u) du \int_{-\infty}^{\infty} \cos(t(x-u)) dt,$$

$$\frac{d^n}{dx^n} f(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(u) du \int_{-\infty}^{\infty} t^n \cos(t(x-u)) + \frac{1}{\pi} n\pi dt,$$

که n می‌تواند هم مثبت و هم منفی باشد. این اولین تعریف مناسب از مشتق کسری برای هر تابع خوش تعریف بود. تا این زمان کاربردی از مشتق و یا انتگرال کسری شناخته شده نبود، اما آبل با استفاده از این ابزار جدید به حل معادله انتگرالی به دست آمده از مسئله خم‌همزمانی پرداخت و گامی اساسی در جهت بیان کاربردهای حسابان کسری برداشت. این مسئله که اهمیت زیادی در فیزیک پیدا کرده بود شامل تعیین یک منحنی در صفحه (x, y) است به طوری که زمان لازم برای رسیدن ذره در حال لغزش به پایین‌ترین نقطه تحت جاذبه گرانشی زمین، مستقل از نقطه شروع (x_0, y_0) بر روی منحنی باشد. بعد از آبل، مطالعات وسیعی را لیوویل^۷ در حسابان کسری انجام داد. در سال ۱۸۴۷ ریمن^۸ تحقیقاتی را روی تعمیم سری تیلور انجام داد. او تعریف انتگرال کسری از مرتبه α از تابع مفروض $f(x)$ را به صورت زیر پیشنهاد کرد

$$D^{-\alpha} f(x) = \frac{1}{\Gamma(\alpha)} \int_c^x (x-t)^{\alpha-1} f(t) dt + \psi(x). \quad (1.1)$$

وجود تابع $\psi(x)$ به علت ابهام در حد پایین انتگرال است. امروزه رابطه (۱.۱) بدون تابع $\psi(x)$ رایج‌ترین تعریف از انتگرال‌گیری کسری است که انتگرال کسری ریمن-لیوویل نام گرفته است. در اوایل قرن بیستم مارچاد^۹، واتانابه^{۱۰} و ریس^{۱۱} تلاش‌هایی را در زمینه تعریف و

⁴Euler

⁵Laplace

⁶Lacroix

⁷Liouville

⁸Riemann

⁹Marchaud

¹⁰Watanabe

¹¹Riesz

تعمیم حسابان کسری انجام دادند. با گذشت زمان و پیدایش کاربردهای مختلفی از حسابان کسری تعاریف متعددی از مشتق و انتگرال کسری که متناسب با مدل‌های فیزیکی بودند بیان شد. تا نیمه قرن بیستم حسابان کسری تا حدی رشد پیدا کرد که اولین کنفرانس در سال ۱۹۷۴ فقط مربوط به نظریات و کاربردهای حسابان کسری برگزار شد.

۲.۱ حسابان کسری مرتبه متغیر

یک تعمیم جالب از نظریه حسابان کسری که اخیراً زیاد به آن توجه شده است مسائلی هستند که شامل مشتق‌های مرتبه کسری و وابسته به زمان هستند، یعنی مرتبه این مشتق‌ها ثابت نیست [۶، ۷]. بسیاری از نویسندگان دریافتند که بسیاری از فرآیندهای دینامیکی ممکن است رفتاری که مرتبه کسری متناظر با زمان یا مکان تغییر می‌کنند از خود نشان دهند. در حقیقت بعضی پدیده‌ها در فیزیک، وقتی عملگر مرتبه کسری ثابت نباشد بهتر توصیف می‌شوند مانند رفتار بعضی از فرآیندهای انتشار در پاسخ به تغییرات دما، فرآیند انتشار در یک محیط ناهمگن یا فرآیندهایی که تغییرات در محیط باعث تغییر دینامیک ذرات می‌شود [۸، ۹، ۱۰]. لورنزو^{۱۲} و هارتلی^{۱۳} (۱۹۹۸) اولین کسانی بودند که عملگر دیفرانسیل کسری مرتبه متغیر را پیشنهاد کردند [۱۱]. اینگمن^{۱۴} در سال ۲۰۰۰، [۱۲]، لورنزو و هارتلی مجدداً در سال ۲۰۰۲، [۱۳]، این کار را دنبال کردند. بعد از آن‌ها کویمبرا^{۱۵} در سال ۲۰۰۳ یک عملگر دیفرانسیل مرتبه متغیر خطی برای مدل‌بندی فیزیکی ارائه کرد [۱۴]. سان^{۱۶} و همکارانش در سال ۲۰۱۱ مشتق مرتبه صحیح، مشتق کسری مرتبه ثابت و دو نوع از مشتقات کسری مرتبه متغیر در توصیف ویژگی حافظه سیستم‌ها را مورد مطالعه قرار دادند [۱۵]. سائرسیک^{۱۷} و همکاران در سال ۲۰۱۱ یک تعریف دیگر از انتگرال-مشتق^{۱۸} مرتبه متغیر پیشنهاد کردند [۱۶].

ملاحظه ۱.۲.۱. در متن پایان‌نامه منظور از مشتق کسری، مشتق کسری از مرتبه ثابت است و منظور از کنترل بهینه کسری نیز کنترل بهینه کسری از مرتبه ثابت است. مگر مواردی که نوع مشتق و مسئله کنترل ذکر شود.

¹²Lorenzo

¹³Hartley

¹⁴Ingman

¹⁵Coimbra

¹⁶Sun

¹⁷Sierociuk

¹⁸Differ-Integral

۳.۱ توابع پرکاربرد در حسابان کسری

در این بخش به معرفی توابع خاص از جمله تابع گاما، بتا و تابع میتاگ-لفلر^{۱۹} که نقش بسیار مهمی در حسابان کسری دارند، می‌پردازیم [۱۷].

تابع گاما

تابع گاما، $\Gamma(z)$ ، با انتگرال $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$ تعریف می‌شود که در نیمه راست صفحه مختلط ($\Re(z) > 0$) همگرا است. برخی از خواص تابع گاما عبارت است از

(۱) با استفاده از روش انتگرال‌گیری جزء به جزء داریم

$$\Gamma(z+1) = \int_0^\infty e^{-t} t^z dt = \lim_{t \rightarrow \infty} [-e^{-t} t^z]_{t=0} + z \int_0^\infty e^{-t} t^{z-1} dt = z\Gamma(z),$$

که $z \in \mathbb{C} \setminus \{0, -1, -2, -3, \dots\}$. برای هر $n \in \mathbb{N}$ می‌توان نتیجه گرفت $\Gamma(n+1) = n!$. از این جاست که می‌توان تابع گاما را تعمیم مفهوم فاکتوریل در نظر گرفت.

(۲) با استفاده از خاصیت اول به راحتی می‌توان نشان داد

$$\Gamma(z) = \frac{\Gamma(z+m)}{z(z+1)\cdots(z+m-1)}, \quad -m < \Re(z) < -m+1, \quad (2.1)$$

که $m \in \mathbb{Z}^+$ و $z \neq \{0, -1, -2, \dots\}$.

(۳) تابع گاما دارای نمایش حدی زیر است

$$\Gamma(z) = \lim_{n \rightarrow \infty} \frac{n! n^z}{z(z+1)\cdots(z+n)}, \quad \Re(z) > 0.$$

تابع بتا

در بسیاری موارد، استفاده از تابعی به نام تابع بتا به جای ترکیب خاصی از مقادیر تابع گاما مناسب‌تر است. تابع بتا به صورت زیر تعریف می‌شود

$$B(z, w) = \int_0^1 t^{z-1} (1-t)^{w-1} dt, \quad (\Re(z) > 0, \Re(w) > 0).$$

لم ۱.۳.۱. [۱۷] رابطه بین تابع بتا و تابع گاما به صورت زیر است

$$B(z, w) = \frac{\Gamma(z)\Gamma(w)}{\Gamma(z+w)}, \quad z, w \in \mathbb{C}.$$

¹⁹Mittag-Leffler

تابع میتاگ-لفلر

تابع میتاگ-لفلر به افتخار ریاضی‌دان سوئدی تابع میتاگ-لفلر [۱۸، ۱۹] نامیده شده است و تابع مهمی است که در زمینه حساب کسری مورد استفاده قرار می‌گیرد. این تابع تعمیم تابع نمایی است. همان‌طور که تابع نمایی نقش مهمی در جواب معادلات دیفرانسیل مرتبه غیرصحیح ایفا می‌کند، تابع میتاگ-لفلر نیز نقش مهمی در جواب معادلات دیفرانسیل مرتبه غیرصحیح دارد. تابع میتاگ-لفلر در دو نوع تابع تک پارامتری و دو پارامتری به‌صورت زیر تعریف می‌شود

$$E_{\alpha}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + 1)}, \quad E_{\alpha, \beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + \beta)}, \quad \alpha, \beta > 0, \quad z \in \mathbb{C}.$$

چند حالت خاص از توابع تک پارامتری و دو پارامتری در زیر بیان شده است

$$\begin{aligned} E_1(z) &= \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(k+1)} = \sum_{k=0}^{\infty} \frac{z^k}{k!} = e^z, \\ E_{1,1}(-z^2) &= \cos(z), \\ E_{1,2}(z) &= \frac{e^z - 1}{z}, \\ E_{2,1}(z^2) &= \sum_{k=0}^{\infty} \frac{z^{2k}}{\Gamma(2k+1)} = \sum_{k=0}^{\infty} \frac{z^{2k}}{(2k)!} = \cosh(z). \end{aligned}$$

۴.۱ انواع مشتق و انتگرال کسری

مفهوم انتگرال مرتبه کسری از اولین مفاهیمی است که در مبحث حسابان کسری درباره آن صحبت می‌شود. تعریف انتگرال مرتبه کسری، در واقع تعمیمی از فرمول کشی^{۲۰} برای محاسبه انتگرال‌های مرتبه صحیح است. از طرف دیگر یکی از مهم‌ترین مفاهیم در سیستم‌های مرتبه کسری، تعریف مشتق است که برخلاف انتگرال مرتبه کسری می‌توان تعریف‌های متفاوتی از آن ارائه نمود [۱۷].

۱.۴.۱ انتگرال و مشتق ریمن-لیوویل

فرض کنید $a \in \mathbb{R}$ عددی روی محور حقیقی $\mathbb{R}$ باشد. تابع $f(t)$ که به ازای $t > a$ تعریف شده است را در نظر بگیرید. انتگرال مرتبه n تابع $f(t)$ طبق فرمول کشی برابر است با

$$\begin{aligned} {}_a I_t^n f(t) &= \int_a^t dt_1 \int_a^{t_1} dt_2 \cdots \int_a^{t_{n-1}} f(t_n) dt_n \\ &= \frac{1}{(n-1)!} \int_a^t (t-\tau)^{n-1} f(\tau) d\tau, \quad n \in \mathbb{N}. \end{aligned} \quad (3.1)$$

از طرفی با توجه به تعریف تابع گاما، (۳.۱) را می‌توان به‌صورت زیر نوشت

$${}_a I_t^n f(t) = \frac{1}{\Gamma(n)} \int_a^t (t-\tau)^{n-1} f(\tau) d\tau, \quad n \in \mathbb{N}.$$

²⁰Cauchy

با الهام گرفتن از فرمول کشی و جانشینی عدد حقیقی $\alpha > 0$ به جای عدد صحیح n ، ${}_a I_t^\alpha f(t)$ و ${}_t I_b^\alpha f(t)$ به صورت زیر تعریف می‌شوند

$$\begin{aligned} {}_a I_t^\alpha f(t) &= \frac{1}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} f(\tau) d\tau, \quad (t > a), \\ {}_t I_b^\alpha f(t) &= \frac{1}{\Gamma(\alpha)} \int_t^b (\tau-t)^{\alpha-1} f(\tau) d\tau, \quad (t < b). \end{aligned} \quad (4.1)$$

این انتگرال‌ها به ترتیب انتگرال‌های چپ و راست ریمن-لیوویل نامیده می‌شوند [۲۰].

قضیه ۱.۴.۱. [۲۱] فرض کنید $f \in L_1[a, b]$ و $\alpha > 0$. در این صورت ${}_a I_t^\alpha f(t)$ تقریباً همه جا روی $t \in [a, b]$ وجود دارد و متعلق به $L_1[a, b]$ است.

تعریف ۱.۴.۱. فضای $AC^n(\Omega)$ عبارت است از مجموعه تمام توابعی مانند f که روی بازه‌ی Ω تا مرتبه $(n-1)$ ام، مشتقات پیوسته دارند و مشتق $(n-1)$ ام آن‌ها مطلقاً پیوسته است. به عبارت دیگر تابع $g \in L_1[0, 1]$ موجود است به طوری که

$$f^{n-1}(t) = f^{(n-1)}(0) + \int_0^t g(\tau) d\tau.$$

فرض کنید $n-1 \leq \alpha < n$ و $f \in AC^n([a, b])$ که $n = [\mathcal{R}(\alpha)] + 1$. علامت $[\mathcal{R}(\alpha)]$ قسمت صحیح $\mathcal{R}(\alpha)$ تعریف می‌شود. مشتقات چپ و راست ریمن-لیوویل از مرتبه α که $\mathcal{R}(\alpha) \geq 0$ به صورت زیر تعریف می‌شوند

$$\begin{aligned} {}_a D_t^\alpha f(t) &= \left(\frac{d}{dt}\right)^n ({}_a I_t^{n-\alpha} f)(t) \\ &= \frac{1}{\Gamma(n-\alpha)} \left(\frac{d}{dt}\right)^n \int_a^t (t-\tau)^{n-\alpha-1} f(\tau) d\tau, \quad (t > a), \\ {}_t D_b^\alpha f(t) &= \left(-\frac{d}{dt}\right)^n ({}_t I_b^{n-\alpha} f)(t) \\ &= \frac{1}{\Gamma(n-\alpha)} \left(-\frac{d}{dt}\right)^n \int_t^b (\tau-t)^{n-\alpha-1} f(\tau) d\tau, \quad (t < b). \end{aligned} \quad (5.1)$$

۲.۴.۱ انتگرال و مشتق لیوویل

در این قسمت مشتق و انتگرال کسری لیوویل روی $\mathbb{R}^+$ بررسی می‌شود. مشتق و انتگرال کسری ریمن-لیوویل روی بازه متناهی $[a, b]$ تعریف می‌شوند که قابل تعمیم به بازه $\mathbb{R}^+$ هستند. انتگرال‌های کسری لیوویل متناظر با (۴.۱) در بازه $\mathbb{R}^+$ عبارت است از

$$\begin{aligned} {}_a^L I_t^\alpha f(t) &= \frac{1}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} f(\tau) d\tau, \quad (t > 0, \mathcal{R}(\alpha) > 0), \\ {}_t^L I_\infty^\alpha f(t) &= \frac{1}{\Gamma(\alpha)} \int_t^\infty (\tau-t)^{\alpha-1} f(\tau) d\tau, \quad (t > 0, \mathcal{R}(\alpha) > 0). \end{aligned} \quad (6.1)$$

مشتقات کسری متناظر با (۵.۱) به ازای $\mathcal{R}(\alpha) \geq 0$ به صورت

$$\begin{aligned} {}_0^L D_t^\alpha f(t) &= \left(\frac{d}{dt}\right)^n ({}_0^L I_t^{n-\alpha} f)(t) \\ &= \frac{1}{\Gamma(n-\alpha)} \left(\frac{d}{dt}\right)^n \int_0^t (t-\tau)^{n-\alpha-1} f(\tau) d\tau, \quad (n = [\mathcal{R}(\alpha)] + 1; t > 0), \\ {}_t^L D_\infty^\alpha f(t) &= \left(-\frac{d}{dt}\right)^n ({}_t^L I_\infty^{n-\alpha} f)(t) \\ &= \frac{1}{\Gamma(n-\alpha)} \left(-\frac{d}{dt}\right)^n \int_t^\infty (\tau-t)^{n-\alpha-1} f(\tau) d\tau, \quad (n = [\mathcal{R}(\alpha)] + 1; t > 0), \end{aligned}$$

تعریف می‌شوند. به‌طور مشابه می‌توان مشتق و انتگرال لیوویل را روی محور حقیقی $\mathbb{R}$ تعریف کرد. برای جزئیات بیشتر [۲۰] را ببینید.

۳.۴.۱ مشتق کسری کاپاتو

ایده تعریف مشتق کاپاتو بسیار مشابه با همان ایده‌ای است که در تعریف مشتق ریمن-لیوویل به کار گرفته شد. در این تعریف مانند تعریف ریمن-لیوویل، مشتق بر اساس مفهوم انتگرال مرتبه کسری توسعه پیدا می‌کند. تفاوت این دو تعریف در واقع در توالی عملگرهای مشتق مرتبه صحیح و عملگر انتگرال مرتبه کسری است. در مشتق کسری کاپاتو برعکس مشتق کسری ریمن-لیوویل ابتدا مشتق مرتبه صحیح از تابع گرفته شده و سپس انتگرال‌گیری مرتبه کسری انجام می‌شود. به بیان ریاضی اگر $\alpha \in \mathbb{R}^+$ عدد مثبت ناصحیحی باشد و $n = [\alpha] + 1$ مشتق کسری کاپاتو از مرتبه α به‌صورت زیر تعریف می‌شود

$${}_a^C D_t^\alpha f(t) = {}_a I_t^{n-\alpha} \{f^{(n)}(t)\}.$$

بنابراین مشتق کسری چپ و راست کاپاتو به ترتیب عبارت است از

$$\begin{aligned} {}_a^C D_t^\alpha f(t) &= \frac{1}{\Gamma(n-\alpha)} \int_a^t (t-\tau)^{n-\alpha-1} f^{(n)}(\tau) d\tau, \\ {}_t^C D_b^\alpha f(t) &= \frac{1}{\Gamma(n-\alpha)} \int_t^b (\tau-t)^{n-\alpha-1} f^{(n)}(\tau) d\tau, \end{aligned}$$

که $n = [\mathcal{R}(\alpha)] + 1$ و $f \in AC^n[a, b]$.

۴.۴.۱ مشتق و انتگرال ریمن-لیوویل اصلاح‌شده یا جوماری

مشتقات کسری ریمن-لیوویل و کاپاتو به دلیل سادگی تعاریف و منطبق بودن آن‌ها با حالت صحیح (جز در موارد خاص) از جمله مهم‌ترین و پرکاربردترین تعاریف از مشتقات کسری هستند. البته این تعاریف دارای ایراداتی نیز هستند. ایراد اساسی مشتق ریمن-لیوویل این است که مشتق تابع ثابت در آن صفر نیست. در واقع اگر $f(t) = C$ تابع ثابت باشد و $0 < \alpha < 1$ داریم

$${}_a D_t^\alpha f(t) = \frac{C}{\Gamma(1-\alpha)} (t-a)^{-\alpha}.$$

همچنین برای محاسبه مشتق کاپاتو از مرتبه $0 < \alpha < 1$ لازم است که تابع f دارای مشتق مرتبه اول کلاسیک باشد. برای این منظور جوماری^{۲۱} [۲۲] تعریف جدیدی از مشتق و انتگرال کسری ارائه داد. انتگرال جوماری به‌صورت زیر تعریف می‌شود

$${}_a^J I_t^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} (f(\tau) - f(a)) d\tau, \quad \alpha > 0.$$

²¹ Jumarie

به ازای $0 < \alpha < 1$ مشتق کسری جوماری به صورت

$${}_a^J D_t^\alpha f(t) = D_a^J I_t^{\alpha-1} f(t) = \frac{1}{\Gamma(1-\alpha)} \left(\frac{d}{dt} \right) \int_a^t (t-\tau)^{\alpha-1} (f(\tau) - f(a)) d\tau,$$

و به ازای $n-1 \leq \alpha < n$ و $n \geq 1$ به صورت

$${}_a^J D_t^\alpha f(t) = D_a^n I_t^{\alpha-n} f(t),$$

محاسبه می‌شود. با توجه به تعریف ارائه شده، مشتق کسری جوماری از عدد ثابت صفر است و برای محاسبه مشتق کسری به ازای $0 < \alpha < 1$ لزومی به وجود مشتق مرتبه اول تابع f نیست. تبدیل لاپلاس از مشتق جوماری به صورت زیر است

$$\mathcal{L}[{}_a^J D_t^\alpha f(t)](s) = s^\alpha \mathcal{L}[f(t)](s) - s^{\alpha-1} f(0), \quad 0 < \alpha < 1.$$

جوماری همچنین بسط تیلور یک تابع بر حسب مشتق کسری جدید خود را با لم زیر بیان کرد

لم ۱.۴.۱. فرض کنید تابع پیوسته $f: \mathbb{R}^+ \rightarrow \mathbb{R}$ دارای مشتقات کسری از مرتبه $k\alpha$ برای مقدار مثبت k و مقدار دلخواه α ، $0 < \alpha < 1$ باشد، آنگاه رابطه زیر برقرار است

$$f(x+h) = \sum_{k=0}^{\infty} \frac{h^{k\alpha}}{\Gamma(1+\alpha k)} {}^J D^{k\alpha} f(x).$$

۵.۴.۱ مشتق و انتگرال کسری گرانوالد-لتنیف

فرض کنید f تابعی پیوسته باشد. در این صورت مشتق مرتبه اول تابع f به صورت زیر تعریف می‌شود

$$f'(t) = \lim_{h \rightarrow 0} \frac{f(t) - f(t-h)}{h}.$$

با به کارگیری تعریف بالا، مشتق مرتبه دوم به صورت زیر به دست می‌آید

$$f''(t) = \lim_{h \rightarrow 0} \frac{f'(t) - f'(t-h)}{h} = \lim_{h \rightarrow 0} \frac{f(t) - 2f(t-h) + f(t-2h)}{h^2}.$$

در حالت کلی برای مشتق مرتبه صحیح و مثبت n ، رابطه زیر را داریم

$$f^{(n)}(t) = \lim_{h \rightarrow 0} \frac{1}{h^n} \sum_{r=0}^n (-1)^r \binom{n}{r} f(t-rh), \quad n \in \mathbb{N}.$$

بنابراین می‌توان مشتق کسری مرتبه α را به صورت زیر تعمیم داد

$${}_a^G D_t^\alpha f(t) = \lim_{h \rightarrow 0} \frac{1}{h^\alpha} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(t-kh), \quad \alpha > 0, \quad (7.1)$$

که $nh = t - a$ ، $a \leq t \leq b$ و $[a] = n + 1$ می‌توان با جا به جایی $-\alpha$ و α در رابطه (۷.۱)، انتگرال کسری گرانوالد-لتنیکف از مرتبه α را تعریف کرد. یعنی با فرض $nh = t - a$ داریم

$$\begin{aligned} {}^G I_t^\alpha f(t) &= \lim_{h \rightarrow 0} h^\alpha \sum_{k=0}^n (-1)^k \binom{-\alpha}{k} f(t - kh), \\ &= \lim_{h \rightarrow 0} \frac{h^\alpha}{\Gamma(\alpha)} \sum_{k=0}^n \frac{\Gamma(k + \alpha)}{\Gamma(k + 1)} f(t - kh), \quad \alpha > 0. \end{aligned}$$

قضیه ۲.۴.۱. [۲۳] فرض کنید تابع $f(t)$ دارای مشتقات مرتبه اول تا مرتبه $n - 1$ پیوسته بوده و $f^{(n)}(t)$ انتگرال پذیر باشد. همچنین فرض کنید $n - 1 < \alpha < n$ ، $n \in \mathbb{N}$. در این صورت مشتق مرتبه α با تعریف گرانوالد-لتنیکف تابع $f(t)$ با مشتق مرتبه α ریمن-لیوویل این تابع برابر خواهد بود، یعنی

$${}^G D_t^\alpha f(t) = {}_a D_t^\alpha f(t).$$

با فرض برقراری شرایط مطرح شده در قضیه ۲.۴.۱ تعاریف گرانوالد-لتنیکف و ریمن-لیوویل معادل بوده و بنابراین خواص مشتق مرتبه کسری ریمن-لیوویل برای مرتبه کسری گرانوالد-لتنیکف نیز برقرار است. با توجه به ماهیت تعریف مشتق کسری گرانوالد-لتنیکف، از این تعریف در روش‌های عددی محاسبه مشتقات مرتبه کسری و همچنین گسسته‌سازی معادلات دیفرانسیل مرتبه کسری استفاده فراوانی می‌شود. همچنین برای تقریب مشتق کسری (۷.۱) در روش‌های عددی، از این تعریف با مجموع متناهی استفاده می‌شود.

تقریباً در تمامی تعاریف مشتقات کسری مشهور موجود، قاعده حاصل ضرب و قاعده زنجیره‌ای در محاسبه مشتق ترکیب توابع برقرار نیست. البته جوماری در [۲۴، ۲۵، ۲۶] روابطی را برای مشتق ریمن-لیوویل اصلاح شده به صورت تقریبی در این زمینه بیان و اثبات نمود اما پس از آن چند مثال در نقض برقراری روابط مطرح شده توسط وی بیان گردید. یکی از تعاریفی که در راستای برقراری این خواص مهم و پرکاربرد مطرح شده است، مشتق کسری همدیس پذیر^{۲۲} است که تعریف آن در زیر آمده است.

۶.۴.۱ مشتق کسری همدیس‌پذیر

این تعریف از مشتق کسری به صورت حدی تعریف می‌شود و به دلیل سادگی و سازگاری با بعضی مدل‌ها در فیزیک و مهندسی کاربردهای فراوانی پیدا کرده است. مشتق کسری چپ همدیس‌پذیر به صورت زیر تعریف می‌شود

$$T_a^\alpha f(t) = \lim_{\varepsilon \rightarrow 0} \frac{f(t + \varepsilon(t - a)^{1-\alpha}) - f(t)}{\varepsilon}.$$

²²Conformable Fractional Derivative

اگر حد فوق موجود باشد، آن گاه تابع از چپ مشتق پذیر است. مشتق کسری راست همدیس پذیر به صورت زیر تعریف می‌شود

$${}_b T^\alpha f(t) = \lim_{\varepsilon \rightarrow 0^+} \frac{f(t + \varepsilon(b-t)^{1-\alpha}) - f(t)}{\varepsilon}.$$

اگر حد فوق موجود باشد، آن گاه تابع از راست مشتق پذیر است. نکته مهم در مورد مشتق همدیس پذیر آن است که به ازای $\alpha = 1$ با تعریف کلاسیک مشتق صحیح برابر است. لازم به ذکر است که این تعریف از مشتق چون به صورت انتگرال بیان نمی‌شود خاصیت حافظه دار بودن را دارا نیست.

۷.۴.۱ مشتق و انتگرال کاپاتو- فابریسیو

تعریفی که کاپاتو از مشتق کسری ارائه داد به صورت زیر است

$${}_a^C D_t^\alpha f(t) = \frac{1}{\Gamma(1-\alpha)} \int_a^t (t-\tau)^{-\alpha} f'(\tau) d\tau, \quad 0 < \alpha < 1. \quad (۸.۱)$$

این تعریف دارای یک نقطه تکیه در $t = \tau$ است. برای رفع این مشکل در (۸.۱) کاپاتو و فابریسیو^{۲۳} تعریف جدیدی را ارائه دادند [۲۷]. با تغییر هسته $(t-\tau)^{-\alpha}$ به تابع $\exp(-\frac{\alpha}{1-\alpha}t)$ و $\frac{1}{\Gamma(1-\alpha)}$ به $\frac{M(\alpha)}{1-\alpha}$ تعریف جدید از مشتق کسری را می‌توان به صورت زیر نوشت

$${}_a^{CF} D_t^\alpha f(t) = \frac{M(\alpha)}{1-\alpha} \int_a^t \exp(-\frac{\alpha(t-\tau)}{1-\alpha}) f'(\tau) d\tau, \quad 0 < \alpha < 1,$$

که $M(\alpha)$ تابع نرمال شده^{۲۴} بر حسب α است به طوری که $M(0) = M(1) = 1$. این تعریف مشتق چپ کاپاتو- فابریسیو است. مشتق راست کاپاتو- فابریسیو به صورت زیر تعریف می‌شود

$${}_t^{CF} D_b^\alpha f(t) = -\frac{M(\alpha)}{1-\alpha} \int_t^b \exp(-\frac{\alpha(\tau-t)}{1-\alpha}) f'(\tau) d\tau, \quad 0 < \alpha < 1.$$

طبق تعریف جدید واضح است که اگر f تابعی ثابت باشد، مشابه تعریف کاپاتو ${}_a^{CF} D_t^\alpha f = 0$. تفاوت اصلی بین مشتق جدید و مشتق کاپاتو این است که مشتق جدید در نقطه $t = z$ دارای نقطه تکیه نیست. از طرفی می‌دانید که تبدیلات لاپلاس نقش مهمی را در مطالعه معادلات دیفرانسیل ایفا می‌کنند. تبدیل لاپلاس در مورد تعریف جدید به صورت زیر است

$$\mathcal{L}[{}_a^{CF} D_t^\alpha f(t)](s) = \frac{(2-\alpha)M(\alpha)}{2(s+\alpha(1-s))} (s\mathcal{L}[f(t)](s) - f(0)), \quad s > 0.$$

۵.۱ خواص انتگرال و مشتقات کسری

در این بخش برخی از خواص مهم انتگرال و مشتقات ذکر شده در بخش قبل و همچنین قضایای مورد استفاده در این رساله بیان می‌شود.

²³Fabrizio

²⁴Normalization Function

خواص مشتق و انتگرال مرتبه کسری ریمن-لیوویل

انتگرال و مشتق مرتبه کسری ریمن-لیوویل عملگر خطی هستند. یعنی برای تمام توابع $f(t)$ و $g(t)$ و تمام اسکالرهای حقیقی λ و γ می توان نشان داد

$$I^\alpha(\lambda f(t) + \gamma g(t)) = \lambda I^\alpha f(t) + \gamma I^\alpha g(t),$$

$$D^\alpha(\lambda f(t) + \gamma g(t)) = \lambda D^\alpha f(t) + \gamma D^\alpha g(t).$$

برای انتگرال مرتبه کسری داریم

$${}_a I_t^\alpha {}_a I_t^\beta f(t) = {}_a I_t^\beta {}_a I_t^\alpha f(t) = {}_a I_t^{\alpha+\beta} f(t), \quad \alpha, \beta \in \mathbb{R}^+ \cup \{0\},$$

$${}_t I_b^\alpha {}_t I_b^\beta f(t) = {}_t I_b^\beta {}_t I_b^\alpha f(t) = {}_t I_b^{\alpha+\beta} f(t), \quad \alpha, \beta \in \mathbb{R}^+ \cup \{0\}.$$

این تساوی بدان معناست که عملگر انتگرال مرتبه کسری خواص جمع پذیری^{۲۵} و جابجایی پذیری^{۲۶} را داراست. اما عملگرهای مشتق مرتبه کسری ریمن-لیوویل در حالت کلی جابجایی پذیر و جمع پذیر نیستند. یعنی تساوی های

$${}_a D_t^\alpha ({}_a D_t^\beta f(t)) = {}_a D_t^\beta ({}_a D_t^\alpha f(t)),$$

$${}_a D_t^\alpha ({}_a D_t^\beta f(t)) = {}_a D_t^{\alpha+\beta} f(t), \quad (9.1)$$

در حالت کلی برقرار نیستند. از طرفی با توجه به اینکه

$${}_a D_t^\alpha ({}_a D_t^\beta f(t)) = {}_a D_t^\beta ({}_a D_t^\alpha f(t)) + \sum_{k=0}^{\beta-1} \frac{f^{(k)}(a)(t-a)^{-\alpha-\beta+k}}{\Gamma(-\alpha-\beta+k+1)},$$

می توان نشان داد که یک شرط کافی برای برقراری تساوی های (۹.۱) این است که برای

$$k = 0, 1, \dots, r-1, \quad r = \max\{[\alpha] + 1, [\beta] + 1\},$$

داشته باشیم $f^{(k)}(a) = 0$.

قضیه ۱.۵.۱ [۲۰] عملگر مشتق کسری ریمن-لیوویل از مرتبه α معکوس چپ عملگر انتگرال کسری از همان مرتبه α است. یعنی

$${}_a D_t^\alpha ({}_a I_t^\beta f(t)) = f(t), \quad {}_t D_b^\alpha ({}_t I_b^\beta f(t)) = f(t).$$

اما عملگر مشتق کسری ریمن-لیوویل یک معکوس راست عملگر انتگرال کسری ریمن-لیوویل نیست. یعنی

$${}_a I_t^\alpha ({}_a D_t^\alpha f(t)) = f(t) - \sum_{j=1}^n [{}_a D_t^{\alpha-j} f(t)]_{t=a} \frac{(t-a)^{\alpha-j}}{\Gamma(\alpha-j+1)},$$

$${}_t I_b^\alpha ({}_t D_b^\alpha f(t)) = f(t) - \sum_{j=1}^n (-1)^{n-j} [{}_t D_b^{\alpha-j} f(t)]_{t=b} \frac{(b-t)^{\alpha-j}}{\Gamma(\alpha-j+1)}.$$

²⁵Additivity Property

²⁶Commutative Property

قواعد اساسی انتگرال و مشتق مرتبه کسری ریمن-لیوویل

الف- فرض کنید $\beta \in \mathbb{C}$ ، $\beta > 0$ و $\mathbb{R} \geq 0$ ، آن‌گاه

$$\begin{aligned}({}_a I_t^\alpha (t-a)^{\beta-1})(t) &= \frac{\Gamma(\beta)}{\Gamma(\beta+\alpha)} (t-a)^{\beta+\alpha-1}, \quad (\Re(\alpha) > 0), \\({}_a D_t^\alpha (t-a)^{\beta-1})(t) &= \frac{\Gamma(\beta)}{\Gamma(\beta-\alpha)} (t-a)^{\beta+\alpha-1}, \quad (\Re(\alpha) \geq 0),\end{aligned}$$

و

$$\begin{aligned}({}_b I_b^\alpha (b-t)^{\beta-1})(t) &= \frac{\Gamma(\beta)}{\Gamma(\beta+\alpha)} (b-t)^{\beta+\alpha-1}, \quad (\Re(\alpha) > 0), \\({}_b D_b^\alpha (b-t)^{\beta-1})(t) &= \frac{\Gamma(\beta)}{\Gamma(\beta-\alpha)} (b-t)^{\beta+\alpha-1}, \quad (\Re(\alpha) \geq 0).\end{aligned}$$

در حالت خاص اگر $\beta = 1$ و $\Re(\alpha) \geq 0$ ، آن‌گاه در حالت کلی مشتق کسری ریمن-لیوویل از تابع ثابت صفر نیست. یعنی

$$({}_a D_t^\alpha 1)(t) = \frac{(t-a)^{-\alpha}}{\Gamma(1-\alpha)}, \quad ({}_b D_b^\alpha 1)(t) = \frac{(b-t)^{-\alpha}}{\Gamma(1-\alpha)}.$$

ب- فرض کنید $c \in \mathbb{R}$ ، آن‌گاه

$$\begin{aligned}I^\alpha e^{ct} &= t^\alpha E_{1,\alpha+1}(ct), & {}_a D_t^\alpha (e^{ct}) &= t^{-\alpha} E_{1,1-\alpha}(ct), \\I^\alpha (\sin(ct)) &= ct^{\alpha+1} E_{\gamma,\alpha+\gamma}(-(ct)^\gamma), & {}_a D_t^\alpha (\sin(ct)) &= ct^{1-\alpha} E_{\gamma,\gamma-\alpha}((ct)^\gamma), \\I^\alpha (\cos(ct)) &= t^{\alpha+1} E_{\gamma,\alpha+1}(-(ct)^\gamma), & {}_a D_t^\alpha (\cos(ct)) &= t^{-\alpha} E_{\gamma,1-\alpha}(-(ct)^\gamma).\end{aligned}$$

۱.۵.۱ خواص مشتق کسری کاپاتو

مشتق مرتبه کسری کاپاتو یک عملگر خطی است، یعنی برای تمام توابع $f(t)$ و $g(t)$ و تمام اسکالرهایی حقیقی λ و γ می‌توان نشان داد

$${}^C D^\alpha (\lambda f(t) + \gamma g(t)) = \lambda {}^C D^\alpha f(t) + \gamma {}^C D^\alpha g(t).$$

برخلاف مشتق کسری ریمن-لیوویل مشتق کسری کاپاتو از عدد ثابت صفر است، یعنی برای هر عدد ثابت C داریم ${}^C D^\alpha C = 0$.

مشتق کسری کاپاتو جابجایی‌پذیر است یعنی

$${}_a^C D_t^m {}_a^C D_t^\alpha f(t) = {}_a^C D_t^\alpha {}_a^C D_t^m f(t) = {}_a^C D_t^{m+\alpha} f(t), \quad n-1 < \alpha < n,$$

اگر و تنها اگر

$$f^{(k)}(a) = 0, \quad (k = n, n+1, \dots, m).$$

اگر $\alpha > 0$ و $f(t) \in L_p(a, b)$ ، $(1 \leq p \leq \infty)$ ، آن‌گاه

$${}_a^C D_t^\alpha ({}_a I_t^\alpha f(t)) = f(t), \quad {}_b^C D_b^\alpha ({}_b I_b^\alpha f(t)) = f(t).$$

این نشان می‌دهد عملگر مشتق کسری ریمن-لیوویل معکوس چپ برای عملگر انتگرال کسری ریمن-لیوویل از همان مرتبه است. اما عملگر مشتق کسری ریمن-لیوویل معکوس راست عملگر انتگرال کسری ریمن-لیوویل نیست، یعنی

$$\begin{aligned} {}_a I_t^{\alpha C} D_t^{\alpha} f(t) &= f(t) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (t-a)^k, \\ {}_t I_b^{\alpha C} D_b^{\alpha} f(t) &= f(t) - \sum_{k=0}^{n-1} \frac{(-1)^k f^{(k)}(b)}{k!} (b-t)^k. \end{aligned}$$

رابطه بین مشتق‌های کسری ریمن-لیوویل و کاپاتو به صورت زیر بیان می‌شود

$$\begin{aligned} {}_a^C D_t^{\alpha} f(t) &= {}_a D_t^{\alpha} f(t) - \sum_{k=0}^{[\alpha]-1} \frac{f^{(k)}(a^+)}{k - \alpha + 1} (t-a)^{k-\alpha}, \\ {}_t^C D_b^{\alpha} f(t) &= {}_t D_b^{\alpha} f(t) - \sum_{k=0}^{[\alpha]-1} \frac{(-1)^k f^{(k)}(b^-)}{k - \alpha + 1} (b-t)^{k-\alpha}. \end{aligned} \quad (10.1)$$

قواعد اساسی مشتق مرتبه کسری کاپاتو

الف- فرض کنید $\beta \in \mathbb{C}$ ، $\beta > 0$ و $\beta \geq 0$ ، آن‌گاه

$$\begin{aligned} ({}_a^C D_t^{\alpha} (t-a)^{\beta-1})(t) &= \frac{\Gamma(\beta)}{\Gamma(\beta-\alpha)} (t-a)^{\beta-\alpha-1}, \quad (\mathcal{R}(\alpha) \geq 0), \\ ({}_t^C D_b^{\alpha} (b-t)^{\beta-1})(t) &= \frac{\Gamma(\beta)}{\Gamma(\beta-\alpha)} (b-t)^{\beta-\alpha-1}, \quad (\mathcal{R}(\alpha) \geq 0), \end{aligned} \quad (11.1)$$

ب- فرض کنید $c \in \mathbb{R}$ آن‌گاه

$$\begin{aligned} {}_a^C D_t^{\alpha} (e^{ct}) &= c^n t^{n-\alpha} E_{\lambda, n-\alpha+1}(ct), \\ {}_a^C D_t^{\alpha} (\sin(ct)) &= -\frac{1}{i} i (ic)^n t^{n-\alpha} [E_{\lambda, n-\alpha+1}(ict) - (-1)^n E_{\lambda, n-\alpha+1}(-ict)], \\ {}_a^C D_t^{\alpha} (\cos(ct)) &= \frac{1}{i} i (ic)^n t^{n-\alpha} [E_{\lambda, n-\alpha+1}(ict) + (-1)^n E_{\lambda, n-\alpha+1}(-ict)]. \end{aligned}$$

۲.۵.۱ خواص مشتق کسری همدیس‌پذیر

لم ۱.۵.۱ [۲۸] فرض کنید $f : [0, \infty) \rightarrow \mathbb{R}$ در نقطه $t_0 > 0$ مشتق‌پذیر از مرتبه $\alpha \in (0, 1]$ باشد (منظور از مشتق‌پذیری کسری، همان مشتق کسری همدیس‌پذیر است). در این صورت f در نقطه t_0 پیوسته است.

لم ۲.۵.۱ [۲۸] فرض کنید $\alpha \in (0, 1]$ و $f : [0, \infty) \rightarrow \mathcal{R}$ تابعی مشتق‌پذیر باشد. در این صورت برای $t_0 > 0$ داریم

$$T^{\alpha}(f)t = t^{1-\alpha} \frac{df}{dt}(t).$$

علاوه بر دو خاصیت فوق، مشتق کسری همدیس‌پذیر خواص دیگری نیز دارد که برخی از آن‌ها در زیر آورده شده‌اند.

لم ۳.۵.۱. [۲۸، ۲۹] برای هر f و g مشتق‌پذیر از مرتبه α

۱. اگر $a, b \in \mathbb{R}$ آن‌گاه $T^\alpha(af + bg) = aT^\alpha(f) + bT^\alpha(g)$.

۲. برای هر $p \in \mathbb{R}$ داریم $T^\alpha(t^p) = pt^{p-\alpha}$.

۳. برای تابع ثابت C ، $f(t) = C$ ، $T^\alpha(C) = 0$.

۴. برای توابع مشتق‌پذیر کسری f و g داریم

$$T^\alpha(fg) = fT^\alpha(g) + gT^\alpha(f), \quad T^\alpha\left(\frac{f}{g}\right) = \frac{gT^\alpha(f) - fT^\alpha(g)}{g^2}.$$

۵- برای تابع پیوسته f داریم $T^\alpha(I^\alpha)(t) = f(t)$.

۶.۱ مشتق‌های کسری مرتبه متغیر

هدف این قسمت بررسی مشتقات مرتبه کسری با α وابسته به زمان است. در این مشتقات مرتبه آن‌ها به صورت یک تابع $\alpha(t)$ در نظر گرفته می‌شود که مقادیر خود را از بازه $(n-1, n)$ که $n = [\max_{a \leq t \leq b}(\alpha(t))] + 1$ ، اختیار می‌کند. با فرض تابع $f(t) \in C^n[a, b]$ تعاریف زیر را داریم

تعریف ۱.۶.۱. مشتق کسری چپ و راست ریمن-لیوویل مرتبه $\alpha(t)$ از تابع $f(t)$ عبارت است از

$${}_a D_t^{\alpha(t)} f(t) = \frac{1}{\Gamma(n - \alpha(t))} \left(\frac{d}{dt}\right)^n \int_a^t (t - \tau)^{n - \alpha(t) - 1} f(\tau) d\tau, \quad (t > a), \quad (12.1)$$

$${}_t D_b^{\alpha(t)} f(t) = \frac{(-1)^n}{\Gamma(n - \alpha(t))} \left(\frac{d}{dt}\right)^n \int_t^b (\tau - t)^{n - \alpha(t) - 1} f(\tau) d\tau, \quad (t < b).$$

تعریف ۲.۶.۱. مشتق کسری چپ و راست کاپاتو مرتبه $\alpha(t)$ از تابع $f(t)$ عبارت است از

$${}_a^C D_t^{\alpha(t)} f(t) = \frac{1}{\Gamma(n - \alpha(t))} \int_a^t (t - \tau)^{n - \alpha(t) - 1} f^{(n)}(\tau) d\tau, \quad (t > a), \quad (13.1)$$

$${}_t^C D_b^{\alpha(t)} f(t) = \frac{(-1)^n}{\Gamma(n - \alpha(t))} \int_t^b (\tau - t)^{n - \alpha(t) - 1} f^{(n)}(\tau) d\tau, \quad (t < b).$$

تعریف ۳.۶.۱. مشتق کسری مرتبه متغیر گرانوالد-لتنیکف از مرتبه $\alpha(t)$ از تابع $f(t)$ به صورت

$${}_a^G D_t^{\alpha(t)} f(t) = \lim_{h \rightarrow 0} \frac{1}{h^{\alpha(t)}} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha(t)}{k} f(t - kh), \quad (14.1)$$

تعریف می‌شود.

یک رابطه بسیار سودمند از مشتق کاپاتو به صورت زیر است

$$\begin{aligned} ({}_a^C D_t^{\alpha(t)} (t - a)^\gamma)(t) &= \frac{\Gamma(\gamma + 1)}{\Gamma(\gamma - \alpha(t) + 1)} (t - a)^{\gamma - \alpha(t)}, \quad (\mathcal{R}(\alpha(t)) \geq 0), \quad \gamma > 0, \\ ({}_t^C D_b^{\alpha(t)} (b - t)^\gamma)(t) &= \frac{\Gamma(\gamma + 1)}{\Gamma(\gamma - \alpha(t) + 1)} (b - t)^{\gamma - \alpha(t)}, \quad (\mathcal{R}(\alpha(t)) \geq 0), \quad \gamma > 0. \end{aligned} \quad (15.1)$$

همچنین با مراجعه به [۳۰] می‌توان چند تعریف دیگر از مشتقات ریمن-لیوویل و کاپاتو از مرتبه متغیر و روابط بین آن‌ها را مشاهده کرد.

۷.۱ چندجمله‌ای‌های مرتبه کسری

چندجمله‌ای‌های کلاسیک متعامد همچون لژاندر^{۲۷}، چیشف^{۲۸}، لاگرانژ^{۲۹} و غیره نقش بسیار مهمی در حل مسائل مربوط به معادلات دیفرانسیل و کنترل بهینه کلاسیک ایفا می‌کنند. این چندجمله‌ای‌ها دارای توان‌های صحیح هستند. مشتق کسری این چندجمله‌ای‌ها، توابعی با توان‌های کسری هستند. این ویژگی باعث می‌شود که با افزایش تعداد چندجمله‌ای‌ها خطا کاهش یابد. یکی از روش‌های مؤثر برای حل مسائل مربوط به حسابان کسری استفاده از توابع پایه به صورت $\Phi_n = \sum_{i=0}^n c_i x^{i\alpha}$ است. مشتق کسری از این چندجمله‌ای‌ها توابعی با توان‌های کسری است که باعث می‌شود نرخ همگرایی این چندجمله‌ای‌ها نسبت به توابع پایه کلاسیک سریع‌تر باشد. در ادامه برخی از توابع پایه از مرتبه کسری معرفی می‌شود.

۱.۷.۱ چندجمله‌ای‌های لاگرانژ مرتبه کسری

فرض کنید $\{t_j\}_{j=0}^{N-1}$ نقاط درونیابی در بازه $[0, 1]$ هستند. در این صورت زامین چندجمله‌ای لاگرانژ $G_j(t)$ از مرتبه $N-1$ به صورت زیر تعریف می‌شود [۳۱]

$$G_j(t) = \prod_{i=0, i \neq j}^{N-1} \frac{t - t_i}{t_j - t_i}, \quad j = 0, 1, \dots, N-1.$$

با استفاده از تغییر متغیر $t = \tau^\alpha$ ، زامین چندجمله‌ای لاگرانژ کسری می‌تواند به صورت زیر تعریف شود

$$G_j^\alpha(\tau) = \prod_{i=0, i \neq j}^{N-1} \frac{\tau^\alpha - t_i}{t_j - t_i}, \quad j = 0, 1, \dots, N-1,$$

که $t_i = \tau_i^\alpha$. مشابه چندجمله‌ای‌های لاگرانژ ویژگی زیر برای چندجمله‌ای لاگرانژ کسری برقرار است

$$G_j^\alpha(\tau) = \begin{cases} 1, & \tau = \tau_j, \\ 0, & \tau \neq \tau_j. \end{cases}$$

²⁷Legendre

²⁸Chebyshev

²⁹Lagrange

زائرنوری^{۳۰} و همکارانش [۳۲] روش دیگری برای محاسبه این چندجمله‌ای‌ها ارائه دادند. آن‌ها چندجمله‌ای‌های لژاندر کسری از درجه $N + \alpha - 1$ را به صورت زیر تعریف کردند

$$h_j^\alpha(x) = \left(\frac{x - x_1}{x_j - x_1}\right)^\alpha \prod_{k=1, k \neq j}^N \frac{x - x_k}{x_j - x_k}, \quad 2 \leq j \leq N.$$

برای مطالعه بیشتر در مورد این چندجمله‌ای‌ها [۳۱، ۳۲] را ببینید.

۲.۷.۱ چندجمله‌ای‌های لژاندر مرتبه کسری

چندجمله‌ای‌های لژاندر چندجمله‌ای‌های متعامد در بازه $[-1, 1]$ هستند که در رابطه بازگشتی زیر صدق می‌کنند [۳۳]

$$L_{i+1} = \frac{2i+1}{i+1} x L_i(x) - \frac{i}{i+1} L_{i-1}(x), \quad i = 1, 2, \dots,$$

که $L_0(x) = 1$ و $L_1(x) = x$. با استفاده از تغییر متغیر $x = 2t - 1$ چندجمله‌ای‌های لژاندر انتقال یافته $P_m(t)$ از مرتبه m روی بازه $[0, 1]$ به صورت زیر نمایش داده می‌شوند

$$P_{m+1}(t) = \frac{2m+1}{m+1} (2t-1) P_m(t) - \frac{m}{m+1} P_{m-1}(t), \quad m = 1, 2, \dots,$$

$$P_0(t) = 1, \quad P_1(t) = 2t - 1.$$

برای چندجمله‌ای‌های $P_m(t)$ داریم

$$\int_0^1 P_m(t) P_n(t) dt = \begin{cases} \frac{1}{2n+1}, & m = n, \\ 0, & m \neq n. \end{cases}$$

فرم تحلیلی چندجمله‌ای‌های لژاندر انتقال یافته $P_n(t)$ به صورت زیر است

$$P_n(t) = \sum_{k=0}^n (-1)^{n-k} \frac{(n+k)!}{(n-k)!(k!)^2}.$$

همچنین این چندجمله‌ای‌ها با استفاده از فرمول رودریگز^{۳۱} روی بازه $[0, 1]$ به صورت زیر قابل نمایش هستند

$$L_n(t) = \frac{1}{2^n n!} \frac{d^n}{dt^n} ((2t-1)^2 - 1)^n.$$

در [۳۴] با جایگزینی مشتق کسری با مشتق معمولی در فرمول رودریگز، تعمیمی از چندجمله‌ای‌های لژاندر از مرتبه کسری به دست می‌آید. پیچیدگی این چندجمله‌ای‌ها امکان استفاده از آن‌ها را در مسائل کاربردی دشوار ساخته است. از این رو در [۳۵] با استفاده از تغییر متغیر $t = x^\alpha$ در فرم بازگشتی (۲.۷.۱) چندجمله‌ای‌های لژاندر مرتبه کسری انتقال یافته در بازه $[0, 1]$ را برای

³⁰Zayernouri

³¹Rodrigues Formula

حل معادلات دیفرانسیل کسری مورد مطالعه قرار داده‌اند. اگر FL_m^α چندجمله‌ای‌های لژاندر مرتبه کسری از درجه m باشد آن‌گاه

$$FL_{m+1}^\alpha(x) = \frac{\Upsilon m + 1}{m + 1}(\Upsilon x^\alpha - 1)FL_m^\alpha(x) - \frac{m}{m + 1}FL_{m-1}^\alpha(x), \quad m = 1, 2, \dots,$$

$$FL_0^\alpha(x) = 1, \quad FL_1^\alpha(x) = \Upsilon x^\alpha - 1.$$

فرم تحلیلی این چندجمله‌ای‌ها به صورت زیر است

$$FL_n^\alpha(x) = \sum_{k=0}^n (-1)^{n+k} \frac{(n+k)!}{(n-k)!(k!)^\Upsilon} x^{k\alpha}.$$

این چندجمله‌ای‌ها با توجه به تابع وزن $w(x) = x^{\alpha-1}$ در بازه $(0, 1]$ متعامد هستند

$$\int_0^1 FL_m^\alpha(x) FL_n^\alpha(x) w(x) dx = \begin{cases} \frac{1}{(\Upsilon n + 1)\alpha}, & m = n, \\ 0, & m \neq n. \end{cases}$$

تابع دلخواه f متعلق به بازه $(0, 1]$ قابل تقریب با استفاده از چندجمله‌ای‌های لژاندر مرتبه کسری به صورت زیر است

$$f(t) = \sum_{n=0}^{\infty} c_n FL_n^\alpha(t), \quad c_n = \alpha(\Upsilon n + 1) \int_0^1 f(t) FL_n^\alpha(t) w(t) dt.$$

۳.۷.۱ چندجمله‌ای‌های چبیشف مرتبه کسری

چندجمله‌ای‌های چبیشف در بازه $[-1, 1]$ تعریف می‌شوند و از فرمول بازگشتی زیر قابل محاسبه هستند [۳۶]

$$T_{i+1}(x) = \Upsilon x T_i(x) - T_{i-1}(x), \quad i = 1, 2, \dots,$$

که $T_0(x) = 1$ و $T_1(x) = x$. رابطه تعامد این چندجمله‌ای‌ها به صورت زیر بیان می‌شود

$$\int_{-1}^1 \frac{T_m(x) T_n(x)}{\sqrt{x}^\Upsilon} dx = \begin{cases} \pi, & m = n = 0, \\ \frac{\pi}{\Upsilon}, & m = n \neq 0, \\ 0, & m \neq n. \end{cases}$$

با استفاده از تغییر متغیر $x = \frac{\Upsilon t}{L} - 1$ چندجمله‌ای‌های چبیشف انتقال یافته $T_m^*(t)$ از مرتبه m در بازه $[0, L]$ به صورت زیر است

$$T_m^*(t) = T_m\left(\frac{\Upsilon t}{L} - 1\right),$$

که $T_0(t) = 1$ و $T_1(t) = \left(\frac{\Upsilon t}{L} - 1\right)$. فرم تحلیلی چندجمله‌ای‌های چبیشف انتقال یافته $T_m^*(t)$ از مرتبه m در بازه $[0, L]$ به صورت زیر است

$$T_m^*(t) = n \sum_{k=0}^n (-1)^{n-k} \frac{(n+k-1)! \Upsilon^{\Upsilon k}}{(n-k)!(\Upsilon k)! L^k} t^k.$$

مشابه چندجمله‌ای‌های لژاندر مرتبه کسری می‌توان با تغییر متغیر $t = x^\alpha$ چندجمله‌ای‌های چبیشف مرتبه کسری زیر را تعریف کرد [۳۷]

$$FT_n^\alpha(x=n) \sum_{k=0}^n (-1)^{n-k} \frac{(n+k-1)! 2^k}{(n-k)! (2k)! L^k} x^{k\alpha}.$$

خاصیت تعامد برای این چندجمله‌ای‌ها به صورت زیر است

$$\int_0^L FT_m^\alpha(x) FT_n^\alpha \frac{x^{\alpha-1}}{\sqrt{Lx^\alpha - x^{2\alpha}}} dx = \delta_{mn} h_n,$$

که $h_n = \frac{b_n \pi}{2\alpha}$ با $b_0 = 2$. تابع دلخواه f متعلق به بازه $(0, L]$ قابل تقریب با استفاده از چندجمله‌ای‌های چبیشف مرتبه کسری به صورت

$$f(t) = \sum_{n=0}^{\infty} c_n FT_n^\alpha(t),$$

است که ضرایب c_n به صورت

$$c_n = \frac{1}{h_n} \int_0^L f(t) FT_n^\alpha(t) w(t) dt, \quad n = 0, 1, 2, \dots,$$

محاسبه می‌شوند.

۴.۷.۱ چندجمله‌ای‌های مونتس-لژاندر

در این بخش چندجمله‌ای‌های مونتس^{۳۲} بررسی می‌شود. همه مفاهیمی که در این بخش بیان می‌شود را می‌توان با جزئیات بیشتر در [۳۸] یافت. فرض کنید $0 \leq \lambda_0 < \lambda_1 < \dots < \infty$. قضیه کلاسیک مونتس-سزار^{۳۳} بیان می‌کند که چندجمله‌ای‌های مونتس به فرم $\sum_{k=0}^n a_k x^{\lambda_k}$ در $L^2[0, 1]$ چگال هستند اگر

$$\sum_{k=0}^n a_k x^{\lambda_k} = \infty.$$

اگر تابع ثابت ۱ متعلق به سیستم باشد یا به عبارت دیگر $\lambda_0 = 0$ ، آنگاه نتایج مشابهی برای $C[0, 1]$ با نرم یکنواخت برقرار است. فرض کنید اعداد مختلط در مجموعه $\Lambda_n = \{\lambda_0, \lambda_1, \dots, \lambda_n\}$ در رابطه $\mathcal{R}(\lambda_k) > -\frac{1}{4}$ برای هر $k \in \mathbb{N}$ صدق کند آنگاه چندجمله‌ای‌های مونتس-لژاندر^{۳۴} روی بازه $[0, 1]$ به صورت زیر تعریف می‌شوند [۳۹]

$$M_n(x) = M_n(x, \Lambda_n) = \sum_{k=0}^n C_{n,k} x^{\lambda_k}, \quad C_{n,k} = \frac{\prod_{\nu=0}^{n-1} (\lambda_k + \bar{\lambda}_\nu + 1)}{\prod_{\nu=0, \nu \neq k}^{n-1} (\lambda_k + \bar{\lambda}_\nu)},$$

³²Muntz

³³Muntz-Szasz

³⁴Muntz-Legendre

که با توجه به ضرب داخلی در بازه $[0, 1]$ متعامد هستند. یعنی

$$(M_n, M_m) = \int_0^1 M_n(x) M_m(x) dx = \frac{\delta_{nm}}{\lambda_n + \bar{\lambda}_n + 1}, \quad (16.1)$$

که δ_{nm} دلتای کرونکر^{۳۵} و $(.,.)$ ضرب داخلی را نشان می‌دهد. با در نظر گرفتن $\lambda_k = \alpha k$ ، چندجمله‌ای‌های مونتمس-لژاندر کسری روی بازه $[0, 1]$ به صورت زیر نمایش داده می‌شوند [۴۰]

$$L_{n,\alpha}(t) = \sum_{k=0}^n C_{n,k} t^{k\alpha}, \quad C_{n,k} = \frac{(-1)^{n+k}}{\alpha^n k! (n-k)!} \prod_{\nu=0}^{n-1} ((k+\nu)\alpha + 1).$$

با توجه به ویژگی تعامد (۱۶.۱) داریم

$$(L_{n,\alpha}, L_{m,\alpha}) = \int_0^1 L_{n,\alpha}(t) L_{m,\alpha}(t) dt = \frac{\delta_{nm}}{\Gamma(n\alpha + 1)}.$$

هر تابع دلخواه $f \in L^2[0, 1]$ توسط چندجمله‌ای‌های مونتمس-لژاندر به صورت زیر قابل تقریب است

$$f(t) = \sum_{n=0}^{\infty} c_n L_{n,\alpha}(t), \quad c_n = (\Gamma(n\alpha + 1)) \int_0^1 f(t) L_{n,\alpha}(t) dt.$$

مهم‌ترین ویژگی چندجمله‌ای‌های مونتمس-لژاندر یعنی مشتق کسری این چندجمله‌ای‌ها به صورت زیر بیان می‌شود

$${}^C D^\alpha L_{n,\alpha}(t) = \sum_{k=\lceil \alpha \rceil}^n D_{n,k} t^{(k-1)\alpha}, \quad D_{n,k} = \frac{\Gamma(k\alpha + 1)}{\Gamma(k\alpha + 1 - \alpha)} C_{n,k}.$$

³⁵Kronecker delta

فصل ۲

پیش‌نیازها از شبکه عصبی و معرفی سیستم‌های دینامیکی

۱.۲ پیش‌نیازها از شبکه عصبی

یکی از روش‌های کارآمد در حل مسائل پیچیده، شکستن آن به زیرمسئله‌های ساده‌تر است که هر کدام از این زیرمسئله‌ها به نحو ساده‌تری قابل درک و توصیف هستند. در حقیقت یک شبکه مجموعه‌ای از این ساختارهای ساده است که در کنار یکدیگر سیستم پیچیده نهایی را توصیف می‌کند. شبکه‌ها انواع مختلفی دارند اما همگی آن‌ها از دو مؤلفه تشکیل می‌شوند

الف) مجموعه‌ای از گره‌ها: هر گره در حقیقت واحد محاسباتی شبکه است که ورودی‌ها را گرفته و بر روی آن پردازش انجام می‌دهد تا خروجی به دست آید. پردازش انجام شده توسط گره می‌تواند از ساده‌ترین نوع پردازش‌ها نظیر جمع کردن ورودی‌ها تا پیچیده‌ترین محاسبات را شامل شود. در حالت خاص، یک گره می‌تواند خود شامل یک شبکه دیگر باشد.

ب) اتصالات بین گره‌ها: این اتصالات نحوه گذر اطلاعات بین گره‌ها را مشخص می‌کند. در

حالت کلی اتصالات می‌توانند تک سویه^۱ یا دوسویه^۲ باشند.

۱.۱.۲ شبکه‌های عصبی بیولوژیکی

مغز انسان سیستم پیچیده‌ای است که قدرت تفکر، تحلیل و یادآوری را دارد و سیستم اعصاب انسان دارای ویژگی‌های شگفت‌آوری است. مغز انسان قادر به انجام کارهایی است که برای پیشرفته‌ترین کامپیوترها غیرممکن است که مهم‌ترین دلیل آن، عملکردها و پردازش‌های موازی و هم‌زمان در شبکه اعصاب است. نرون^۳ یک واحد سلولی از سیستم مغز است و مغز انسان تقریباً از صد میلیون نرون ساخته شده است. نرون‌ها با وجود تفاوت‌های زیاد از نظر اندازه و شکل ظاهری، مشخصه‌های مشترکی دارند. همان‌طور که در شکل ۱.۲ نشان داده شده است از تنه سلولی نرون تعدادی شاخک کوتاه خارج می‌شود که دندریت^۴ نام دارد. دندریت‌ها و تنه سلولی، سیگنال‌ها را از نرون‌های مجاور دریافت می‌کنند و از طریق یک لوله باریک به نام آکسون^۵ به نرون‌های دیگر منتقل می‌شوند.

آکسون در انتهای خود به تعدادی رشته جانبی باریک تقسیم می‌شود که پایانه آکسونی نام دارد و با دندریت‌های سایر نرون‌ها مرتبط است. ارتباط بین آکسون یک نرون و دندریت نرون دیگر را سیناپس^۶ می‌نامند. همه سیگنال‌های جمع‌بندی شده در تنه نرون ترکیب می‌شوند و اگر وسعت سیگنال‌های ترکیب شده به آستانه نرون برسد، مرحله تحریک شدن^۷ فعال می‌شود و یک سیگنال خروجی تولید می‌شود. این سیگنال به صورت یک پالس منفرد یا بخشی از پالس‌ها در یک میزان خاص به موازات آکسون به پایانه‌های سیناپسی انتقال می‌یابد و موجب ترشح موادی به نام عصب-رسانه می‌شود. عصب-رسانه در داخل شکاف سیناپسی پخش می‌شود و نرون بعدی را تحریک می‌کند و به این ترتیب یک سیگنال از یک نرون به دیگری انتقال می‌یابد. تعداد بسیار زیادی آکسون از نرون‌های مختلف با دندریت‌های یک نرون به این صورت ارتباط برقرار می‌کنند. نرون‌ها با توجه به عملکردشان به سه دسته تقسیم می‌شوند

(۱) **نرون‌های حسی:** این نرون‌ها اطلاعات را از ارگان‌های حسی به مغز و نخاع می‌فرستند.

(۲) **نرون‌های حرکتی:** این نرون‌ها سیگنال‌های فرمان را از مغز و نخاع به ماهیچه‌ها و غدد هدایت می‌کنند.

(۳) **نرون‌های رابط:** این نرون‌ها پیام‌ها را از نرون‌های حسی می‌گیرند و به یک نرون رابط دیگر انتقال می‌دهند و یا پیام‌ها را به یک نرون حرکتی می‌رسانند. نرون‌های رابط فقط در

¹Unidirectional

²Bidirectional

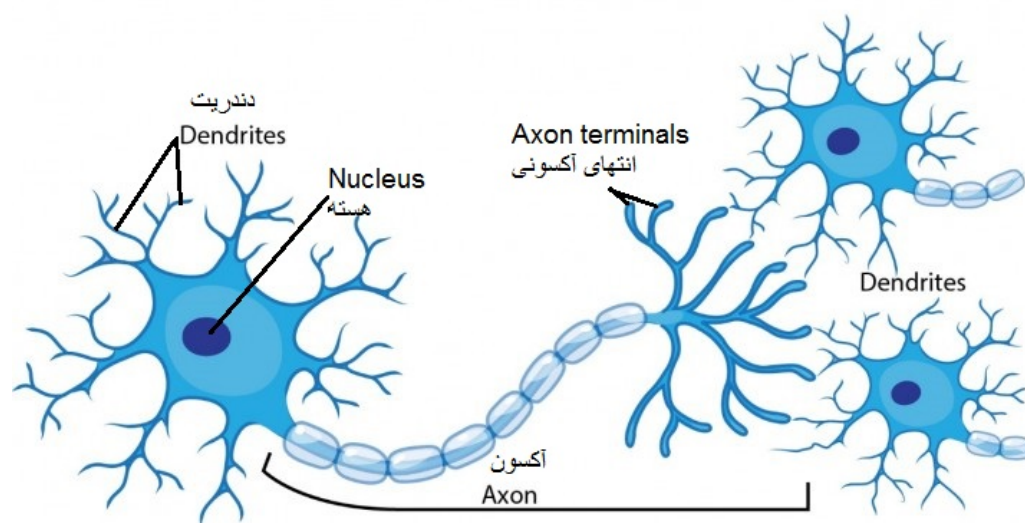
³Neuron

⁴Dendrite

⁵Axon

⁶Synapse

⁷Firing



شکل ۱.۲: ساختار نرون

مغز، چشم و طناب نخاعی یافت می‌شوند.

سوال این است که با این تعریف، یادگیری چیست؟ پاسخ این سوال در شبکه‌های عصبی مصنوعی مهم خواهد بود.

تعریف ۱.۱.۲. عمل یادگیری تغییر در نوع فعالیت شیمیایی دندریت‌ها و بدنه سلول و تنظیم مجدد آن فعالیت‌ها است.

در نتیجه حفظ اطلاعات و ذخیره‌سازی آن‌ها در مغز چیزی جز فعالیت‌های شیمیایی در بدنه سلول و پیوندهای آن نیست. نتیجه دیگر اینکه هنگام عبور سیگنال از دندریت‌ها و بدنه سلول و تا رسیدن به سیناپس‌ها دستخوش تغییراتی می‌شود. اگر به واسطه فرآیندهای شیمیایی این تغییرات با قوانین خاص جدیدی صورت گیرد یک یادگیری جدید رخ داده است. همین مقدمه و معرفی از شبکه‌های عصبی زیستی کافی است تا بتوان شبکه‌های عصبی مصنوعی را بیان کرد.

۲.۱.۲ شبکه‌های عصبی مصنوعی

شبکه‌های عصبی مصنوعی^۸ پدیده نسبتاً جدیدی است. هر چند شروع تحقیقات در زمینه شبکه‌های عصبی مصنوعی به دهه ۱۹۴۰ میلادی برمی‌گردد اما گسترش آن در دهه ۱۹۸۰ صورت پذیرفت [۱]. شبکه‌های عصبی شاخه‌ای از هوش مصنوعی است و در واقع تلاشی برای پیاده‌سازی نرون‌های عصبی مغز انسان به صورت مصنوعی است. در یک شبکه عصبی

^۸Artificial Neural Network

مصنوعی سلول‌های عصبی مغز انسان یا همان نرون‌ها به صورت یک مدل ریاضی بیان می‌شوند و با مرتب کردن سلول‌های عصبی در اشکال مختلف می‌توان برخی از عملکردهای مغز انسان مانند ذخیره اطلاعات در حافظه، الگوها، فرآیند یادگیری و ارتباط برقرار کردن بین الگوها را پیاده‌سازی کرد.

مهم‌ترین ویژگی شبکه‌های عصبی مصنوعی قابلیت یادگیری در آن‌ها است. به این مفهوم که این شبکه بر مبنای یک قاعده مشخص و ثابت برنامه‌ریزی نشده و با گذشت زمان و یا در هر تکرار شبکه وضعیت آن تغییر می‌کند یا به اصطلاح آموزش می‌بیند. مبحث شبکه‌های عصبی مصنوعی مربوط به الگوبرداری از قوه یادگیری در انسان و پیاده‌سازی آن به صورت الگوریتم‌های کامپیوتری است. بنابراین شبکه‌های عصبی این قابلیت را دارند که کارهای پیچیده‌ای که برای سیستم‌های مبتنی بر قاعده مشکل هستند را فرا گیرند.

یک شبکه عصبی مصنوعی نمی‌تواند در مورد کارهایی که انجام می‌دهد استدلال کند و این یک نقطه ضعف برای شبکه عصبی مصنوعی محسوب می‌شود. به عنوان مثال هنگامی که یک شبکه عصبی عمل تشخیص چهره را آموخت، نمی‌تواند توضیح دهد که چگونه این کار را انجام داده است. لذا در حال حاضر تحقیقات بر روی دسته‌ای از سیستم‌های هوشمند که مرکب از شبکه‌های عصبی با قابلیت یادگیری و سیستم‌های مبتنی بر قاعده (با قابلیت توضیح و استدلال) هستند انجام می‌شود.

سیر تکامل شبکه عصبی مصنوعی

از قرن نوزدهم به‌طور همزمان اما جداگانه از سویی نروفیزیولوژیست‌ها سعی کردند سیستم یادگیری و تجزیه و تحلیل مغز را کشف کنند و از سوی دیگر ریاضی‌دانان تلاش کردند مدل ریاضی بسازند که قابلیت فراگیری و تجزیه و تحلیل عمومی مسائل را دارا باشد. اولین کوشش‌ها در شبیه‌سازی با استفاده از یک مدل منطقی در اوایل دهه‌ی ۱۹۴۰ توسط وارن مک‌کالک^۹ و والتر پیتز^{۱۰} انجام شد که امروزه معیار اصلی سازنده اکثر شبکه‌های عصبی مصنوعی است. عملکرد این مدل مبتنی بر جمع ورودی‌ها و ایجاد خروجی با استفاده از شبکه‌ای از نرون‌ها است. اگر حاصل جمع ورودی‌ها از مقدار آستانه بیشتر باشد اصطلاحاً نرون برانگیخته می‌شود. نتیجه این مدل اجرای ترکیبی از توابع منطقی بود. در سال ۱۹۴۹ دونالد هب^{۱۱} قانون یادگیری را برای شبکه‌های عصبی طراحی کرد [۴۱]. در سال ۱۹۵۸ شبکه پرسپترون^{۱۲} توسط روزنبلات^{۱۳} معرفی گردید. این شبکه نظیر واحدهای مدل شده قبلی بود. پرسپترون دارای سه لایه است که شامل لایه‌ی ورودی، لایه‌ی خروجی و لایه‌ی میانی است. این سیستم می‌تواند یاد بگیرد که با روشی تکرارشونده وزن‌ها را به گونه‌ای تنظیم کند که

⁹Warren McCulloch

¹⁰Walter Pitts

¹¹Donald Hebb

¹²Perceptron

¹³Rosenblatt

شبکه توان بازتولید جفت‌های ورودی و خروجی را داشته باشد [۴۲]. روش دیگر مدل خطی تطبیقی نرون است که در سال ۱۹۶۰ توسط برنارد ویدرو^{۱۴} و مارسیان هاف^{۱۵} در دانشگاه استنفورد به وجود آمد که اولین شبکه‌های عصبی به کار گرفته شده در مسائل واقعی بود. آن‌ها یک مدل سه سطحی از نرون‌ها را ابداع کردند که آدلاین^{۱۶} نام‌گذاری شد و مختصر شده حروف "عنصر خطی تطبیقی"^{۱۷} بود [۴۳، ۴۴]. همچنین برای آدلاین یک قانون تعلیم بر مبنای مشتق‌گیری بیان نمودند. این قانون تعلیم به $\alpha - LMS$ معروف است. از متصل کردن چند آدلاین به یکدیگر مادلاین^{۱۸} ساخته شد. ویدرو و هاف از شبکه عصبی مادلاین برای حذف اکو از خطوط تلفن استفاده کردند. در سال ۱۹۶۹ میسکی و پاپرت^{۱۹} کتابی نوشتند که محدودیت‌های سیستم‌های تک لایه و چند لایه پرسپترون را تشریح کردند. نتیجه این کتاب پیش‌داوری و قطع سرمایه‌گذاری برای تحقیقات در زمینه شبیه‌سازی شبکه‌های عصبی بود. آن‌ها با طرح اینکه شبکه پرسپترون قادر به حل هیچ مسئله جالبی نیست، تحقیقات در این زمینه را برای مدت چند سال متوقف کردند. پال ورباس^{۲۰} در سال ۱۹۷۴ شیوه آموزش پس انتشار^{۲۱} خطا را ایجاد کرد که یک شبکه پرسپترون چندلایه البته با قوانین نیرومندتر آموزشی بود. پیشرفت‌هایی که در سال ۱۹۷۰ تا ۱۹۸۰ به دست آمد برای جلب توجه به شبکه‌های عصبی بسیار مهم بود. سال‌های ۱۹۸۲ تا ۱۹۸۶ را می‌توان تولد دوباره شبکه عصبی دانست. در این سال‌ها دو اتفاق مهم در این زمینه رخ داد. اول ارائه شبکه عصبی بازگشتی و مفهوم تابع انرژی توسط جان هاپفیلد^{۲۲} بود که در سال ۱۹۸۲ منتشر شد [۴۵]. دومین اتفاق، حل مسئله فروشنده دوره‌گرد (TSP) توسط شبکه هاپفیلد بود [۴۶]. هاپفیلد نشان داد که از این شبکه بازگشتی می‌توان برای حل مسائل بهینه‌سازی استفاده کرد و افق جدیدی را در شبکه‌های عصبی مصنوعی گشود.

مدل ریاضی نرون عصبی مصنوعی

در هر شبکه عصبی مصنوعی اطلاعاتی که قرار است پردازش شود، در محلی که نرون نام دارد ذخیره می‌شود. هر کدام از این اطلاعات یا سیگنال‌ها با یک نوع ارتباط به نرون‌های دیگر متصل می‌شوند. در یک نگاه ساده مدل یک نرون باید شامل $n + 1$ ورودی x_j ، $(j = 1, 2, \dots, n + 1)$ باشد. این ورودی‌ها در وزن‌های سیناپسی^{۲۳} w_j ، $(j = 1, 2, \dots, n + 1)$ که

¹⁴Bernard Widrow

¹⁵Marcian Hoff

¹⁶Adeline

¹⁷Adaptive Linear Elements

¹⁸Madaline

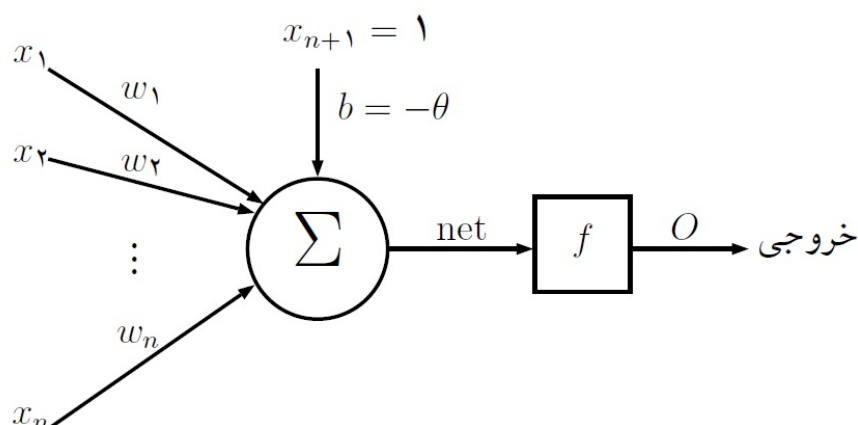
¹⁹Papert and Minsky

²⁰Paul Werbos

²¹Back Propagation

²²John Hopfield

²³Short Tetm Memory



شکل ۲.۲: مدل ریاضی نرون عصبی مصنوعی

روی آکسون‌ها واقع هستند و نقش ساختار سیناپسی را ایفا می‌کنند ضرب می‌شوند. وزن اولین ورودی که مقدار آن همواره ۱- است را بایاس^{۲۴} می‌نامند [۲]. مجموع حاصل ضرب ورودی‌ها در وزن‌های متناظرشان توسط تابعی به نام تابع فعال‌ساز^{۲۵} به یک مقدار دیگر نگاشته می‌شود. تابع فعال‌ساز تصمیم‌گیری می‌کند که آیا نرون فعال شود یا خیر و اگر جواب مثبت باشد میزان خروجی را مشخص می‌سازد. بنابراین شبکه عصبی مصنوعی با استفاده از مدل ساده شده نرون واقعی به پردازش اطلاعات می‌پردازد. با توجه به این توضیحات می‌توان مدل ساده‌ای برای توصیف یک نرون (یک گره در شبکه عصبی مصنوعی) پیشنهاد کرد. این مدل در شکل ۲.۲ نشان داده شده است. جدای از ساده‌سازی‌های اعمال شده تفاوت اصلی این مدل با واقعیت در این است که در شبکه واقعی ورودی‌ها سیگنال‌های زمانی‌اند، حال آن‌که در این مدل ورودی اعداد حقیقی هستند. مشخصه‌های اصلی یک شبکه عصبی که در طراحی آن باید مد نظر قرار گیرند عبارت است از

معماری شبکه عصبی: نحوه اتصالات بین نرون‌ها، تعداد آن‌ها و تعداد لایه‌های تشکیل دهنده بخش نرون‌های ارتباطی را معماری شبکه عصبی گویند.

تابع‌های فعال‌ساز: این‌که چه تابعی روی نرون قرار گیرد تا بر ورودی‌های نرون اثر کرده و خروجی نرون را تولید کند تابع فعال‌ساز آن نرون گویند.

الگوریتم آموزش: روشی که وزن‌های روی آکسون‌ها بر اساس آن تغییر می‌کنند تا خروجی شبکه عصبی به حالت مطلوب درآید، الگوریتم آموزش نامیده می‌شود.

توابع فعال‌ساز

در اینجا چند نمونه از توابع فعال‌ساز معرفی می‌شود

²⁴Soma

²⁵Activation Function

تابع پله‌ای: این تابع با حد آستانه‌ای θ به صورت زیر است

$$f(x) = \begin{cases} 1 & x \geq \theta, \\ 0 & x < \theta. \end{cases}$$

تابع پله‌ای دوقطبی: این تابع با حد آستانه‌ای θ به صورت زیر تعریف می‌شود

$$f(x) = \begin{cases} 1 & x \geq \theta, \\ -1 & x < \theta. \end{cases}$$

تابع سیگموئید: تابع سیگموئید^{۲۶}، تابعی یکنوا، کراندار و مشتق‌پذیر است و می‌توان مشتق تابع را به کمک خود تابع بیان کرد که به شکل زیر است

$$f(x) = \frac{1}{1 + e^{-x}}, \quad f'(x) = f(x)(1 - f(x)).$$

تابع سیگموئید قطبی: این تابع همان خواص تابع سیگموئید را دارد و به صورت زیر تعریف می‌شود

$$g(x) = 2f(x) - 1 = 2\left(\frac{1}{1 + e^{-x}}\right) - 1 = \frac{1 - e^{-x}}{1 + e^{-x}}, \quad g'(x) = \frac{1}{4}(1 + g(x))(1 - g(x)).$$

تابع همانی: این تابع بیشتر برای نرون‌های ورودی مورد استفاده قرار می‌گیرد و به صورت زیر بیان می‌شود

$$f(x) = x.$$

۳.۱.۲ انواع یادگیری برای شبکه‌های عصبی

در یک شبکه چند لایه، یادگیری عبارت است از یافتن وزن‌ها، بایاس‌ها و توابع فعال‌ساز به‌طوری‌که در یک دامنه کاری از هر ورودی، خروجی مطلوب حاصل شود. قواعد یادگیری به سه بخش عمده تقسیم می‌شوند: یادگیری با ناظر، یادگیری تشدید و یادگیری بدون ناظر. در اینجا توضیح مختصری درمورد نحوه عملکرد این قواعد ارائه می‌شود.

یادگیری با ناظر^{۲۷}

در این یادگیری به ازای هر ورودی، خروجی مطلوب معلوم است و برای یافتن وزن‌ها، ورودی‌ها به شبکه اعمال می‌شوند و خروجی شبکه به دست می‌آید، سپس این خروجی با خروجی مطلوب مقایسه می‌شود. در صورت وجود اختلاف هنوز یادگیری کامل نشده و باید وزن‌ها و بایاس‌ها تغییر یابند. این عمل تا رسیدن به پارامترهای بهینه شبکه را یادگیری با ناظر می‌گویند.

^{۲۶}Sigmoid

^{۲۷}Supervised Learning

یادگیری تشدید یا یادگیری تقویتی

یک اشکال یادگیری با ناظر این است که شبکه عصبی ممکن است بدون معلم نتواند مواضع جدیدی را که توسط مجموعه داده‌های جدید تجربی پوشانده نشده است یاد بگیرد. یادگیری از نوع تشدید این محدودیت را برطرف می‌کند. این نوع یادگیری به‌طور آنلاین^{۲۸} صورت می‌گیرد در حالی که یادگیری با ناظر را به دو صورت آنلاین و آفلاین^{۲۹} می‌توان انجام داد. در حالت آفلاین می‌توان از یک سیستم محاسب با در اختیار داشتن داده‌های یادگیری استفاده کرد و طراحی شبکه عصبی را به پایان رساند. پس از مرحله طراحی و یادگیری، شبکه عصبی به عنوان یک سیستم استاتیکی عمل می‌کند. اما در یادگیری آنلاین شبکه عصبی همراه با خود سیستم یادگیر در حال انجام کار است و از این رو مثل یک سیستم دینامیکی عمل می‌کند. یادگیری از نوع تشدید یک یادگیری آنلاین از یک نگاشت ورودی-خروجی است. این کار از طریق یک پروسه سعی و خطا به‌صورتی انجام می‌پذیرد که یک شاخص اجرایی موسوم به سیگنال تشدید بیشینه شود. این بدین معنی است که اگر شبکه عصبی پارامترهایش را به گونه‌ای تغییر داد که منجر به یک حالت مساعد شد آن‌گاه تمایل سیستم یادگیر جهت تولید آن عمل خاص تقویت یا تشدید می‌شود. در غیر این‌صورت تمایل شبکه عصبی جهت تولید آن عمل خاص تضعیف می‌شود. یادگیری تقویتی مثل یادگیری با ناظر نیست و این الگوریتم بیشتر برای سیستم‌های کنترلی کاربرد دارد.

یادگیری بدون ناظر^{۳۰}

در یادگیری بدون ناظر یا یادگیری خود سامانده پارامترهای شبکه عصبی تنها توسط پاسخ سیستم اصلاح و تنظیم می‌شوند. به عبارتی تنها اطلاعات دریافتی از محیط به شبکه را بردارهای ورودی تشکیل می‌دهند و در مقایسه با یادگیری با ناظر بردار جواب مطلوب به شبکه اعمال نمی‌شود. به عبارتی به شبکه عصبی هیچ نمونه‌ای از تابعی که قرار است بیاموزد داده نمی‌شود بلکه یک شاخص برای آموزش وجود دارد (مانند تابع هزینه). در عمل می‌بینیم که یادگیری با ناظر در مورد شبکه‌هایی که از تعداد زیادی لایه‌های نرونی تشکیل شده باشند بسیار کند عمل می‌کند و در این‌گونه موارد تلفیق یادگیری با ناظر و بدون ناظر پیشنهاد می‌گردد.

۴.۱.۲ انواع شبکه‌های عصبی مصنوعی

هر چند نحوه مدل کردن نرون جزء اساسی‌ترین نکات کلیدی در کارایی شبکه عصبی است اما نحوه برقراری اتصالات و چیدمان (توپولوژی) شبکه نیز فاکتور بسیار مهم و اثرگذاری است.

²⁸On-Line

²⁹Off-Line

³⁰Unsupervised Machine Learning

باید توجه داشت که توپولوژی مغز انسان آن قدر پیچیده است که نمی‌توان از آن به عنوان مدلی برای اعمال به شبکه عصبی استفاده نمود، چرا که مدلی که استفاده می‌کنیم، یک مدل ساده شده است در حالی که چیدمان مغز از المان‌های بسیار زیادی استفاده می‌کند. انواع مختلفی از مدل‌های محاسباتی تحت عنوان کلی شبکه‌های عصبی مصنوعی معرفی شده‌اند که هر یک برای دسته‌ای از کاربردها قابل استفاده هستند و در هر کدام از وجه مشخصی از قابلیت‌ها و خصوصیات مغز انسان الهام گرفته شده است. در همه این مدل‌ها، یک ساختار ریاضی در نظر گرفته شده است که البته به صورت گرافیکی هم قابل نمایش دادن است و یک سری پارامترها و پیچ‌های تنظیم دارد. این ساختار کلی، توسط یک الگوریتم یادگیری یا تربیت^{۳۱} آن قدر تنظیم و بهینه می‌شود، که بتواند رفتار مناسبی از خود نشان دهد. در ادامه یک مرور کوتاه از انواع مختلف شبکه‌های عصبی خواهیم داشت که مطالعه آن در ایجاد یک آشنایی اولیه بسیار موثر خواهد بود.

شبکه‌های پس‌خور

در شبکه‌های پس‌خور^{۳۲} یا برگشتی حداقل یک سیگنال برگشتی از یک نرون به همان نرون یا نرون‌های همان لایه یا نرون‌های لایه‌های قبل وجود دارد و اگر نرونی دارای فیدبک باشد بدین مفهوم است که خروجی نرون در لحظه حال نه تنها به ورودی در آن لحظه بلکه به مقدار خروجی خود نرون در لحظه‌ی گذشته نیز وابسته است. به عنوان مثال، شبکه عصبی هاپفیلد^{۳۳} یک شبکه بازگشتی است، به این مفهوم که در آن خروجی هر نرون به ورودی تمام نرون‌های دیگر متصل است. پس از اعمال ورودی اولیه به شبکه خروجی آن محاسبه می‌شود و در قدم بعد به عنوان ورودی به شبکه اعمال می‌شود. این فرآیند آنقدر تکرار می‌شود تا خروجی شبکه به نقطه ثابت همگرا شود. در اینجا شبکه به عنوان یک سیستم دینامیکی که وضعیتش در هر تکرار تغییر می‌کند، عمل می‌کند و همگرایی شبکه، یعنی رسیدن پاسخ شبکه به جوابی که بر اثر تکرار تغییر نکند. در این صورت می‌توان گفت که شبکه به حالت پایدار یا تعادل خود رسیده است.

شبکه‌های پیش‌خور

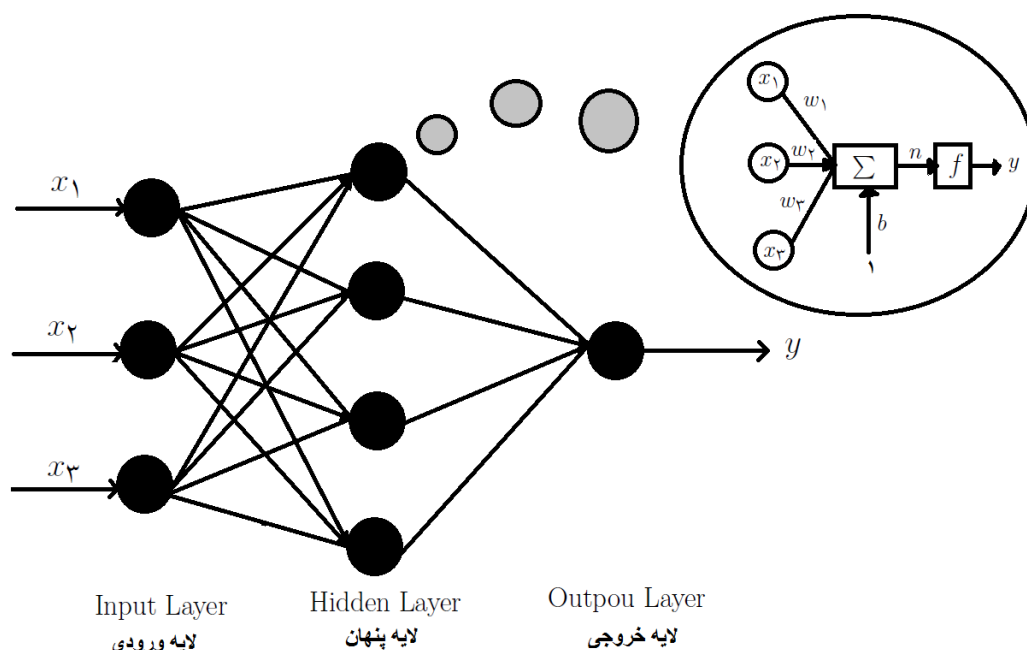
شبکه‌های پیش‌خور^{۳۴} شبکه‌هایی هستند که مسیر پاسخ در آن‌ها همواره رو به جلو پردازش شده و به نرون‌های لایه‌های قبل خود باز نمی‌گردد. در این نوع شبکه به سیگنال‌ها تنها اجازه عبور از مسیر یک‌طرفه (از ورودی تا خروجی) داده می‌شود. بنابراین بازخورد یا فیدبک وجود ندارد به این معنی که خروجی هر لایه تنها بر لایه بعد اثر می‌گذارد و در لایه‌ی خودش تغییری

³¹Training Algorithm

³²FeedBack Neural Network

³³Hopfield

³⁴Feedforward Neural Network



شکل ۳.۲: ساختار یک نرون با لایه ورودی، لایه پنهان و لایه خروجی

ایجاد نمی‌کند مانند شبکه پرسپترون.

پرسپترون چندلایه یا MLP

یکی از ساده‌ترین و در عین حال کارآمدترین چیدمان پیشنهادی برای استفاده در مدل‌سازی عصب‌های واقعی، مدل پرسپترون چندلایه (MLP)^{۳۵} است که از یک لایه ورودی (نرون‌های حسی)، یک یا چند لایه پنهان (نرون‌های ارتباطی) و یک لایه خروجی (نرون‌های حرکتی) تشکیل شده است. در این ساختار، تمام نرون‌های یک لایه به تمام نرون‌های لایه بعد متصل هستند. این چیدمان اصطلاحاً یک شبکه با اتصالات کامل را تشکیل می‌دهد. شکل ۳.۲ شمای یک شبکه پرسپترون سه لایه را نشان می‌دهد. به سادگی می‌توان استنباط نمود که تعداد نرون‌های هر لایه مستقل از تعداد نرون‌های دیگر لایه‌ها است. در پرسپترون‌های چندلایه واحدهای ورودی به واحدهای خروجی توسط یک نگاشت غیرخطی بهم متصل می‌شوند. مهم‌ترین کاربرد پرسپترون چندلایه توانایی آن‌ها در تقریب تابع است. از قضیه وجودی کولوموگوروف^{۳۶} می‌دانیم که یک پرسپترون سه لایه با $n(2n+1)$ گره می‌تواند از هر تابع پیوسته با n متغیر محاسبه شود [۴۷]. دقت تقریب به تعداد نرون‌ها در لایه‌های پنهان شده بستگی دارد و به تعداد لایه‌های پنهان شده بستگی ندارد. شبکه‌های عصبی چندلایه پیش‌خور قابل یادگیری هستند و در برخی موارد مسئله آموزشی این شبکه‌ها بدحالت است.

^{۳۵}Multi Layer Perceptron

^{۳۶}Existence Colomogorov Theorem

۲.۲ سیستم‌های دینامیکی

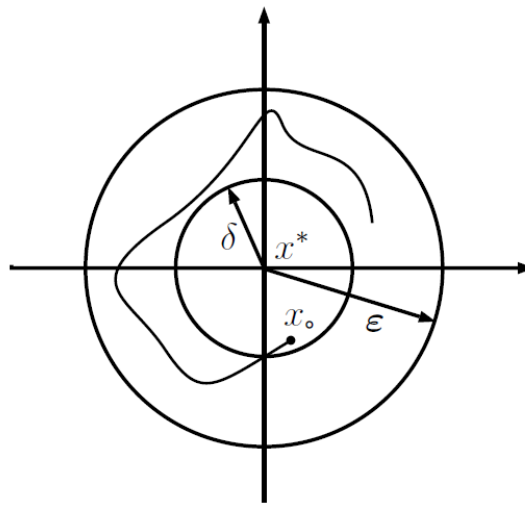
تعریف ۱.۲.۲. سیستم دینامیکی زیر را در نظر بگیرید

$$\dot{x} = f(x(t)), \quad x(t_0) = x_0 \in \mathbb{R}^n, \quad (1.2)$$

که f یک تابع از $\mathbb{R}^n$ به $\mathbb{R}^n$ است. x^* یک نقطه تعادل^{۳۷} (۱.۲) نامیده می‌شود اگر $f(x^*) = 0$. اگر همسایگی $\Omega^* \subseteq \mathbb{R}^n$ از x^* وجود داشته باشد که $f(x^*) = 0$ و $f(x) \neq 0, \forall x \in \Omega^* \setminus \{x^*\}$ ، x^* آن‌گاه یک نقطه تعادل تنها^{۳۸} نامیده می‌شود.

تعریف ۲.۲.۲ (پایداری به مفهوم لیاپانوف^{۳۹}). فرض کنید $x(t)$ یک جواب (۱.۲) است. x^* نقطه تعادل پایدار تنها، به مفهوم لیاپانوف است اگر برای هر $x_0 = x(t_0)$ و هر $\varepsilon > 0$ ، یک $\delta > 0$ وجود داشته باشد به طوری که اگر $\|x(t_0) - x^*\| < \delta$ آن‌گاه

$$\|x(t) - x^*\| < \varepsilon, \quad \forall t \geq t_0.$$



شکل ۴.۲: پایداری لیاپانوف.

تعریف ۳.۲.۲. سیستم دینامیکی (۱.۲) همگرای سراسری^{۴۰} به مجموعه جواب‌های (۱.۲) که با Ω^* نمایش داده می‌شود، است اگر برای هر نقطه آغازین دلخواه، مسیر $x(t)$ در رابطه زیر صدق کند

$$\lim_{t \rightarrow \infty} \text{dist}(x(t), \Omega^*) = 0,$$

³⁷Equilibrium Point

³⁸Isolated Equilibrium Point

³⁹Stability in the Sense of Lyapunov

⁴⁰Globally Convergent

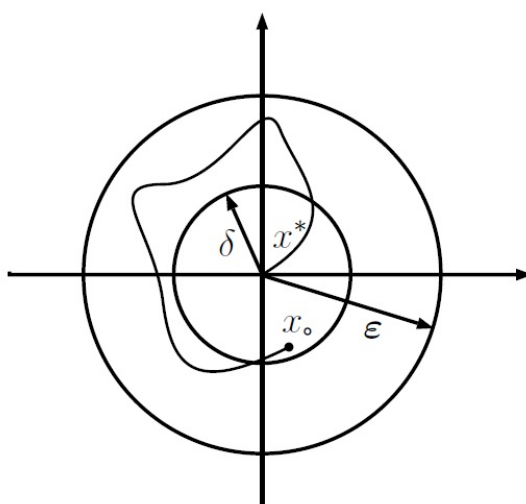
که

$$\text{dist}(x, \Omega^*) = \inf_{y \in \Omega^*} \|x - y\|,$$

و $\|\cdot\|$ نرم l_2 را نشان می‌دهد.

تعریف ۴.۲.۲. سیستم دینامیکی (۱.۲) که دارای نقطه تعادل x^* یکتا است، پایدار مجانبی سراسری^{۴۱} نامیده می‌شود اگر x^* به مفهوم لیاپانوف پایدار باشد و در شرط زیر صدق کند

$$\lim_{t \rightarrow \infty} x(t) = x^*.$$



شکل ۵.۲: پایداری مجانبی سراسری.

تعریف ۵.۲.۲. سیستم دینامیکی (۱.۲) در نقطه x^* پایدار نمایی سراسری^{۴۲} با نرخ β گفته می‌شود اگر برای هر نقطه آغازین دلخواه، مسیر $x(t)$ در رابطه زیر صدق کند

$$\forall t \geq t_0, \quad \|x(t) - x^*\| \leq \alpha \|x(t_0) - x^*\| e^{-\beta(t-t_0)},$$

که α و β ثابت‌های مثبت و مستقل از نقاط آغازین هستند. واضح است که پایداری نمایی سراسری، پایداری مجانبی سراسری را نتیجه می‌دهد.

تعریف ۶.۲.۲. مجموعه $M \subseteq \mathbb{R}^n$ یک مجموعه تغییرناپذیر^{۴۳} نسبت به سیستم (۱.۲) گفته می‌شود اگر $x(t_0) \in M$ ، $t_0 \geq 0$ و به ازای هر $t \geq t_0$ ، $x(t) \in M$ باشد.

⁴¹Globally Asymptotically Stable

⁴²Globally Exponential Stable

⁴³Invariant Set

تعریف ۷.۲.۲. تابع $F : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ روی E پیوسته لیپ‌شیتز^{۴۴} است، اگر ثابت L وجود داشته باشد به طوری که برای هر دو نقطه $x, y \in E$ داشته باشیم

$$\|F(x) - F(y)\| \leq L \|x - y\|.$$

تابع F پیوسته لیپ‌شیتز محلی^{۴۵} روی E نامیده می‌شود اگر برای نقطه $x_0 \in E$ ، یک $\varepsilon -$ همسایگی مانند $N_\varepsilon(x_0) \subset E$ و یک ثابت L_0 وجود داشته باشد به طوری که برای هر دو نقطه $x, y \in N_\varepsilon(x_0)$ داشته باشیم

$$\|F(x) - F(y)\| \leq L_0 \|x - y\|.$$

قضیه ۱.۲.۲. [۴۸] تابع $F : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ را در نظر بگیرید. اگر $F \in C^1(E)$ باشد، آن‌گاه F روی E لیپ‌شیتز محلی است.

تعریف ۸.۲.۲. فرض کنید $V : \mathbb{R}^n \rightarrow \mathbb{R}$ یک تابع به طور پیوسته مشتق پذیر باشد. $V(x)$ به طور شعاعی بی‌کران گفته می‌شود اگر

$$\|x\| \rightarrow \infty \Rightarrow V(x) \rightarrow \infty.$$

لم ۱.۲.۲. [۴۹] فرض کنید $g : \Omega_1 \subset \mathbb{R}^l \rightarrow \mathbb{R}$ باشد که Ω_1 بی‌کران است. آن‌گاه همه مجموعه‌های سطح g کراندار هستند اگر فقط اگر

$$\lim_{k \rightarrow \infty} g(u^k) = +\infty,$$

$$\text{که } \{u^k\} \subset \Omega_1 \text{ و } \lim_{k \rightarrow \infty} \|u^k\| = +\infty.$$

لم ۲.۲.۲. [۴۹] فرض کنید $g : \Omega_1 \subset \mathbb{R}^l \rightarrow \mathbb{R}$ روی مجموعه بسته Ω_1 پیوسته باشد آن‌گاه همه مجموعه‌های سطح g کراندار هستند اگر و فقط اگر مجموعه کمینه‌کننده‌های g غیرتهی و کراندار باشند.

قضیه ۲.۲.۲. (قضیه نقطه ثابت بروئر^{۴۶}) [۴۹] تابع پیوسته $f : \Omega \rightarrow \Omega$ ، که Ω یک مجموعه فشرده و محدب است حداقل یک نقطه ثابت دارد.

قضیه ۳.۲.۲. [۴۸] در سیستم دینامیکی (۱.۲) فرض کنید f یک تابع پیوسته از $\mathbb{R}^n$ به $\mathbb{R}^n$ باشد، آن‌گاه برای هر $t_0 \geq 0$ و $x_0 \in \mathbb{R}^n$ ، یک جواب محلی $x(t)$ به ازای $t \in [t_0, \tau)$ که $\tau > t_0$ وجود دارد. علاوه بر این اگر f در x_0 پیوسته لیپ‌شیتز محلی باشد آن‌گاه جواب یکتاست و اگر f در $\mathbb{R}^n$ پیوسته لیپ‌شیتز باشد آن‌گاه τ می‌تواند تا $+\infty$ توسعه داده شود.

^{۴۴}Lipschitz Continuous

^{۴۵}Locally Lipschitz Continuous

^{۴۶}Brouwer's Fixed Point Theorem

تعریف ۹.۲.۲. اگر یک جواب محلی تعریف شده در بازه $[t_0, \tau)$ نتواند به یک جواب محلی روی یک بازه بزرگتر $[t_0, \tau_1)$ که $\tau_1 > \tau$ است گسترش یابد، آن‌گاه یک جواب ماکسیمال نامیده می‌شود و بازه $[t_0, \tau)$ ، بازه ماکسیمال وجود جواب نامیده می‌شود. بازه ماکسیمال وجود جواب مربوط به x_0 اغلب به صورت $[t_0, \tau(x_0))$ تعریف می‌شود.

قضیه ۴.۲.۲. [۵۰] فرض کنید $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ یک تابع پیوسته است. اگر $x(t)$ ، $t \in [t_0, \tau(x_0))$ یک جواب ماکسیمال باشد و $\tau(x_0) < +\infty$ آن‌گاه

$$\lim_{t \rightarrow \tau(x_0)} \|x(t)\| = +\infty.$$

قضیه ۵.۲.۲. (قضیه پایداری لیاپانوف) [۵۱] فرض کنید که $x = 0$ یک نقطه تعادل سیستم (۱.۲) است و $V: \mathbb{R}^n \rightarrow \mathbb{R}$ یک تابع به‌طور پیوسته مشتق‌پذیر است به‌طوری‌که

$$1. \quad V(0) = 0$$

$$2. \quad \text{به ازای هر } x \in \mathbb{R}^n \setminus \{0\}, \quad V(x) > 0$$

$$3. \quad \text{به ازای هر } x \in \mathbb{R}^n \setminus \{0\}, \quad \dot{V}(x) \leq 0$$

آن‌گاه سیستم در $x = 0$ پایدار است.

قضیه ۶.۲.۲. (قضیه پایداری مجانبی) [۵۱] تحت شرایط قضیه ۵.۲.۲ اگر $V(0)$ در شرایط زیر صدق کند

$$1. \quad V(0) = 0$$

$$2. \quad \text{به ازای هر } x \in \mathbb{R}^n \setminus \{0\}, \quad V(x) > 0$$

$$3. \quad \text{به ازای هر } x \in \mathbb{R}^n \setminus \{0\}, \quad \dot{V}(x) < 0$$

آن‌گاه سیستم در $x = 0$ پایدار مجانبی است.

قضیه ۷.۲.۲. (قضیه پایداری مجانبی سراسری) [۵۱] تحت شرایط قضیه ۵.۲.۲ اگر $V(0)$ به‌طور شعاعی بی‌کران باشد آن‌گاه نقطه $x = 0$ پایدار مجانبی سراسری است.

قضیه ۸.۲.۲. (اصل تغییرناپذیری لسال^{۴۷}) [۵۱] فرض کنید $V: \mathbb{R}^n \rightarrow \mathbb{R}$ یک تابع به‌طور پیوسته مشتق‌پذیر باشد، همچنین فرض کنید

$$1. \quad M \subset \mathbb{R}^n \text{ یک مجموعه فشرده و پایدار نسبت به جواب سیستم (۱.۲) باشد.}$$

$$2. \quad \text{به ازای } x \in M, \quad \dot{V}(x) \leq 0.$$

⁴⁷LaSalle Principle of Invariance

۳. E مجموعه همه نقاط M باشد به طوری که $\dot{V}(x) = 0$.

۴. N بزرگترین مجموعه پایدار در E باشد.

آن گاه به ازای هر $x(t_0) \in M$ که $t_0 \geq 0$ است داریم

$$\lim_{t \rightarrow \infty} \text{dist}(x(t), N) = 0.$$

فصل ۳

مسئله کنترل بهینه و شرایط بهینگی

۱.۳ مسئله کنترل بهینه

کنترل بهینه شاخه‌ای از ریاضیات است که به مسائل کاربردی می‌پردازد و با اصول تحلیل و طراحی سیستم‌های کنترل سروکار دارد. کنترل بهینه در بسیاری از رشته‌های مختلف از قبیل مهندسی، اقتصاد و علوم زیستی کاربردهای زیادی دارد. در کنترل بهینه منظور از کنترل این است که عاملی روی نحوه عملکرد سیستم تاثیر بگذارد تا عملکرد مطلوبی حاصل شود. این اثرگذاری و کنترل می‌تواند توسط خود سیستم و یا شخصی خارج و مستقل از سیستم انجام شود. نظریه کنترل علاوه بر سیستم‌های خطی در سیستم‌های غیرخطی نیز کاربرد دارد. در یک مسئله کنترل بهینه، هدف پیدا کردن ورودی یک سیستم به عنوان کنترل کننده است که در شرایط فیزیکی دستگاه صدق کند و تابع هدف را کمینه یا بیشینه کند.

۱.۱.۳ مسئله کنترل بهینه کلاسیک

شکل کلی یک مسئله کنترل بهینه با مشتق مرتبه صحیح به صورت زیر است

$$\text{Minimize } J = h(x(t_f), t_f) + \int_{t_0}^{t_f} f_0(x(t), u(t), t) dt, \quad (1.3)$$

$$\text{subject to } \dot{x} = g(x(t), u(t), t), \quad (2.3)$$

$$x(t_0) = x_0, \quad x(t_f) = x_f, \quad (3.3)$$

که $x(t) = (x_1(t), \dots, x_p(t))^T \in \mathbb{R}^p$ بردار حالت و $u(t) = (u_1(t), \dots, u_q(t))^T \in \mathbb{R}^q$ بردار کنترل نامیده می‌شود.

حساب تغییرات و معادله اویلر-لاگرانژ

حساب تغییرات به عنوان شاخه‌ای از ریاضیات نوعی تعمیم محاسبات است و در حل مسائل بهینه‌سازی بسیار مفید است. در حساب تغییرات هدف یافتن مجموعه‌ای از مسیرها، خم‌ها و غیره است که به عنوان توابعی پیوسته و مشتق‌پذیر دارای اکسترمم هستند که از مسئله پیشنهاد شده (مسئله بارچيستوچرون)^۱ در سال ۱۶۹۶ به وسیله ریاضی‌دان سوئیسی جان برنولی^۲ ناشی شده است [۵۲].

تعریف ۱.۱.۳. [۵۳] نمو تابعی را می‌توان به صورت زیر تعریف کرد

$$\Delta J(x, \delta x) = \delta J(x, \delta x) + g(x, \delta x) \|\delta x\|,$$

که δJ قسمت خطی بر حسب δx است. اگر $\lim_{\|\delta x\| \rightarrow 0} g(x, \delta x) = 0$ ، آن‌گاه J را دیفرانسیل پذیر^۳ گویند و δJ تغییرات تابعی J نامیده می‌شود.

قضیه ۱.۱.۳. [۵۳] (قضیه اساسی حساب تغییرات). اگر x^* یک منحنی اکسترمم باشد، تغییرات J روی x^* باید صفر شود، یعنی به ازای هر δx قابل قبول $\delta J(x^*, \delta x) = 0$.

لم ۱.۱.۳. [۵۳] (لم اساسی حساب تغییرات). اگر تابع h پیوسته بوده و برای هر تابع پیوسته مانند δx در محدوده $[t_0, t_f]$ داشته باشیم

$$\int_{t_0}^{t_f} h(t) \delta x(t) dt = 0,$$

آن‌گاه h باید در تمام محدوده $[t_0, t_f]$ صفر باشد.

¹Barchistochrone

²John Bernoulli

³Differentiable

تابعی زیر را با فرض اینکه f_0 دارای مشتقات نسبی مرتبه اول و دوم پیوسته نسبت به هر یک از متغیرهای خود است و t_0 و t_f ثابت و نقاط انتهایی منحنی به صورت x_0 و x_f هستند را در نظر بگیرید

$$J(x) = \int_{t_0}^{t_f} f_0(x(t), \dot{x}(t), t) dt. \quad (4.3)$$

در جستجوی یک منحنی هستیم (در صورت وجود) که $J(x)$ را به اکسترمم خود ببرد. تغییرات $\delta J(x, \delta x)$ به صورت زیر به دست می آید

$$\delta J(x, \delta x) = \int_{t_0}^{t_f} \left\{ \left[\frac{\partial f_0}{\partial x}(x(t), \dot{x}(t), t) \right] \delta x(t) + \left[\frac{\partial f_0}{\partial \dot{x}}(x(t), \dot{x}(t), t) \right] \delta \dot{x} \right\} dt.$$

اگر منحنی اکسترممی مانند x^* را در نظر بگیرید با اعمال قضیه اساسی حساب تغییرات داریم

$$\delta J(x^*, \delta x) = 0 = \int_{t_0}^{t_f} \left\{ \frac{\partial f_0}{\partial x}(x^*(t), \dot{x}^*(t), t) - \frac{d}{dt} \left[\frac{\partial f_0}{\partial \dot{x}}(x^*(t), \dot{x}^*(t), t) \right] \right\} \delta x(t) dt.$$

حال با اعمال لم اساسی حساب تغییرات شرط لازم برای اینکه x^* به ازای هر $t \in [t_0, t_f]$ اکسترمم باشد به صورت زیر به دست می آید

$$\frac{\partial f_0}{\partial x}(x^*(t), \dot{x}^*(t), t) - \frac{d}{dt} \left[\frac{\partial f_0}{\partial \dot{x}}(x^*(t), \dot{x}^*(t), t) \right] = 0. \quad (5.3)$$

معادله (۵.۳) شرط لازم بهینگی برای (۴.۳) است و معادله‌ی اوایلر-لاگرانژ نامیده می شود [۵۲].

شرایط لازم برای کنترل بهینه کلاسیک و اصل کمینه پونتریاگین

هدف تعیین کنترل قابل قبول مانند u^* است که باعث شود دستگاه (۳.۳) با طی کردن منحنی مسیر قابل قبول x^* تابعی معیار زیر را کمینه کند

$$J(u) = h(x(t_f), t_f) + \int_{t_0}^{t_f} f_0(x(t), u(t), t) dt. \quad (6.3)$$

در بخش قبل نشان داده شد شرط لازم برای اینکه x^* به ازای هر $t \in [t_0, t_f]$ تابعی $J(x)$ را اکسترمم کند معادله اوایلر-لاگرانژ است. لذا برای تشخیص اینکه x^* منحنی بهینه است باید شرایط زیر برقرار باشند

(۱) اگر t_f مشخص و x_f آزاد باشد شرایط لازم برای اینکه x^* ، اکسترمم تابعی

$$J(x) = \int_{t_0}^{t_f} f_0(x(t), \dot{x}(t), t) dt,$$

باشد عبارت است از برقراری معادله اوایلر و همچنین برقراری معادله

$$\frac{\partial f_0}{\partial \dot{x}}(x^*(t_f), \dot{x}^*(t_f), t_f) = 0.$$

این معادله شرط مرزی طبیعی^۴ نامیده می شود.

⁴Natural Boundary Condition

(۲) اگر t_f آزاد و x_f مشخص باشد، شرایط لازم عبارت است از برقراری معادله اوایلر و برقراری معادله

$$f_0(x^*(t_f), \dot{x}^*(t_f), t_f) - \left[\frac{\partial f_0}{\partial \dot{x}}(x^*(t_f), \dot{x}^*(t_f), t_f) \right] \dot{x}^*(t_f) = 0. \quad (7.3)$$

(۳) در مسائلی که t_f و x_f هر دو آزاد هستند دو حالت زیر را در نظر می‌گیریم

(الف) t_f و x_f مرتبط نیستند. در این صورت شرایط لازم عبارت است از برقراری معادله اوایلر و برقراری معادله $f_0(x^*(t_f), \dot{x}^*(t_f), t_f) = 0$

(ب) t_f و x_f مرتبط هستند، به عنوان مثال $x_f = \theta(t_f)$. در این حالت علاوه بر برقراری معادله اوایلر، معادله

$$\left[\frac{\partial f_0}{\partial \dot{x}}(x^*(t_f), \dot{x}^*(t_f), t_f) \right] \left[\frac{d\theta}{dt}(t_f) - \dot{x}^*(t_f) \right] + f_0(x^*(t_f), \dot{x}^*(t_f), t_f) = 0,$$

که شرط اربیبی^۵ نامیده می‌شود نیز باید برقرار باشد.

با اعمال قضیه اساسی حساب تغییرات و تعریف تابع $H(x(t), u(t), \lambda(t), t)$ به نام تابع همیلتونی به صورت

$$H(x(t), u(t), \lambda(t), t) \triangleq f_0(x(t), u(t), t) + \lambda^T g(x(t), u(t), t), \quad (8.3)$$

که λ ضرب‌گر لاگرانژ است [۵۲]، شرایط لازم عبارت است از

$$\frac{\partial H}{\partial x}(x^*(t), u^*(t), \lambda^*(t), t) = -\dot{\lambda}(t), \quad (9.3)$$

$$\frac{\partial H}{\partial u}(x^*(t), u^*(t), \lambda^*(t), t) = 0, \quad (10.3)$$

$$\frac{\partial H}{\partial \lambda}(x^*(t), u^*(t), \lambda^*(t), t) = \dot{x}(t), \quad (11.3)$$

$$\left[\frac{\partial h}{\partial x}(x^*(t_f), t_f) - \lambda^*(t_f) \right]^T \delta x_f + \left[H(x^*(t_f), u^*(t_f), \lambda^*(t_f), t_f) + \frac{\partial h}{\partial t}(x^*(t_f), t_f) \right] \delta t_f = 0. \quad (12.3)$$

قضیه ۲.۱.۳. (اصل کمینه پونتریاگین) [۵۲] فرض کنید $u^*(t)$ یک کنترل قابل قبول است و مسیر متناظر آن نیز $x^* = (x_1^*, x_2^*)$ است که دستگاه را از x^0 در زمان $t = t_0$ به x^1 در زمان t_1 هدایت می‌کند. برای اینکه u^* و x^* بهینه باشد (یعنی J را کمینه کند) لازم است که بردار غیرصفر $\lambda = (\lambda_0, \lambda_1, \lambda_2)^T$ در $\lambda_i = -\frac{\partial H}{\partial x_i}$ ، $i = 0, 1, 2$ صدق کند و نیز تابع عددی

$$H(\lambda, x, u) = \lambda_0 f_0(x, u) + \lambda_1 f_1(x, u) + \lambda_2 f_2(x, u), \quad (13.3)$$

موجود باشد به طوری که

⁵Transversality Condition

(الف) به ازای هر $t_0 \leq t \leq t_1$ ، H به کمینه خود نسبت به u در $u = u^*(t)$ برسد.

(ب) $H(\lambda^*, x^*, u^*) = 0$ و در $t = t_1$ ، $\lambda \leq 0$ ، که $\lambda^*(t)$ جواب $\dot{\lambda}_i = -\frac{\partial H}{\partial x_i}$ ، $i = 1, 2$ به ازای $u = u^*$ است.

به علاوه می توان نشان داد که، عدد ثابت $H(\lambda^*, x^*, u^*) = 0$ ، و نیز، عدد ثابت $\lambda_0(t) = 0$. بنابراین به ازای هر نقطه روی مسیر بهینه، $H = 0$ و $\lambda_0 \leq 0$.

۲.۱.۳ مسئله کنترل بهینه کسری و شرایط لازم بهینگی

مسئله کنترل بهینه کسری تعمیمی از مسئله کنترل بهینه کلاسیک است که تحت سیستم های دینامیکی شامل مشتقات کسری است و امروزه یکی از موضوعات بسیار جالب و مورد توجه در سیستم های کسری است. یک مسئله کنترل بهینه کسری می تواند نسبت به مشتق های مختلفی تعریف شود که مهم ترین آن ها مشتق های کسری کاپاتو و ریمان-لیوویل است. ابتدا فرم کلی مسئله کنترل بهینه کسری با مرتبه ثابت را به صورت زیر تعریف می کنیم و سپس قضایای شرایط بهینگی بیان می شود.

$$\text{Minimize} \quad J(x, u) = \psi(T, x(T)) + \int_0^T f(t, x(t), u(t)) dt, \quad (14.3)$$

$$\text{subject to} \quad M \dot{x}(t) + N {}^\circ D_t^\alpha x(t) = g(t, x(t), u(t)), \quad (15.3)$$

$$x(0) = x_0, \quad (16.3)$$

که $M, N \in \mathbb{R}^p$ ، $x_0 \in \mathbb{R}^p$ ، $0 < \alpha < 1$ و $(M, N) \neq (0, 0)$ هستند و

$$f, g : \mathbb{R} \times \mathbb{R}^p \times \mathbb{R}^q \rightarrow \mathbb{R}, \quad \psi : \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}^p,$$

توابعی مشتق پذیر هستند. $x(\cdot)$ و $u(\cdot)$ به ترتیب متغیرهای وضعیت و کنترل هستند و $\dot{x}(t)$ مشتق معمولی را نشان می دهد. حال اگر در (۱۵.۳) داشته باشیم ${}^\circ D_t^\alpha x(t)$ ، مشتق کسری چپ کاپوتو از مرتبه α و اگر داشته باشیم ${}^\circ D_t^\alpha x(t)$ ، مشتق کسری چپ ریمان لیوویل از مرتبه α را نشان می دهد. در مسئله (۱۴.۳)–(۱۶.۳)، T و $x(T)$ می توانند ثابت و یا آزاد باشند. بنابراین توجه در این مسئله، علاوه بر متغیر وضعیت $x(\cdot)$ و متغیر کنترل $u(\cdot)$ ، روی زمان نهایی T مورد نظر نیز معطوف است تا بتوان تابعی J را با توجه به حالت های T و $x(T)$ بهینه کرد. در مسئله (۱۴.۳)–(۱۶.۳)، اگر $N = 0$ باشد مسئله کنترل بهینه کلاسیک را خواهیم داشت.

تعمیم شرایط لازم بهینگی برای مسائل کنترل بهینه کسری با مشتق کاپاتو

برای اینکه شرایط لازم بهینگی نتیجه شود یعنی جواب بهینه سه تایی (x, u, T) در مسئله کنترل بهینه کسری (۱۴.۳)–(۱۶.۳) صدق کند، ضریب لاگرانژ $\lambda(\cdot)$ را به کار برده و تابع همیلتونی

H را به صورت زیر تعریف کنید [۵۴، ۵۵]

$$H(t, x, u, \lambda) = f(t, x, u) + \lambda g(t, x, u). \quad (۱۷.۳)$$

قضیه ۳.۱.۳. [۵۵] فرض کنید (x, u, T) جواب بهینه برای مسئله (۱۴.۳) – (۱۶.۳) است. در این صورت تابع $\lambda(\cdot)$ وجود دارد به طوری که سه تایی (x, u, λ) در شرایط زیر صدق می کند (الف) دستگاه همیلتونی

$$M\dot{\lambda}(t) - N_t D_T^\alpha \lambda(t) = \frac{-\partial H}{\partial x}(t, x(t), u(t), \lambda(t)), \quad t \in [0, T], \quad (۱۸.۳)$$

$$M\dot{x}(t) + N_0^C D_t^\alpha x(t) = \frac{\partial H}{\partial \lambda}(t, x(t), u(t), \lambda(t)), \quad t \in [0, T], \quad (۱۹.۳)$$

(ب) شرایط ایستایی

$$\frac{\partial H}{\partial u}(t, x(t), u(t), \lambda(t)) = 0, \quad t \in [0, T], \quad (۲۰.۳)$$

(ج) شرایط تراگردی

$$\begin{aligned} & \left[H((t, x(t), u(t), \lambda(t)) - N\lambda(t)_0^C D_t^\alpha x(t) + N\dot{x}(t)_t I_T^{1-\alpha} \lambda(t) + \frac{\partial \psi}{\partial t}(t, x(t)) \right]_{t=T} = 0, \\ & \left[M\lambda(t) + N_t I_T^{1-\alpha} \lambda(t) - \frac{\partial \psi}{\partial x}(t, x(t)) \right]_{t=T} = 0, \end{aligned} \quad (۲۱.۳)$$

که همیلتونی H در (۱۷.۳) تعریف شده است.

با توجه به زمان نهایی T و مکان نهایی $x(T)$ ، حالت های خاص جالبی برای شرایط بهینگی به وجود می آید که نتیجه زیر چند مورد از این حالت ها را شامل می شود.

نتیجه ۱.۱.۳. فرض کنید (x, u) ، جواب بهینه برای مسئله کنترل بهینه کسری (۱۴.۳) – (۱۶.۳) است.

(الف) اگر T ثابت و $x(T)$ آزاد باشد، آن گاه قضیه ۳.۱.۳ با شرایط تراگردی زیر برقرار است که جایگزین (۲۱.۳) می شود

$$\left[M\lambda(t) + N_t I_T^{1-\alpha} \lambda(t) - \frac{\partial \psi}{\partial x}(t, x(t)) \right]_{t=T} = 0.$$

(ب) اگر T آزاد و $x(T)$ ثابت باشد، آن گاه قضیه ۳.۱.۳ با شرایط تراگردی زیر برقرار است که جایگزین (۲۱.۳) می شود

$$\left[H(t, x(t), u(t), \lambda(t)) - N\lambda(t)_0^C D_t^\alpha x(t) + N\dot{x}(t)_t I_T^{1-\alpha} \lambda(t) + \frac{\partial \psi}{\partial t}(t, x(t)) \right]_{t=T} = 0.$$

(ج) اگر T و $x(T)$ هر دو ثابت باشند، آن گاه قضیه ۳.۱.۳ بدون شرایط تراگردی برقرار خواهد بود.

(د) اگر مکان نهایی متعلق به یک منحنی ثابت باشد یعنی $x(T) = \gamma(T)$ ، که γ منحنی مشتق پذیر است، آن گاه قضیه ۳.۱.۳ با شرایط تراگردی زیر برقرار است که جایگزین (۲۱.۳) می شود

$$\left[H(t, x(t), u(t), \lambda(t)) - N\lambda(t) {}^C D_t^\alpha x(t) + N\dot{x}(t) {}^I_T^{1-\alpha} \lambda(t) + \frac{\partial \psi}{\partial t}(t, x(t)) - \dot{\gamma}(t) \left(M\lambda(t) + N {}^I_T^{1-\alpha} \lambda(t) - \frac{\partial \psi}{\partial t}(t, x(t)) \right) \right]_{t=T} = 0.$$

(ه) اگر T ثابت و برای برخی مقادیر $K, K \in \mathbb{R}$ باشد $x(T) \geq K$ ، با شرایط تراگردی زیر برقرار است که جایگزین (۲۱.۳) می شود

$$\left[M\lambda(t) + N {}^I_T^{1-\alpha} \lambda(t) - \frac{\partial \psi}{\partial x}(t, x(t)) \right]_{t=T} \leq 0, \\ (x(T) - K) \left[M\lambda(t) + N {}^I_T^{1-\alpha} \lambda(t) - \frac{\partial \psi}{\partial x}(t, x(t)) \right]_{t=T} = 0.$$

(و) اگر $x(T)$ ثابت باشد و برای برخی مقادیر $K, K \in \mathbb{R}$ باشد $T \leq K$ ، با شرایط تراگردی زیر برقرار است که جایگزین (۲۱.۳) می شود

$$\left[H(t, x(t), u(t), \lambda(t)) - N\lambda(t) {}^C D_t^\alpha x(t) + N\dot{x}(t) {}^I_T^{1-\alpha} \lambda(t) + \frac{\partial \psi}{\partial t}(t, x(t)) \right]_{t=T} \geq 0, \\ \left[H(t, x(t), u(t), \lambda(t)) - N\lambda(t) {}^C D_t^\alpha x(t) + N\dot{x}(t) {}^I_T^{1-\alpha} \lambda(t) + \frac{\partial \psi}{\partial t}(t, x(t)) \right]_{t=T} \\ \times (T - K) = 0.$$

تعمیم شرایط لازم بهینگی برای مسائل کنترل بهینه کسری با مشتق ریمن-لیوویل

در اینجا مسئله کنترل بهینه کسری

$$\begin{aligned} \text{Minimize} \quad & J(x, u) = \int_0^1 F(x(t), u(t), t) dt, \\ \text{subject to} \quad & {}^0 D_t^\alpha x(t) = G(x(t), u(t), t), \\ & x(0) = x_0, \end{aligned}$$

در شرایط مشتق کسری چپ و راست ریمن - لیوویل فرمول بندی می شود [۵۶]. برای یافتن شرایط لازم بهینگی از رهیافت سنتی قبل تابع هدف را به صورت زیر در نظر بگیرید

$$\bar{J}(u) = \int_0^1 [F(x, u, t) + \lambda(G(x, u, t) - {}^0 D_t^\alpha x)] dt, \quad (22.3)$$

که λ ضربگر لاگرانژ است. با در نظر گرفتن تغییرات تابعی (۲۲.۳) داریم

$$\delta \bar{J}(u) = \int_0^1 \left[\frac{\partial F}{\partial x} \delta x + \frac{\partial F}{\partial u} \delta u + \delta \lambda (G(x, u, t) - {}_0 D_t^\alpha x) + \lambda \left(\frac{\partial G}{\partial x} \delta x + \frac{\partial G}{\partial u} \delta u - \delta({}_0 D_t^\alpha x) \right) \right] dt, \quad (23.3)$$

که δx ، δu و $\delta \lambda$ به ترتیب تغییرات x ، u و λ هستند و طبق شرط تراگردی $\lambda(1) = 0$ است. ريو^۶ در [۵۷] نشان داد که برای $\alpha > 0$ رابطه زیر برقرار است

$$\int_a^b \frac{d^\alpha f(t)}{d(t-a)^\alpha} g(t) dt = (-1)^{-\alpha} \int_a^b f(t) \frac{d^\alpha g(t)}{d(t-b)^\alpha} dt, \quad (24.3)$$

به شرطی که برای $k = 0, \dots, n-1$ در $t = a$ و $t = b$ داشته باشیم $\frac{d^k f(t)}{dt^k} = 0$ و یا $\frac{d^k g(t)}{dt^k} = 0$ در (۲۴.۳)

$$\frac{d^\alpha f(t)}{d(t-a)^\alpha} = {}_0 D_t^\alpha f(t), \quad (25.3)$$

مشتق کسری چپ ریمن-لیوویل تابع $f(t)$ از مرتبه α است و n کوچکترین عدد صحیح بزرگتر از α است. با استفاده از نمادگذاری‌های استفاده شده، (۲۴.۳) به صورت زیر نوشته می شود

$$\int_a^b ({}_a D_t^\alpha f(t)) g(t) dt = \int_a^b f(t) ({}_t D_b^\alpha g(t)) dt. \quad (26.3)$$

رابطه (۲۶.۳) فرمول انتگرال گیری جز به جز کسری نامیده می شود که در [۵۸] آورده شده است. با استفاده از (۲۶.۳) آخرین انتگرال در (۲۳.۳) به صورت زیر نوشته می شود

$$\int_0^1 \lambda \delta({}_0 D_t^\alpha x) dt = \int_0^1 \delta x ({}_t D_1^\alpha \lambda) dt. \quad (27.3)$$

ثابت می شود که $\delta x(0) = 0$ یا $\lambda(0) = 0$ و $x(1) = 0$ یا $\lambda(1) = 0$. چون $x(0) = 0$ مشخص است، داریم $\delta x(0) = 0$ و چون $x(1)$ مشخص نیست، لذا $\lambda(1) = 0$. با این فرضیات رابطه (۲۷.۳) برقرار است. با استفاده از روابط (۲۳.۳) و (۲۷.۳) داریم

$$\delta \bar{J}(u) = \int_0^1 \left[\delta \lambda (G(x, u, t) - {}_0 D_t^\alpha x) + \delta x \left(\frac{\partial F}{\partial x} + \lambda \frac{\partial G}{\partial x} - {}_t D_1^\alpha \lambda \right) + \delta u \left(\frac{\partial F}{\partial u} + \lambda \frac{\partial G}{\partial u} \right) \right] dt. \quad (28.3)$$

مقدار کمینه $\bar{J}(u)$ (و در نتیجه مقدار کمینه $J(u)$) زمانی اتفاق می افتد که ضرایب $\delta \lambda$ ، δx و δu در رابطه (۲۸.۳) صفر شود. در نتیجه شرایط لازم بهینگی به صورت زیر به دست می آید

$$\begin{cases} {}_0 D_t^\alpha x = G(x, u, t), \\ {}_t D_1^\alpha \lambda = \frac{\partial F}{\partial x} + \lambda \frac{\partial G}{\partial x}, \\ \frac{\partial F}{\partial u} + \lambda \frac{\partial G}{\partial u} = 0, \\ x(0) = x_0, \quad \lambda(1) = 0. \end{cases}$$

۳.۱.۳ مسئله کنترل بهینه از مرتبه متغیر و شرایط بهینگی

مسئله کنترل بهینه از مرتبه متغیر را به صورت زیر در نظر می‌گیرید

$$\begin{aligned} J(u) &= \int_a^b \varphi(x, u, t) dt, \\ {}_a D_t^{\alpha(t)} x(t) &= \psi(x, u, t), \\ x(a) &= x_a, \quad x(b) = x_b. \end{aligned} \quad (29.3)$$

مشتق کسری مرتبه متغیر ارائه شده در اینجا مشتق کسری ریمن-لیوویل است و می‌توان مشتق‌های کسری دیگر هم در نظر گرفت. با توجه به [۵۹] تابع همیلتونی برای مسئله (۲۹.۳) به صورت زیر است

$$H(x, u, \lambda, t) = \varphi(x, u, t) + \lambda \psi(x, u, t), \quad (30.3)$$

که λ ضرب‌گر لاگرانژ است. شرایط لازم بهینگی برای مسئله کنترل بهینه (۲۹.۳) با مرتبه متغیر به صورت

$$\begin{cases} {}_a D_t^{\alpha(t)} x(t) = \frac{\partial H}{\partial \lambda}, \\ {}_t D_b^{\alpha(t)} \lambda(t) = \frac{\partial H}{\partial x}, \\ \frac{\partial H}{\partial u} = 0, \end{cases} \quad (31.3)$$

است.

۲.۳ بررسی روش‌های عددی برای مسائل کنترل بهینه کسری

در بیشتر موارد جواب‌های تحلیلی برای مسائل کسری ممکن است به آسانی به دست نیاید و اغلب از روش‌های عددی استفاده کنند که ثابت شده است که این روش‌ها معتبر و کارا هستند. در دهه‌های اخیر روش‌های عددی بسیاری برای حل مسائل کنترل بهینه ارائه شده است. رویکردها برای حل عددی مسائل کنترل بهینه کسری ممکن است به دو دسته بزرگ تقسیم شوند: روش‌های مستقیم و روش‌های غیرمستقیم. روش مستقیم، مسائل کنترل بهینه را برای سیستم‌های غیرخطی و کاملاً پیوسته حل می‌کند که بر پایه یک گسسته‌سازی نسبت به زمان است. این بدان معناست که مسئله برای یافتن متغیرهای کنترل و وضعیت بهینه، به برنامه‌ریزی غیرخطی با ارزش‌گذاری مقادیر کنترل و وضعیت فقط در یک تعداد متناهی از نقاط هم‌محلی زمانی تبدیل می‌شود. با تغییر دادن مقادیر کنترل و وضعیت در این نقاط، تابع هزینه به طور مستقیم بهینه می‌شود. مسئله به دست آمده می‌تواند با روش‌های برنامه‌ریزی غیرخطی حل شود.

روش‌های غیرمستقیم بر پایه جواب‌هایی هستند که در معادلات اوایلر-لاگرانژ که شرایط لازم بهینگی هستند و یا در معادلات همیلتونی-ژاکوبی-بلمن^۷ که شرایط کافی بهینگی هستند صدق کنند. در این روش، به جای کمینه کردن تابع هزینه در روش مستقیم، جواب بهینه، از صادق بودن در شرایط بهینگی به دست می‌آید. با توجه به مسئله کنترل بهینه داده شده، شرایط بهینگی منجر به دستیابی به مسئله با مقادیر کران دونقطه‌ای می‌شود که شامل معادلات دیفرانسیل برای متغیر وضعیت پیوسته و متغیرهای الحاقی پیوسته، معادله جبری برای کنترل پیوسته و شرایط کران‌داری برای متغیرهای وضعیت، متغیرهای الحاقی و زمان است. در برخی موارد خاص پیوستگی کامل سیستم‌های خطی و هزینه‌های درجه دوم، می‌توان یک راه‌حل تحلیلی برای مسئله با شرایط مرزی پیدا کرد. روش‌های غیرمستقیم بر پایه بهینگی سپس گسسته‌سازی مسئله کنترل بهینه اصلی هستند در حالی که روش‌های مستقیم بر پایه گسسته‌سازی و سپس بهینگی مسئله کنترل بهینه اصلی هستند.

برای برخی از کارهای عددی چه به صورت مستقیم و چه به صورت غیرمستقیم می‌توان به [۶۰]–[۷۴] رجوع کرد. در میان آن‌ها نویسندگان در [۶۱] یک فرم کلی از مسئله کنترل بهینه کسری را ارائه کردند و روش حل آن‌ها بر پایه پیدا کردن جواب عددی معادله همیلتونی-ژاکوبی-بلمن است که توسط روش هم‌محلی^۸ گاوس-لژاندر پایه‌ریزی شده است. در [۶۲] یک رده‌ای از مسائل کنترل بهینه کسری که در آن سیستم‌های دینامیکی شامل مشتق‌های کسری و کلاسیک هستند مطرح شد. در این کار نویسندگان برای حل عددی از چندجمله‌ای‌های لاگر استفاده کردند. در [۶۵] یک روش عددی برای حل مسائل کنترل بهینه کسری با کمک موجک‌ها^۹ مطالعه شده است. در [۷۵، ۷۶] می‌توان چند روش عددی برای حل مسائل کسری از مرتبه متغیر را مشاهده کرد. به طور معمول این روش‌های عددی به صورت یک گسسته‌سازی حول زمان یا جایگزین کردن عملگرهای کسری با یک تجزیه مناسب صورت می‌گیرد.

گرچه این روش‌ها تقریب‌های خوبی برای جواب فراهم می‌کنند اما آن‌ها نیاز به گسسته‌سازی دامنه از طریق مش‌بندی دارند که ممکن است در ابعاد دو یا بالاتر چالش برانگیز باشد. جواب‌های تقریبی به دست آمده ناپیوسته هستند و به طور جدی روی پایداری جواب تاثیر می‌گذارند. همچنین برای به دست آوردن جواب‌های مناسب و مطلوب ممکن است نیاز باشد تا مش‌بندی‌های زیادی انجام شود که این خود به طور قابل توجهی هزینه‌های محاسباتی را افزایش می‌دهد. از طرف دیگر برخی از روش‌های عددی از ماتریس‌های عملگر برای حل استفاده می‌کنند در حالی که پیدا کردن ماتریس‌های عملگر معمولاً سخت است و این ماتریس‌های عملگر در مرتبه‌های بالاتر باعث پیچیدگی محاسباتی نیز خواهند شد. یک رویکرد امیدوار کننده برای غلبه بر این محدودیت‌ها استفاده از شبکه‌های عصبی مصنوعی بر پایه

⁷Hamilton-Jacobi-Bellman

⁸Collocation Method

⁹Haar Wavelet

یک توپولوژی شبکه، الگوی اتصال، خواص فعال‌سازی عصبی، استراتژی یادگیری و توانایی پردازش داده‌هاست. جواب‌های به‌دست آمده از طرح‌های شبکه عصبی مشتق‌پذیر و نزدیک به جواب‌های تحلیلی هستند.

بسیاری منابع درباره قضایا، کاربردها، مدل‌بندی، طراحی و ساختار ریاضی از شبکه‌های عصبی وجود دارند (مراجع [۷۷]–[۷۹] را ببینید). همچنین حل عددی معادلات دیفرانسیل معمولی و جزئی [۸۰]–[۸۲]، مسائل کنترل بهینه [۸۳]–[۸۶]، مسائل برنامه ریزی خطی [۸۷]–[۹۴] و حل عددی معادلات دیفرانسیل کسری و کنترل بهینه کسری [۶۰، ۹۵، ۹۶] با ساختار شبکه عصبی را می‌توان در این مراجع ذکر شده مشاهده کرد. در میان آن‌ها نویسندگان در [۶۰] یک شبکه عصبی پرسپترون ارائه کردند تا جواب مسئله کنترل بهینه کسری را تقریب بزنند. شرایط بهینگی منجر به دست‌یابی به یک مسئله با مقادیر کران دونقطه‌ای شد. همچنین نویسندگان در [۹۵] یک کاربرد از شبکه‌های عصبی برای حل معادلات دیفرانسیل معمولی از مرتبه کسری را بیان کردند. روش ارائه شده از یک سری توانی مناسب استفاده می‌کند. در [۸۵] یک رده‌ای از مسائل کنترل بهینه غیرخطی با محدودیت‌های نامساوی بحث شده است که بر پایه جواب شرایط بهینگی حل شده است. در [۹۶] جواب معادلات دیفرانسیل تاخیری کسری با استفاده از یک رویکرد جدید بر پایه شبکه عصبی مصنوعی تقریب زده شد. نویسندگان در [۹۶] معادلات دیفرانسیل کسری از مرتبه متغیر را با هسته میتاگ–لفلر بررسی کردند.

فصل ۴

مسائل کنترل بهینه کسری با شبکه عصبی و طرح‌های بهینه‌سازی دینامیکی

۱.۴ مقدمه

یک سیستم دینامیکی کسری، سیستمی است که دینامیک‌ها توسط معادلات دیفرانسیل مرتبه کسری توصیف می‌شوند و یک مسئله کنترل بهینه کسری یک مسئله کنترل بهینه با یک مشتق کسری است. این مشتق کسری می‌تواند از مرتبه ثابت یا از مرتبه متغیر باشد. در این فصل یک روش عددی غیرمستقیم بر پایه شبکه عصبی برای حل مسائل کنترل بهینه کسری از مرتبه ثابت ارائه خواهد شد. با توجه به مطالب ارائه شده در فصل قبل در مورد مسائل کنترل بهینه از مرتبه ثابت و مرتبه متغیر، این روش به راحتی قابل تعمیم برای مسائل کنترل بهینه کسری از مرتبه متغیر است. یکی از بهترین خاصیت‌های روش غیرمستقیم دقت بالای جواب‌های تقریبی به دست آمده از حل مسئله کنترل بهینه اصلی است. این خاصیت ویژه بر پایه صدق کردن در شرایط لازم بهینگی است که از حساب تغییرات و اصل کمینه پونتریاگین ناشی شده است. به عنوان یک کاربرد از روش‌های هوش محاسباتی، صبوری و عفتی [۶۰] یک شبکه عصبی را برای حل مسائل کنترل بهینه کسری با مشتق کسری کاپاتو ارائه کردند.

متناظر با شرایط لازم بهینگی آن‌ها مسئله کنترل بهینه را به یک سیستم از معادلات انتگرال و لترا تبدیل کردند و با استفاده از قاعده سیمپسون مسئله به یک سیستم از معادلات غیرخطی تبدیل شده است. پیچیدگی محاسباتی قاعده سیمپسون یکی از معایب این روش است. در روش ارائه شده در این فصل در محاسبات عددی، مشتق کسری ریمن-لیوویل با کمک مشتق کسری گرانوالد-لتنیکف تقریب زده می‌شود. بنابراین قبل از ارائه روش، عملگرهای گرانوالد-لتنیکف معرفی می‌شوند.

۲.۴ عملگرهای گرانوالد-لتنیکف

عملگرهای گرانوالد-لتنیکف بر پایه یک مفهوم کلاسیک برای مشتقات مرتبه صحیح می‌تواند به صورت محدودیت‌های تفاضل متناهی نمایش داده شود. با فرض $l \in \mathbb{N}$, $y \in C^l([a, b])$ واضح است که مشتق مرتبه l از y در $t \in (a, b]$ می‌تواند به صورت زیر به دست آید

$${}_a D_t^l y(t) = \lim_{h \rightarrow 0} \frac{1}{h^l} \sum_{k=0}^l (-1)^k \binom{l}{k} y(t - kh), \quad (1.4)$$

که ضرایب دوجمله‌ای به صورت

$$\binom{l}{k} = \frac{l(l-1)\dots(l-k+1)}{k!} = \frac{l!}{k!(l-k)!}, \quad (2.4)$$

تعریف می‌شوند. با توجه به اینکه $l \in \mathbb{N}$ داریم $\Gamma(l+1) = l!$. تابع گاما می‌تواند به عنوان یک بسط از فاکتوریل با آرگومان‌های حقیقی در نظر گرفته شود و با جایگزین کردن l با هر مقدار حقیقی مثبت α در (۲.۴)، ضرایب دوجمله‌ای (۲.۴) را می‌توان به صورت زیر بازنویسی کرد

$$\binom{\alpha}{k} = \frac{\Gamma(\alpha+1)}{\Gamma(k+1)\Gamma(\alpha-k+1)}. \quad (3.4)$$

از طرفی چون رابطه $\Gamma(z+1) = z\Gamma(z)$ این اجازه را می‌دهد که از تابع گاما برای $z < 0$ وقتی $z \notin \{0, -1, -2, -3, \dots\}$ استفاده شود پس (۲.۴) را می‌توان برای $\alpha < 0$ با $\alpha \notin \{0, -1, -2, -3, \dots\}$ به کار برد [۲۰]. بنابراین برای $l = [\alpha]$, $y \in C^l([a, b])$ این امکان وجود دارد تا (۱.۴) به مرتبه کسری $\alpha > 0$ برای $t \in (a, b]$ به صورت زیر بسط داده شود

$${}_a^G D_t^\alpha y(t) = \lim_{N \rightarrow \infty} \frac{1}{h^\alpha} \sum_{k=0}^N w_k^{(\alpha)} y(t - kh), \quad (4.4)$$

که $h = \frac{t-a}{N}$ و وزن‌های $w_k^{(\alpha)}$ به صورت

$$w_k^{(\alpha)} = (-1)^k \binom{\alpha}{k},$$

تعریف می‌شوند. همچنین وزن‌ها می‌توانند به صورت رابطه بازگشتی زیر تعریف شوند

$$w_o^{(\alpha)} = 1, \quad w_k^{(\alpha)} = \left(1 - \frac{\alpha + 1}{k}\right) w_{k-1}^{(\alpha)}. \quad (5.4)$$

وقتی $h > 0$ ، رابطه (۴.۴) مشتق چپ گرانوالد-لتنیکف و وقتی $h < 0$ ، رابطه (۴.۴) مشتق کسری راست گرانوالد-لتنیکف نامیده می‌شود. از طرفی همان‌طور که در فصل اول بیان شد مشتق کسری مرتبه متغیر گرانوالد-لتنیکف از مرتبه $\alpha(t)$ از تابع $f(t)$ به صورت

$${}_a^G D_t^{\alpha(t)} f(t) = \lim_{h \rightarrow 0} \frac{1}{h^{\alpha(t)}} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha(t)}{k} f(t - kh), \quad (6.4)$$

تعریف می‌شود. وقتی $h > 0$ ، رابطه (۶.۴) مشتق چپ گرانوالد-لتنیکف از مرتبه متغیر و وقتی $h < 0$ ، رابطه (۶.۴) مشتق کسری راست گرانوالد-لتنیکف از مرتبه متغیر نامیده می‌شود.

همچنین رابطه بازگشتی تابع وزن $w_k^{(\alpha(t))} = (-1)^k \binom{\alpha(t)}{k}$ به صورت

$$w_o^{(\alpha(t))} = 1, \quad w_k^{(\alpha(t))} = \left(1 - \frac{\alpha(t) + 1}{k}\right) w_{k-1}^{(\alpha(t))}, \quad (7.4)$$

تعریف می‌شود.

۳.۴ مسئله کنترل بهینه کسری و روش حل مسئله

مسئله کنترل بهینه زیر را در نظر بگیرید

$$\begin{aligned} &\text{minimize} && J(x(t), u(t)) = \int_{t_o}^{t_f} F(t, x(t), u(t)) dt, \\ &\text{subject to} && {}_{t_o} D_{t_f}^{\alpha} x(t) = G(t, x(t), u(t)), \\ &&& x(t_o) = x_o, \quad x(t_f) = x_f, \end{aligned} \quad (8.4)$$

که $x(t) \in \mathbb{R}^p$ متغیر وضعیت، $u(t) \in \mathbb{R}^q$ متغیر کنترل و $t \in \mathbb{R}$. فرض کنید F دارای مشتقات جزئی مرتبه اول و دوم نسبت به تمام آرگومان‌هایش و G پیوسته لیپ‌شیتز روی مجموعه $\Omega \subseteq \mathbb{R}^p$ است. همچنین $n = [\alpha_i] + 1$ ، $(i = 1, \dots, p)$ ، که $[\alpha_i]$ قسمت صحیح α_i است. با استفاده از شرایط لازم بهینگی ارائه شده در فصل ۳، اگر (x, u) یک کمینه از (۸.۴) باشد آن‌گاه یک $\lambda(t)$ وجود دارد که سه تایی (x, u, λ) در زیر صدق می‌کند

$$\begin{cases} \frac{\partial H}{\partial x}(t, x, u, \lambda) = {}_t D_{t_f}^{\alpha} \lambda(t), \\ \frac{\partial H}{\partial \lambda}(t, x, u, \lambda) = {}_{t_o} D_t^{\alpha} x(t), \\ \frac{\partial H}{\partial u}(t, x, u, \lambda) = 0, \\ x(t_o) = x_o, \quad x(t_f) = x_f, \end{cases} \quad (9.4)$$

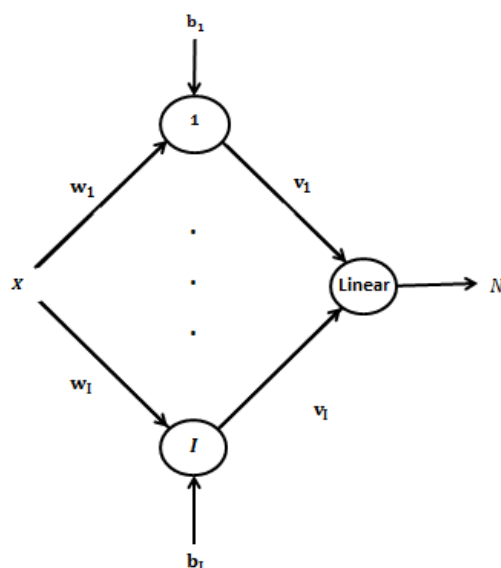
که H تابع همیلتونی است و به صورت

$$H(t, x, u, \lambda) = F(t, x, u) + \lambda G(t, x, u),$$

تعریف می‌شود. حال یک طرح تقریبی برای حل دستگاه (۹.۴) ارائه می‌دهیم. سه شبکه عصبی پیش‌خور پرسپترون به‌طور جداگانه برای هر یک از متغیرهای وضعیت (شبکه عصبی آن N_x)، ضرب‌گرهای لاگرانژ (شبکه عصبی آن N_λ) و کنترل (شبکه عصبی آن N_u) برای $i = 1, 2, \dots, I$ به‌صورت زیر تعریف می‌کنیم

$$\begin{cases} N_x(t) = \sum_{i=1}^I v_x^i \sigma(z_x^i), & z_x^i = w_x^i t + b_x^i, \\ N_\lambda(t) = \sum_{i=1}^I v_\lambda^i \sigma(z_\lambda^i), & z_\lambda^i = w_\lambda^i t + b_\lambda^i, \\ N_u(t) = \sum_{i=1}^I v_u^i \sigma(z_u^i), & z_u^i = w_u^i t + b_u^i, \end{cases} \quad (10.4)$$

که I تعداد نرون‌هاست که می‌تواند برای هر شبکه متفاوت باشد، w^i ، پارامتر وزن از ورودی به i -امین واحد پنهان، v^i ، i -امین پارامتر وزن از لایه پنهان i به لایه خروجی، b^i ، i -امین بایاس برای i واحد در لایه پنهان، z^i ، ورودی از i -امین واحد پنهان و σ یک تابع فعال‌ساز دلخواه است (شکل ۳.۴ را ببینید). در مثال‌های عددی ارائه شده در این فصل از رساله تابع



شکل ۱۰.۴: ساختار شبکه عصبی پرسپترون

سیگموئید به‌صورت

$$\sigma(x) = \frac{1}{1 + e^{-x}}, \quad (11.4)$$

استفاده شده است. بنابر قضیه کولوموگوروف، می‌توان هر تابع پیوسته را با پرسپترون چندلایه ایجاد کرد. به وسیله این قضیه می‌توان از هر شبکه عصبی دلخواه در تخمین تابع استفاده کرد [۴۷]. حال می‌توان رابطه‌ای بین جواب‌های (۹.۴) و جواب‌های یک مسئله کمینه‌سازی نامقید معادل از طریق جواب‌های آزمایشی^۱ پیدا کرد [۹۷]. این جواب‌های آزمایشی به‌صورت

¹trial solutions

مجموع دو بخش نوشته می‌شوند. بخش اول شامل پارامترهای قابل تنظیم نیست و در شرایط اولیه و مرزی صدق می‌کند. بخش دوم شامل یک شبکه عصبی پیش‌خور با پارامترهای قابل تنظیم است. لازم به ذکر است که در قسمت دوم وزن‌ها طوری تنظیم شده‌اند که فقط با مسئله کمینه‌سازی سروکار داشته باشند و در شرایط مرزی مقدار آن‌ها صفر شود. بر پایه این حقیقت و شبکه‌های عصبی پیشنهاد شده، جواب‌های آزمایشی برای تابع‌های وضعیت، ضرب‌گرهای لاگرانژ و کنترل به ترتیب در ساختار زیر تعریف می‌شوند

$$\begin{cases} x_T(t) = \frac{(t - t_f)x(t_0) - (t - t_0)x(t_f)}{t_0 - t_f} + (t - t_0)(t - t_f)N_x(t), \\ \lambda_T(t) = N_\lambda(t), \\ u_T(t) = N_u(t). \end{cases} \quad (12.4)$$

به آسانی می‌توان بررسی کرد که x_T در شرایط اولیه صدق می‌کند، یعنی $x_T(t_0) = x_0$ و $x_T(t_f) = x_f$. نکته این است که برای نقاط انتهایی آزاد داریم $\lambda(\cdot) = 0$. برای مثال اگر $x(t_0)$ آزاد باشد باید داشته باشیم $\lambda(t_0) = 0$ و بنابراین می‌توان λ_T در (12.4) را به صورت $\lambda_T = (t - t_0)N_\lambda$ تعریف کرد. برای هر شرایط اولیه و مرزی دیگر نیز می‌توان جواب‌های آزمایشی مناسب را ساخت. جواب‌های تقریبی (12.4) تقریب‌های جامع هستند و باید در (9.4) صدق کنند. بنابراین

$$\begin{cases} \frac{\partial H_T}{\partial x_T} = {}_t D_{t_f}^\alpha \lambda_T(t), \\ \frac{\partial H_T}{\partial \lambda_T} = {}_{t_0} D_t^\alpha x_T(t), \\ \frac{\partial H_T}{\partial u_T} = 0, \end{cases} \quad (13.4)$$

که

$$H_T = H(t, x_T(t), u_T(t), \lambda_T(t)).$$

برای تبدیل (13.4) به یک مسئله بهینه‌سازی نامقید ابتدا سیستم بهیگی (13.4) را روی $N + 1$ نقطه $t_k = t_0 + \frac{t_f - t_0}{N}k$ ، $k = 0, 1, \dots, N$ از بازه $[t_0, t_f]$ گسسته‌سازی می‌کنیم. برای بازه زمانی گسسته‌سازی شده $[t_0, t_f]$ با $N + 1$ گره با فاصله مساوی، رابطه (4.4) را می‌توان به صورت تقریبی زیر نوشت

$${}^G D^\alpha y(t_k) \simeq \frac{1}{h^\alpha} \sum_{j=0}^k w_j^{(\alpha)} y(t_{k-j}), \quad k = 0, \dots, N. \quad (14.4)$$

متناظر با (14.4)، مشتقات چپ و راست گرانوالد-لتنیکف در (13.4) در نقطه k به شرح زیر است

$${}_{t_0} D_{t_k}^\alpha x_T \simeq \frac{1}{h^\alpha} \sum_{j=0}^k w_j^{(\alpha)} x_T(t_{k-j}), \quad k = 1, \dots, N, \quad (15.4)$$

$${}_{t_k}D_{t_f}^{\alpha}\lambda_T \simeq \frac{1}{h^{\alpha}} \sum_{j=\circ}^{N-k} w_j^{(\alpha)} \lambda_T(t_{k+j}), \quad k = N-1, \dots, \circ. \quad (16.4)$$

بنابراین مسئله بهینه‌سازی به صورت زیر تعریف می‌شود

$$\text{minimize}_y \quad E(y) = \frac{1}{\Upsilon} \sum_{k=\circ}^N \{E_{\lambda}(t_k, y) + E_{\Upsilon}(t_k, y) + E_{\Upsilon}(t_k, y)\}, \quad (17.4)$$

که $y = (w_x, w_{\lambda}, w_u, b_x, b_{\lambda}, b_u, v_x, v_{\lambda}, v_u)^T \in \mathbb{R}^{\Upsilon I(\Upsilon p+q)}$ و

$$\begin{cases} E_{\lambda}(t_k, y) = \left[\frac{\partial H_T}{\partial x_T} - \frac{1}{h^{\alpha}} \sum_{j=\circ}^{N-k} w_j^{(\alpha)} \lambda_T(t_{k+j}) \right]^{\Upsilon}, & k = \circ, \dots, N-1, \\ E_{\Upsilon}(t_k, y) = \left[\frac{\partial H_T}{\partial \lambda_T} - \frac{1}{h^{\alpha}} \sum_{j=\circ}^k w_j^{(\alpha)} x_T(t_{k-j}) \right]^{\Upsilon}, & k = 1, \dots, N, \\ E_{\Upsilon}(t_k, y) = \left[\frac{\partial H_T}{\partial u_T} \right]^{\Upsilon}, & k = 1, \dots, N. \end{cases} \quad (18.4)$$

لم ۱.۳.۴. اگر $y^* = (w_x^*, w_{\lambda}^*, w_u^*, b_x^*, b_{\lambda}^*, b_u^*, v_x^*, v_{\lambda}^*, v_u^*)^T$ در دستگاه زیر صدق کند

$$\eta(y) = \begin{bmatrix} E_{\lambda}(t_1, y) \\ E_{\lambda}(t_{\Upsilon}, y) \\ \vdots \\ E_{\lambda}(t_N, y) \\ E_{\Upsilon}(t_1, y) \\ E_{\Upsilon}(t_{\Upsilon}, y) \\ \vdots \\ E_{\Upsilon}(t_N, y) \\ E_{\Upsilon}(t_1, y) \\ E_{\Upsilon}(t_{\Upsilon}, y) \\ \vdots \\ E_{\Upsilon}(t_N, y) \end{bmatrix} = \circ, \quad (19.4)$$

آن گاه y^* جواب بهینه (۱۷.۴) است.

برهان. فرض کنید $\eta(y^*) = \circ$. آن گاه برای $i = 1, 2, 3$ و $k = 1, \dots, N$ داریم $E_i(t_k, y^*) = \circ$ چون در (۱۷.۴) $E(y) \geq \circ$ ، بنابراین y^* یک جواب بهینه (۱۷.۴) است. $\square$

حال با توجه به لم ۱.۳.۶، به راحتی می‌توان بررسی کرد مسئله کمینه‌سازی (۱۷.۴) معادل با مسئله زیر است

$$\text{minimize}_y \quad E(y) = \frac{1}{\Upsilon} \|\eta(y)\|^{\Upsilon}. \quad (20.4)$$

برای حل (۲۰.۴) که یک مسئله بهینه‌سازی نامقید است، می‌توان از الگوریتم‌های بهینه‌سازی مانند تندترین کاهش، نیوتن، شبه‌نیوتن، گرادیان مزدوج و غیره استفاده کرد [۹۸، ۹۹]. همچنین می‌توان از الگوریتم‌های ابتکاری مانند الگوریتم ژنتیک، بهینه‌سازی گروه پرندگان، الگوریتم مورچگان و غیره استفاده کرد [۱۰۰].

۴.۴ الگوریتم آموزش

فرض کنید $v_u(\cdot)$ ، $v_\lambda(\cdot)$ ، $v_x(\cdot)$ ، $b_u(\cdot)$ ، $b_\lambda(\cdot)$ ، $b_x(\cdot)$ ، $w_u(\cdot)$ ، $w_\lambda(\cdot)$ ، $w_x(\cdot)$ متغیرهای وابسته به زمان هستند. هدف طراحی یک مدل دینامیکی است که به نقطه تعادل خود، که یک نقطه تعادل از تابع انرژی $E(y)$ است میل کند. یک مدل دینامیکی برای حل مسئله (۲۰.۴) به صورت زیر پیشنهاد می‌شود

$$\begin{cases} \frac{dy(t)}{dt} = -\kappa \nabla E(y(t)), & \kappa > 0, \\ y(0) = y_0, \end{cases} \quad (21.4)$$

که κ نرخ همگرایی برای (۲۱.۴) است. اکنون برخی از نتایج اولیه در زمینه معادلات دیفرانسیل معمولی که بعدها مورد استفاده قرار می‌گیرد و در تجزیه و تحلیل‌های بعدی نقش مهمی را ایفا می‌کنند یادآوری می‌کنیم. در ادامه مطلب $\|\cdot\|$ ، نرم L^2 از $\mathbb{R}^p$ ، T ، ترانهاده و $x = (x_1, x_2, \dots, x_p)^T$.

اگر $F: \mathbb{R}^p \rightarrow \mathbb{R}$ ، تابع مشتق‌پذیر باشد، آن‌گاه $\nabla F \in \mathbb{R}^p$ ، گرادیان F است. برای هر نگاشت مشتق‌پذیر $F: \mathbb{R}^p \rightarrow \mathbb{R}^q$ ، $F = (F_1, \dots, F_q)^T: \mathbb{R}^p \rightarrow \mathbb{R}^q$ ، $\nabla F = [\nabla F_1(x), \dots, \nabla F_q(x)] \in \mathbb{R}^{p \times q}$ ، ژاکوبین F در x است.

تعریف ۱.۴.۴. [۱۰۱] فرض کنید $\Omega \subseteq \mathbb{R}^p$ یک همسایگی باز از $\bar{x}$ است. به تابع مشتق‌پذیر و پیوسته $\zeta: \mathbb{R}^p \rightarrow \mathbb{R}$ ، یک تابع لیاپانوف در $\bar{x}$ برای سیستم $\dot{x} = F(x)$ گفته می‌شود، اگر

$$\begin{cases} \zeta(\bar{x}) = 0, & \zeta(x) > 0, \forall x \in \Omega \setminus \{\bar{x}\}, \\ \frac{d\zeta(x(t))}{dt} = [\nabla \zeta(x(t))]^T F(x(t)) \leq 0, & \forall x \in \Omega. \end{cases}$$

لم ۱.۴.۴. [۱۰۱]

(a) نقطه تعادل تنها x^* از سیستم $\dot{x} = F(x)$ پایدار لیاپانوف است اگر همسایگی Ω^* از x^* روی تابع لیاپانوف وجود داشته باشد.

²Isolated Equilibrium Point

(b) نقطه تعادل تنها x^* از سیستم $\dot{x} = F(x)$ پایدار مجانبی است هرگاه یک همسایگی Ω^* از x^* روی تابع لیاپانوف وجود داشته باشد به طوری که $\forall x \in \Omega^* \setminus \{x^*\}$ داشته باشیم

$$\frac{d\zeta(x(t))}{dt} < 0.$$

تعریف ۲.۴.۴. [۹۹] فرض کنید $x(t)$ یک مسیر از سیستم $\dot{x} = F(x)$ و X^* مجموعه‌ای از نقاط تعادل این سیستم است. مسیر $x(t)$ را به مجموعه X^* همگرای سراسری گویند اگر $x(t)$ در

$$\lim_{t \rightarrow \infty} \text{dist}(x(t), X^*) = 0,$$

صدق کند که $\text{dist}(x(t), X^*) = \inf_{y \in X^*} \|x - y\|$. به طور خاص، اگر مجموعه X^* فقط شامل نقطه x^* باشد آن گاه $\lim_{t \rightarrow \infty} x(t) = x^*$ و سیستم در نقطه x^* پایدار مجانبی لیاپانوف است.

لم ۲.۴.۴. [۹۹] اگر A یک ماتریس $n \times n$ ناتکین باشد آن گاه دستگاه همگن $AX = 0$ دارای جواب بدیهی $X = 0$ است.

قضیه ۱.۴.۴. [۹۹] مسئله بهینه‌سازی زیر را در نظر بگیرید

$$\begin{aligned} &\text{minimize} && f(x), \\ &\text{subject to} && x \in \mathbb{R}^n, \end{aligned}$$

که $f: \mathbb{R}^n \rightarrow \mathbb{R}$ در نقطه $x^* \in \mathbb{R}^n$ مشتق پذیر است. اگر x^* یک جواب موضعی در مسئله بهینه‌سازی باشد، آن گاه $\nabla f(x^*) = 0$.

۵.۴ پایداری و همگرایی

قضیه ۱.۵.۴. فرض کنید y^* یک نقطه تعادل^۳ از مدل (۲۱.۴) و ماتریس ژاکوبین $\eta(y)$ در (۱۹.۴) معکوس پذیر است. آن گاه y^* یک جواب بهینه برای (۲۰.۴) است. برعکس، اگر y^* یک جواب بهینه برای (۲۰.۴) باشد، آن گاه y^* یک نقطه تعادل برای (۲۱.۴) است.

برهان. فرض کنید y^* یک نقطه تعادل برای مدل دینامیکی (۲۱.۴) است. واضح است که $\nabla E(y^*) = 0$. با یک محاسبه ساده می‌توان نشان داد

$$\nabla E(y) = (\nabla \eta(y))^T \eta(y),$$

که $\nabla \eta(y)$ ماتریس ژاکوبین $\eta(y)$ است. طبق لم ۲.۴.۴، $\eta(y^*) = 0$ ، یعنی y^* یک جواب بهینه برای (۲۰.۴) است. با استفاده از قضیه ۱.۴.۴ عکس قضیه نیز اثبات می‌شود. $\square$

³Equilibrium Point

قضیه ۲.۵.۴. فرض کنید y^* یک نقطه تعادل تنها برای (۲۱.۴) است و ماتریس ژاکوبین $\eta(y)$ در (۱۹.۴) معکوس پذیر است. آن گاه سیستم (۲۱.۴) پایدار مجانبی در y^* است.

برهان. ابتدا توجه کنید که $E(y) \geq 0$ و $E(y^*) = 0$. علاوه بر این، چون y^* یک نقطه تعادل تنها از (۲۱.۴) است، پس یک همسایگی $\Omega^* \subseteq \mathbb{R}^{I(2p+q)}$ از y^* وجود دارد به طوری که $E(y^*) = 0$ و

$$\nabla E(y) \neq 0, \quad \forall y \in \Omega^* \setminus \{y^*\}.$$

ادعا می کنیم که برای هر $y \in \Omega^* \setminus \{y^*\}$ ، $E(y) > 0$. در غیر این صورت اگر $y \in \Omega^* \setminus \{y^*\}$ وجود داشته باشد که در $E(y) = 0$ صدق کند آن گاه $\nabla E(y) = 0$. یعنی y نیز یک نقطه تعادل برای (۲۱.۴) است که به وضوح با فرض اینکه y^* یک نقطه تعادل تنها برای Ω^* باشد در تناقض است. علاوه بر این

$$\frac{dE(y(t))}{dt} = \nabla[E(y(t))]^T \frac{dy(t)}{dt} = -\kappa \|\nabla E(y(t))\|^2 \leq 0, \quad \forall y \in \Omega^*.$$

در نتیجه

$$\frac{dE(y(t))}{dt} < 0, \quad \forall y(t) \in \Omega^*, \quad y(t) \neq y^*.$$

با استفاده از لم ۱.۴.۴ قسمت (b)، سیستم (۲۱.۴) پایدار مجانبی در y^* است. $\square$

گزاره ۱.۵.۴. (a) برای هر حالت اولیه $y_0 = y(t_0)$ ، حداکثر یک جواب ماکسیمال $y(t, y_0)$ با $t \in [t_0, \beta(y_0))$ برای مدل (۲۱.۴) وجود دارد.

(b) اگر مجموعه سطح $L(y_0) = \{y \in \mathbb{R}^{I(2p+q)} : E(y) \leq E(y_0)\}$ کراندار باشد آن گاه

$$\beta(y_0) \rightarrow +\infty.$$

برهان. (a) چون η دیفرانسیل پذیر پیوسته است، در نتیجه $\nabla \eta$ نیز پیوسته است. بنابراین در یک همسایگی فشرده محلی از y کراندار است و این نتیجه می دهد که $\nabla E(y)$ به طور محلی پیوسته لیپشیتز است. بنابر لم ۱.۲ در [۱۰۲] نتیجه مطلوب حاصل می شود. (b) مجموعه سطح $L(y_0)$ کراندار است. نتیجه را با استفاده از برهان خلف اثبات می کنیم. فرض می کنیم $\beta(y_0) < \infty$ ، آن گاه

$$\lim_{t \rightarrow \beta(y_0)} \|y(t)\| = \infty.$$

متمم $L(y_0)$ و β_0 را به صورت زیر تعریف می کنیم

$$L^C(y_0) := \mathbb{R}^{I(2p+q)} \setminus L(y_0), \quad (22.4)$$

$$\beta_0 := \inf\{s \geq 0 \mid s < \beta(y_0), y(s) \in L^C(y_0)\} < \infty.$$

می‌دانیم که $y(\beta_0)$ روی مرز $L(y_0)$ و $L^C(y_0)$ قرار دارد. از طرفی طبق فرض، $L(y_0)$ کران‌دار است و با توجه به پیوستگی $E(y)$ بسته نیز است. در نتیجه $L(y_0)$ فشرده است. بنابراین $y(\beta_0) \in L(y_0)$ و $\beta_0 < \beta(y_0)$. پس برای برخی از $s \in (\beta_0, \beta(y_0))$ داریم

$$E(y(s)) > E(y_0) > E(y(\beta_0)). \quad (23.4)$$

از طرفی از قضیه ۲.۵.۴ نتیجه می‌شود که E روی بازه $[t_0, \beta(y_0))$ ، غیرافزایشی است، که با (۲۳.۴) در تناقض است. در نتیجه اثبات (b) تمام است. $\square$

قضیه ۳.۵.۴. فرض کنید $y = y(t, y_0)$ یک مسیر از (۲۱.۴) در نقطه ابتدایی $y_0 = y(0, y_0)$ است و مجموعه سطح $L(y_0) = \{y \in \mathbb{R}^{3I(\gamma_{p+q})} : E(y) \leq E(y_0)\}$ کراندار باشد. آن‌گاه

$$(a) \quad \gamma^+(y_0) = \{y(t, y_0) | t \geq 0\} \text{ کراندار است.}$$

$$(b) \quad \bar{y} \text{ وجود دارد به طوری که } \lim_{t \rightarrow \infty} y(t, y_0) = \bar{y}$$

(c) فرض کنید ماتریس ژاکوبین از $\eta(y)$ در (۱۹.۴) معکوس‌پذیر است. آن‌گاه $\forall y_0 \in \mathbb{R}^{3I(\gamma_{p+q})}$ مسیر مربوطه از (۲۱.۴) به یک جواب بهینه از (۲۰.۴) همگرا می‌شود.

برهان. (a) با محاسبه مشتق $E(y)$ در امتداد مسیر برای $y(t, y_0)$ ($t \geq 0$) داریم

$$\frac{dE(y(t))}{dt} = \nabla[E(y(t))]^T \frac{dy(t)}{dt} = -\kappa \|\nabla E(y(t))\|^2 \leq 0. \quad (24.4)$$

بدین ترتیب در امتداد مسیر $y = y(t, y_0)$ ، $(t \geq 0)$ ، E به طور یکنواخت کاهش می‌یابد. از این رو $\gamma^+(y_0) \subseteq L(y_0)$ ، یعنی $\gamma^+(y_0) = \{y(t, y_0) | t \geq 0\}$ کراندار است.

(b) با استفاده از قسمت (a)، $\gamma^+(y_0) = \{y(t, y_0) | t \geq 0\}$ یک مجموعه کراندار از نقاط است، پس دنباله $\{\bar{t}_p\}$ ، $0 < \bar{t}_1 < \dots < \bar{t}_p \rightarrow \infty$ را یک دنباله اکیداً یکنواخت افزایشی در نظر می‌گیریم. در نتیجه $\{y(\bar{t}_p, y_0)\}$ دنباله‌ای کراندار است که از بی‌نهایت نقطه تشکیل شده است. بنابراین نقطه حدی $\bar{y}$ وجود دارد. به این معنا که یک زیردنباله $\{t_p\} \subseteq \{\bar{t}_p\}$ ، $t_p \rightarrow \infty$ وجود دارد به طوری که $\lim_{p \rightarrow \infty} (t_p, y_0) = \bar{y}$ ، که نشان می‌دهد $\bar{y}$ نقطه w -حدی از $\gamma^+(y_0)$ است. با استفاده از قضیه لسل ۸.۲.۲، زمانی که $t \rightarrow \infty$ داریم $y(t, y_0) \rightarrow \bar{y} \in M$ که M بزرگترین مجموعه در $K = \{y(t, y_0) | \frac{dE(y(t, y_0))}{dt} = 0\}$ است.

(c) با استفاده از (a) و (b)، $\bar{y}$ یک نقطه تعادل از (۲۱.۴) است. چون $\nabla \eta$ معکوس‌پذیر است از (۲۴.۴) نتیجه می‌شود

$$\frac{dw_x}{dt} = \frac{dw_p}{dt} = \dots = \frac{d\nu_u}{dt} = 0 \iff \frac{dE(y(t))}{dt} = 0.$$

بنابراین با استفاده از $D^* \subseteq K \subseteq M$ که D^* مجموعه جواب‌های بهینه (۲۰.۴) معرفی شده است داریم $\bar{y} \in D^*$. در نتیجه از هر نقطه اولیه y_0 ، مسیر $y(t, y_0)$ به یک جواب بهینه از (۲۰.۴) همگرا می‌شود. قضیه‌های ۲.۵.۴ و ۳.۵.۴ نشان می‌دهند که مدل ارائه شده در (۲۱.۴) به طور مجانبی پایدار است و به یک جواب دقیق از (۲۰.۴) همگرا است. به ویژه اگر (۲۰.۴) یک نقطه کمینه یکتا داشته باشد (یعنی یک نقطه تعادل منحصربه‌فرد در (۲۱.۴) وجود داشته باشد) آن‌گاه مسیر $y(t, y_0)$ همگرای سراسری است. $\square$

۶.۴ مثال‌های عددی

در این بخش، سعی می‌شود که الگوریتم پیشنهادی برای چند مثال بررسی شود. در دو مثال اول علاوه بر بررسی مسائل کنترل بهینه کسری با مرتبه ثابت، همان مسائل با مرتبه متغیر هم به‌طور جداگانه بررسی خواهد شد و روش پیشنهاد شده روی آن‌ها اعمال می‌شود. برای هر مسئله ۵ پارامتر برای وزن‌های ورودی، خروجی و بایاس استفاده می‌کنیم و مشتقات کسری هم با کمک تعریف مشتق گرانوالد-لتنیکف تقریب زده می‌شود. در همه مثال‌ها از نرم‌افزار MATLAB برای شبیه‌سازی استفاده شده است.

مثال ۱.۶.۴. مسئله

$$\begin{aligned} \text{minimize} \quad & J = \frac{1}{T} \int_0^1 (x^2(t) + u^2(t)) dt, \\ \text{subject to} \quad & {}^\circ D_t^\alpha x(t) = -x(t) + u(t), \end{aligned} \quad (25.4)$$

با شرط اولیه

$$x(0) = 1,$$

را در نظر بگیرید. در این سیستم مشتق کسری از مرتبه ثابت α است. از [۱۰۳] می‌دانیم که جواب تحلیلی برای مسئله ارائه شده برای $\alpha = 1$ به‌صورت زیر است

$$\begin{aligned} x(t) &= \cosh(\sqrt{2}t) + \beta \sinh(\sqrt{2}t), \\ u(t) &= (1 + \sqrt{2}\beta) \cosh(\sqrt{2}t) + (\sqrt{2} + \beta) \sinh(\sqrt{2}t), \end{aligned}$$

که

$$\beta = -\frac{\cosh(\sqrt{2}t) + \sqrt{2} \sinh(\sqrt{2}t)}{\sqrt{2} \cosh(\sqrt{2}t) + \sinh(\sqrt{2}t)} \approx -0.98.$$

متناظر با (۹.۴) داریم

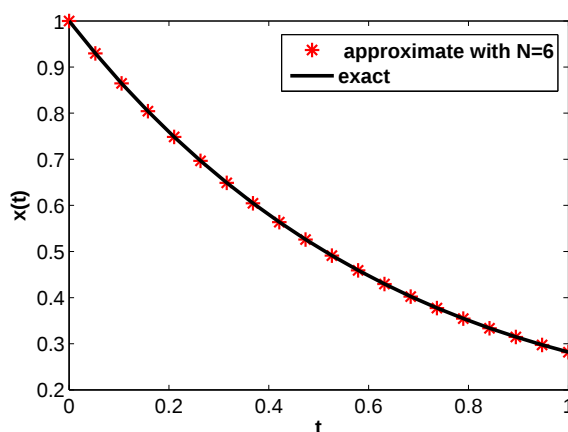
$$\begin{cases} {}_t D_T^\alpha \lambda(t) = x(t) - \lambda(t), \\ {}^\circ D_t^\alpha x(t) = -x(t) + u(t), \\ u(t) + \lambda(t) = 0, \\ x(0) = 1, \quad \lambda(1) = 0. \end{cases} \quad (26.4)$$

چون $x(1)$ آزاد است، داریم $\lambda(1) = 0$. با توجه به این شرط و شرط اولیه $x(0) = 1$ ، جواب‌های آزمایشی را می‌توان به‌صورت زیر انتخاب کرد

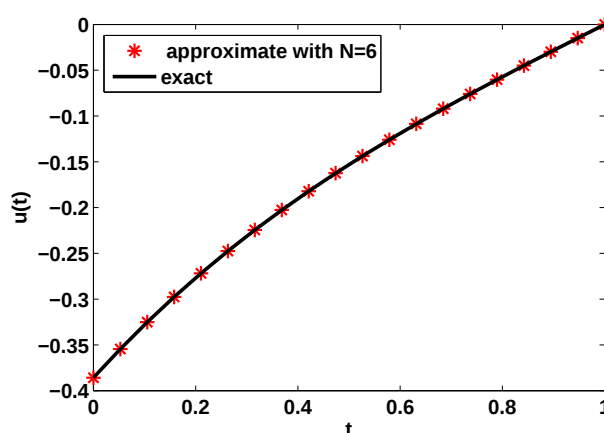
$$\begin{cases} x_T = 1 + tN_x, \\ \lambda_T = (t-1)N_\lambda, \\ u_T = N_u. \end{cases} \quad (27.4)$$

۶۰ مسائل کنترل بهینه کسری با شبکه عصبی و طرح‌های بهینه‌سازی دینامیکی

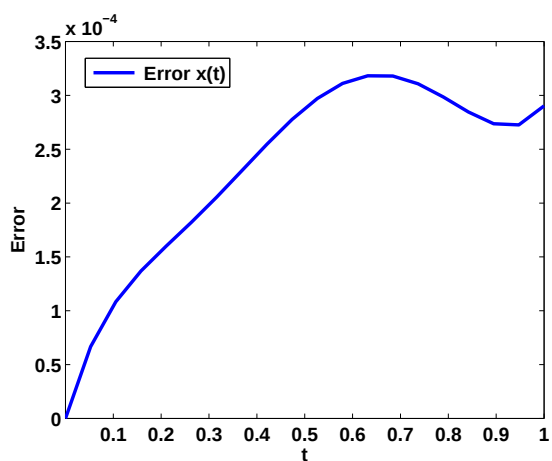
جواب‌های آزمایشی به‌دست آمده را در شرایط لازم بهینگی (۲۶.۴) جایگذاری می‌کنیم. با روند ارائه شده در این فصل برای تبدیل شرایط بهینگی به یک مسئله بهینه‌سازی نامقید، سیستم بهینگی (۲۶.۴) را روی $N = 6$ نقطه از بازه $[0, 1]$ گسسته‌سازی می‌کنیم. با کمک نرم‌افزار مطلب، با یک روش بهینه‌سازی نامقید مسئله را حل می‌کنیم. وزن‌های به‌دست آمده را در جواب‌های آزمایشی (۲۷.۴) جایگذاری کرده و سپس مسیر بهینه را رسم می‌کنیم. در شکل‌های ۲.۴ و ۳.۴ جواب‌های تقریبی توابع وضعیت و کنترل برای روش پیشنهاد شده با جواب‌های دقیق مقایسه شده‌اند. خطای مطلق توابع در شکل‌های ۴.۴ و ۵.۴ نشان داده شده است. همچنین جواب‌های تقریبی توابع وضعیت و کنترل با $N = 6$ برای چند مقدار α به ترتیب در شکل‌های ۶.۴ و ۷.۴ رسم شده است.



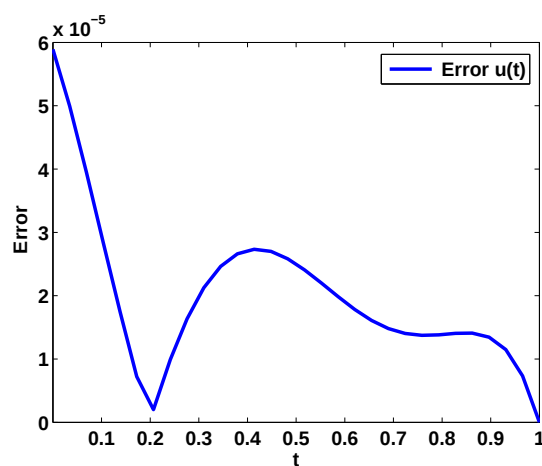
شکل ۲.۴: مسیر بهینه $x(\cdot)$ برای $\alpha = 1$ در مثال ۱.۶.۴.



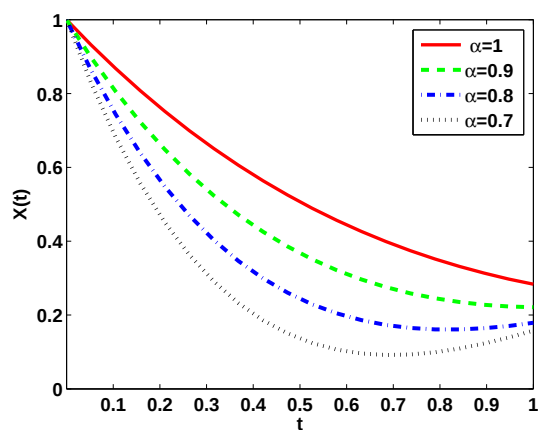
شکل ۳.۴: کنترل بهینه $x(\cdot)$ برای $\alpha = 1$ در مثال ۱.۶.۴.



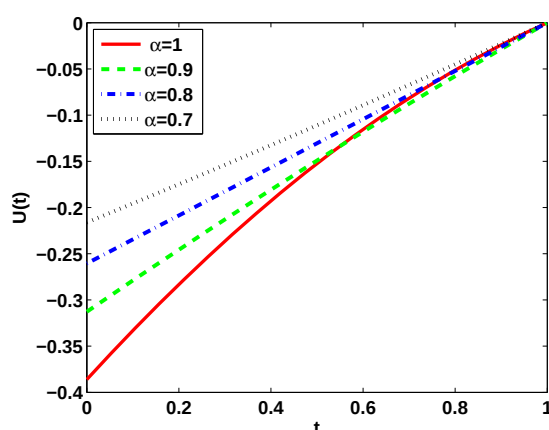
شکل ۴.۴: خطای تقریب تابع وضعیت برای $\alpha = 1$ در مثال ۱.۶.۴.



شکل ۵.۴: خطای تقریب تابع کنترل برای $\alpha = 1$ در مثال ۱.۶.۴.



شکل ۶.۴: جواب تقریبی مسیر بهینه $x(\cdot)$ برای مقادیر مختلف α در مثال ۱.۶.۴.



شکل ۷.۴: جواب تقریبی کنترل بهینه $u(\cdot)$ برای مقادیر مختلف α در مثال ۱.۶.۴.

حال سیستم (۲۵.۴) را با مشتق کسری از مرتبه متغیر $\alpha(t)$ به صورت زیر در نظر بگیرید

$$\begin{aligned} \text{minimize} \quad & J = \frac{1}{\tau} \int_0^1 (x^\tau(t) + u^\tau(t)) dt, \\ \text{subject to} \quad & {}^\circ D_t^{\alpha(t)} x(t) = -x(t) + u(t), \end{aligned} \quad (28.4)$$

با شرط اولیه

$$x(\circ) = 1.$$

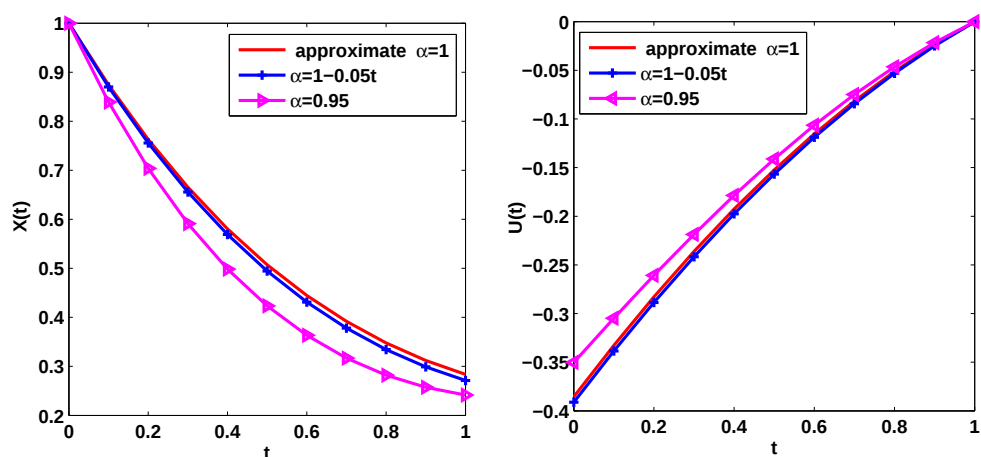
شرایط لازم مسئله عبارت است از

$$\begin{cases} {}_t D_1^{\alpha(t)} \lambda(t) = x(t) - \lambda(t), \\ {}^\circ D_t^{\alpha(t)} x(t) = -x(t) + u(t), \\ u(t) + \lambda(t) = \circ, \\ x(\circ) = 1, \quad \lambda(1) = \circ. \end{cases} \quad (29.4)$$

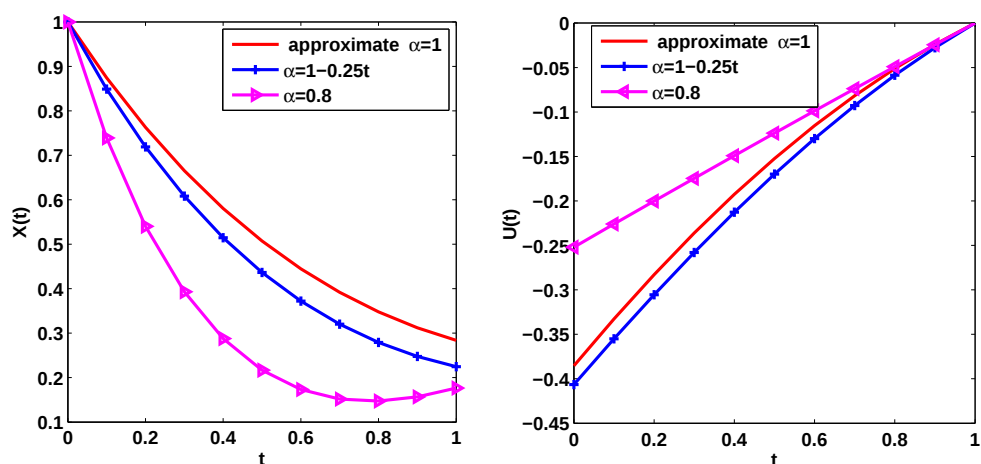
همان ساختار شبکه و جواب‌های آزمایشی (۲۷.۴) که برای مشتق مرتبه ثابت ارائه شد را در شرایط بهینگی (۲۹.۴) جایگذاری می‌کنیم. بعد از حل با نرم افزار، شکل ۸.۴ جواب مسئله برای مرتبه متغیر $\alpha(t) = 1 - \circ/5t$ ، مرتبه ثابت $\alpha = 1$ و $\alpha = \circ/95$ را نشان می‌دهد. شکل ۹.۴ جواب مسئله برای مرتبه متغیر $\alpha(t) = 1 - \circ/25t$ و مرتبه ثابت $\alpha = \circ/8$ را نشان می‌دهد.

مثال ۲.۶.۴. [۵۶] مسئله کنترل بهینه کسری وابسته به زمان را به صورت زیر در نظر بگیرید

$$\begin{aligned} \text{minimize} \quad & J = \frac{1}{\tau} \int_0^1 (x^\tau(t) + u^\tau(t)) dt, \\ \text{subject to} \quad & {}^\circ D_t^\alpha x(t) = tx(t) + u(t), \\ & x(\circ) = 1. \end{aligned} \quad (30.4)$$



شکل ۸.۴: جواب تقریبی مسیر بهینه $x(\cdot)$ و کنترل بهینه $u(\cdot)$ برای مقادیر α و $\alpha(t)$ در مثال ۱.۶.۴.



شکل ۹.۴: جواب تقریبی مسیر بهینه $x(\cdot)$ و کنترل بهینه $u(\cdot)$ برای مقادیر α و $\alpha(t)$ در مثال ۱.۶.۴.

متناظر با شرایط لازم بهینگی (۹.۴) داریم

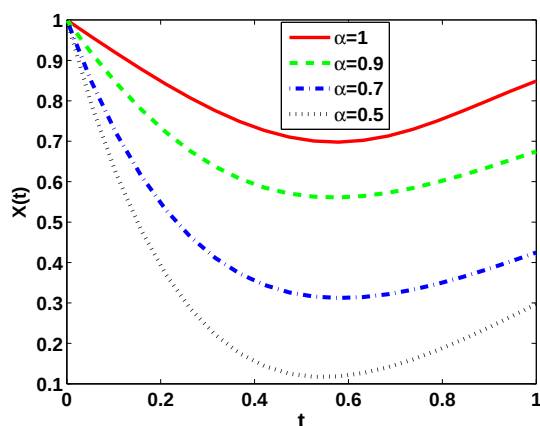
$$\begin{cases} {}_t D_t^\alpha \lambda(t) = x(t) + t\lambda(t), \\ {}^\circ D_t^\alpha x(t) = tx(t) + u(t), \\ u(t) + \lambda(t) = {}^\circ, \\ x({}^\circ) = 1, \quad \lambda(1) = {}^\circ. \end{cases}$$

چون $x(1)$ آزاد است، داریم $\lambda(1) = {}^\circ$. با در نظر گرفتن این شرط و شرط اولیه $x({}^\circ) = 1$

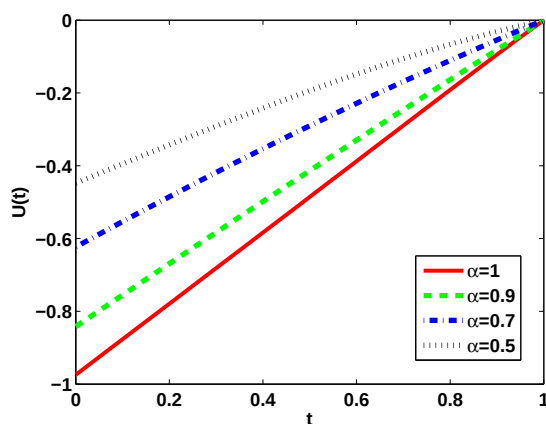
می‌توان جواب‌های آزمایشی را به صورت زیر انتخاب کرد

$$\begin{cases} x_T = 1 + tN_x, \\ \lambda_T = (t - 1)N_\lambda, \\ u_T = N_u. \end{cases}$$

رفتار جواب‌های عددی این مسئله برای مقادیر مختلف α با $N = 6$ برای توابع وضعیت و کنترل به ترتیب در شکل‌های ۱۰.۴ و ۱۱.۴ نشان داده می‌شود.



شکل ۱۰.۴: جواب تقریبی مسیر بهینه $x(\cdot)$ برای مقادیر مختلف α در مثال ۲.۶.۴.



شکل ۱۱.۴: جواب تقریبی کنترل بهینه $u(\cdot)$ برای مقادیر مختلف α در مثال ۲.۶.۴.

مانند مثال قبل مسئله (۳۰.۴) را با مشتق کسری از مرتبه متغیر $\alpha(t)$ به صورت زیر در نظر

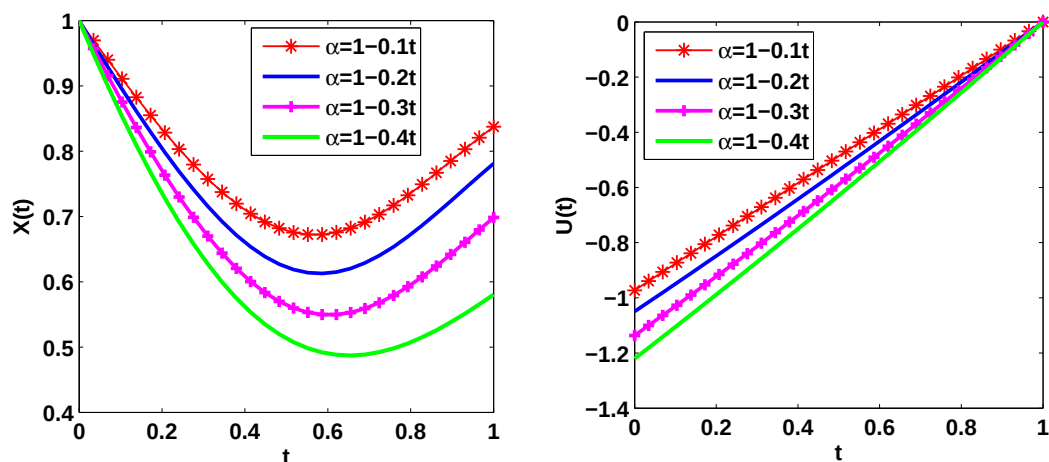
بگیرید

$$\begin{aligned} \text{minimize} \quad & J = \frac{1}{\tau} \int_0^1 (x^\tau(t) + u^\tau(t)) dt, \\ \text{subject to} \quad & {}_0 D_t^{\alpha(t)} x(t) = tx(t) + u(t), \\ & x(0) = 1. \end{aligned} \quad (31.4)$$

شرایط لازم عبارت است از

$$\begin{cases} {}_t D_1^{\alpha(t)} \lambda(t) = x(t) + t\lambda(t), \\ {}_0 D_t^{\alpha(t)} x(t) = tx(t) + u(t), \\ u(t) + \lambda(t) = 0, \\ x(0) = 1, \quad \lambda(1) = 0. \end{cases} \quad (32.4)$$

جواب‌های آزمایشی انتخاب شده در حالت مرتبه ثابت را در شرایط (۳۲.۴) جایگذاری می‌کنیم. بعد از حل با نرم افزار مسیر بهینه و کنترل بهینه برای چند مقدار با α های متغیر را می‌توان در شکل ۱۲.۴ مشاهده کرد.



شکل ۱۲.۴: جواب تقریبی مسیر بهینه $x(\cdot)$ و کنترل بهینه $u(\cdot)$ برای مقادیر مختلف $\alpha(t)$ در مثال ۲.۶.۴.

مثال ۳.۶.۴. مسئله کنترل بهینه زیر را در نظر بگیرید

$$\begin{aligned} \text{minimize} \quad & J = \int_0^1 [(x(t) - t^\tau)^2 + (u(t) + t^\tau - \frac{\tau \circ t^{\frac{\tau}{\tau_0}}}{\Gamma(\frac{\tau}{\tau_0})})^2] dt, \\ \text{subject to} \quad & {}_0 D_t^\lambda x(t) = y(t), \\ & {}_0 D_t^{\alpha/\lambda} y(t) = t^\tau x(t) + u(t), \\ & x(0) = 0, \quad y(0) = 0. \end{aligned} \quad (33.4)$$

جواب‌های تحلیلی مسئله به صورت زیر است

$$x(t) = t^2, \quad y(t) = 2t, \quad u(t) = \frac{2 \circ t^{\frac{9}{10}}}{9\Gamma(\frac{9}{10})} - t^4.$$

با تعریف تابع همیلتونی به صورت

$$H = (x(t) - t^2)^2 + (u(t) + t^4 - \frac{2 \circ t^{\frac{9}{10}}}{9\Gamma(\frac{9}{10})})^2 + \lambda_1(t)(y(t)) + \lambda_2(t)(t^2 x(t) + u(t)),$$

و متناظر با (۹.۴) شرایط لازم را می‌توان به صورت زیر نوشت

$$\begin{cases} \circ D_t^1 x(t) = y(t) \\ {}_t D_1^1 \lambda_1(t) = -2(x(t) - t^2) - \lambda_2(t)t^2 \\ \circ D_t^{\circ/1} y(t) = t^2 x(t) + u(t), \\ {}_t D_1^{\circ/1} \lambda_2(t) = \lambda_1(t), \\ 2(u(t) + t^4 - \frac{2 \circ t^{\frac{9}{10}}}{9\Gamma(\frac{9}{10})}) + \lambda_2(t) = 0, \\ x(\circ) = 0, \quad y(\circ) = 0, \quad \lambda_1(1) = 0, \quad \lambda_2(1) = 0. \end{cases}$$

چون $x(1)$ و $y(1)$ آزاد هستند داریم $\lambda_1(1) = 0$ و $\lambda_2(1) = 0$. با در نظر گرفتن این شرایط و شرایط اولیه $x(\circ) = 0$ و $y(\circ) = 0$ می‌توان جواب‌های آزمایشی را به صورت زیر تعریف کرد

$$\begin{cases} x_T = tN_x, \\ y_T = tN_y, \\ \lambda_{1T} = (t-1)N_{\lambda_1}, \\ \lambda_{2T} = (t-1)N_{\lambda_2}, \\ u_T = N_u. \end{cases}$$

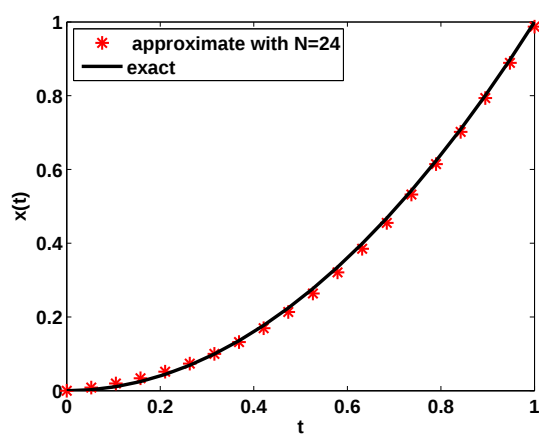
در شکل‌های ۱۳.۴ و ۱۴.۴ جواب‌های تقریبی $x(t)$ و $u(t)$ برای $N = 24$ با جواب‌های دقیق مقایسه شده‌اند. توابع خطای متناظر با توابع مسیر و کنترل در شکل‌های ۱۵.۴ و ۱۶.۴ رسم شده‌اند.

مثال ۴.۶.۴. مسئله کنترل بهینه زیر را در نظر بگیرید

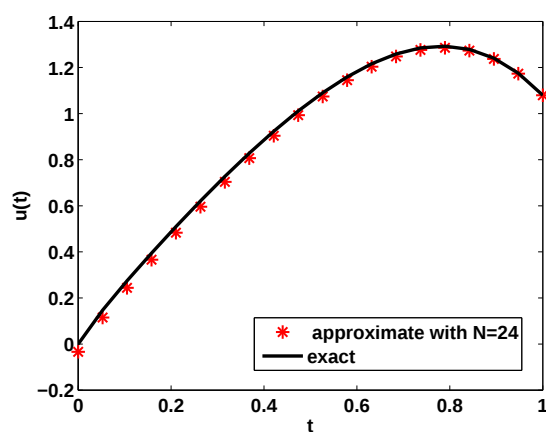
$$\begin{aligned} \text{minimize} \quad & J = \frac{1}{T} \int_0^1 (3x^2(t) + u^2(t))dt \\ \text{subject to} \quad & \circ D_t^\alpha x(t) = -x(t) + u(t), \\ & x(\circ) = 0, \quad x(1) = 2. \end{aligned} \quad (34.4)$$

جواب‌های دقیق در حالت $\alpha = 1$ به صورت

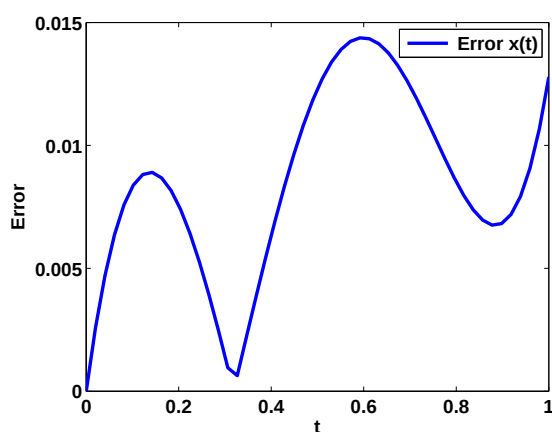
$$\begin{cases} x(t) = \frac{2}{\sinh 2} \sinh(2t), \\ u(t) = \frac{2}{\sinh 2} (2 \cosh(2t) + \sinh(2t)), \end{cases}$$



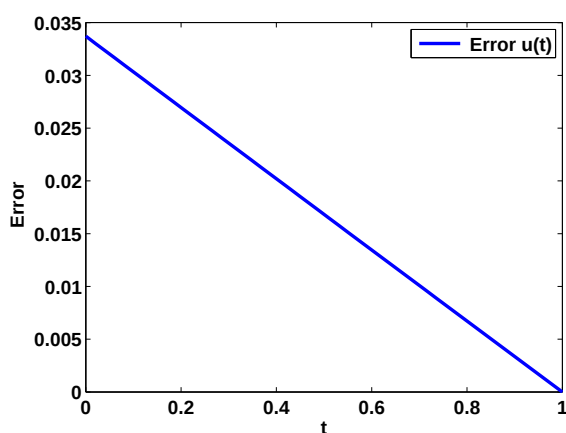
شکل ۱۳.۴: مسیر بهینه $x(\cdot)$ برای $\alpha = 1$ در مثال ۳.۶.۴.



شکل ۱۴.۴: کنترل بهینه $u(\cdot)$ برای $\alpha = 1$ در مثال ۳.۶.۴.



شکل ۱۵.۴: خطای تقریب تابع وضعیت برای $\alpha = 1$ در مثال ۳.۶.۴.



شکل ۱۶.۴: خطای تقریب تابع کنترل برای $\alpha = 1$ در مثال ۳.۶.۴.

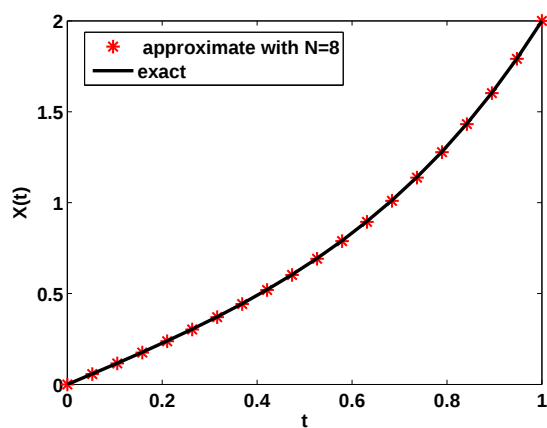
است. مشابه مثال‌های قبلی و رابطه (۹.۴) داریم

$$\begin{cases} {}_t D_1^\alpha \lambda(t) = \mathfrak{V}x(t) - \lambda(t), \\ {}^\circ D_t^\alpha x(t) = -x(t) + u(t), \\ u(t) + \lambda(t) = 0, \\ x(0) = 0, \quad x(1) = 2. \end{cases}$$

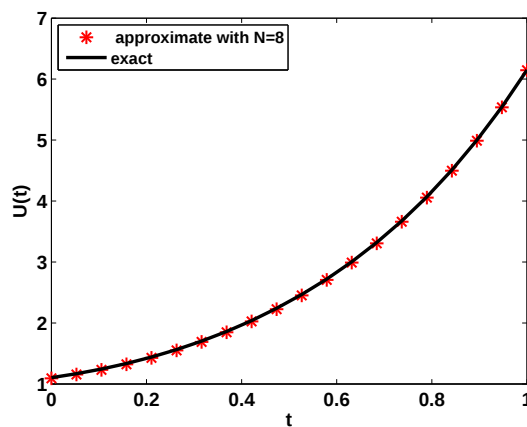
جواب‌های آزمایشی را می‌توان به صورت زیر انتخاب کرد

$$\begin{cases} x_T = 2t + t(t-1)N_x, \\ \lambda_T = N_\lambda, \\ u_T = N_u. \end{cases}$$

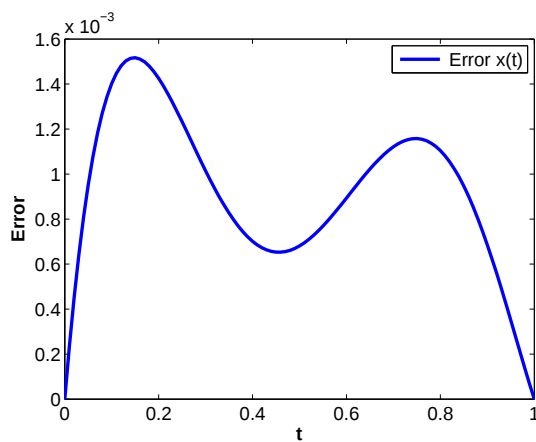
کنترل بهینه $u(\cdot)$ ، مسیر متناظر $x(\cdot)$ و خطای مطلق توابع وضعیت و کنترل در شکل‌های ۱۷.۴ و ۲۰.۴ رسم شده‌اند. همچنین جواب‌های عددی توابع وضعیت و کنترل با $N = 8$ برای چند مقدار α در شکل‌های ۲۱.۴ و ۲۲.۴ به ترتیب رسم شده‌اند.



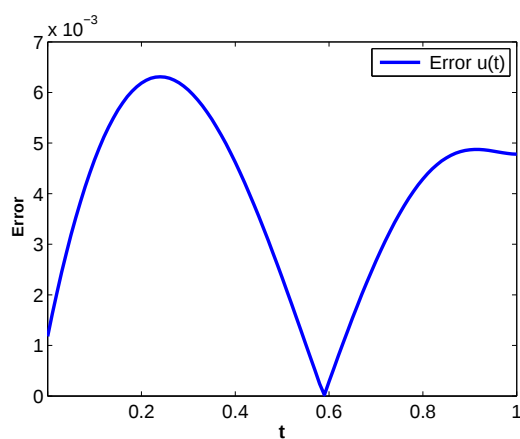
شکل ۱۷.۴: مسیر بهینه $x(\cdot)$ برای $\alpha = 1$ در مثال ۴.۶.۴.



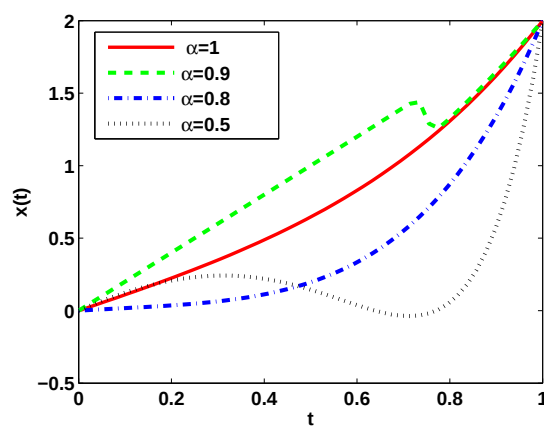
شکل ۱۸.۴: کنترل بهینه $u(\cdot)$ برای $\alpha = 1$ در مثال ۴.۶.۴.



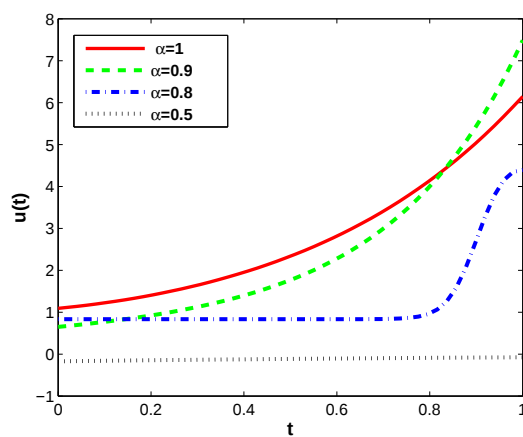
شکل ۱۹.۴: خطای تقریب تابع وضعیت برای $\alpha = 1$ در مثال ۴.۶.۴.



شکل ۲۰.۴: خطای تقریب تابع کنترل برای $\alpha = 1$ در مثال ۴.۶.۴.



شکل ۲۱.۴: مسیر بهینه جواب تقریبی $x(\cdot)$ با مقادیر مختلف α در مثال ۴.۶.۴.



شکل ۲۲.۴: مسیر کنترل جواب تقریبی $u(\cdot)$ با مقادیر مختلف α در مثال ۴.۶.۴.

فصل ۵

حل رده‌ای از مسائل کنترل بهینه کسری با محدودیت‌های مساوی و نامساوی با دو رویکرد متفاوت از شبکه‌های عصبی

در این فصل یک فرم کلی از مسائل کنترل بهینه کسری با محدودیت‌های مساوی و نامساوی با دو روش متفاوت بررسی می‌شود. روش اول ترکیبی از روش‌های سری توانی و شبکه‌های عصبی و روش دوم یک ترکیب از تابع میتاگ-لفلر و شبکه‌های عصبی است. همچنین در روش‌هایی که از شرایط لازم بهینگی استفاده می‌کنند مانند روش‌هایی که در این فصل ارائه خواهد شد علاوه بر مشتق چپ، مشتق راست هم ظاهر می‌شود که در اکثر روش‌های ارائه شده برای محاسبه مشتق راست از انتگرال‌گیری عددی استفاده شده است. در روش اول با انتخاب یک ساختار مناسب از مدل شبکه عصبی می‌توان علاوه بر مشتق چپ، مشتق راست را بدون استفاده از انتگرال‌گیری عددی محاسبه کرد که این خود باعث می‌شود پیچیدگی محاسباتی به‌طور قابل توجهی کاهش یابد. همچنین در انتهای فصل یک مسئله کسری با مشتق مرتبه متغیر با روشی که ارائه می‌شود حل و مورد بررسی قرار می‌گیرد.

$$\text{minimize} \quad J(x, u) = \Phi(t, x)|_b + \int_a^b L(t, x, u) dt, \quad (1.5)$$

$$\text{subject to} \quad A \dot{x}(t) + B {}^c D_t^\alpha x(t) = f(t, x, u), \quad (2.5)$$

$$G(t, x, u) \leq 0, \quad (3.5)$$

$$x(a) = x_a, \quad (4.5)$$

که $x(t) \in \mathbb{R}^p$ ، $(A, B) \neq (0, 0)$ ، $0 < \alpha < 1$ و $x(a)$ ثابت است. $x(b)$ می‌تواند آزاد یا ثابت باشد.

۱.۵ بررسی شرایط لازم بهینگی

از [۵۵] و [۸۵]، اگر (x, u) یک جواب کمینه از مسئله (۱.۵) تا (۴.۵) باشد آن‌گاه $\lambda(t)$ و $\mu(t)$ ای وجود دارند که (x, u, λ, μ) در شرایط زیر صدق می‌کنند

$$\frac{\partial H}{\partial x}(t, x, u, \lambda, \mu) = -A \dot{\lambda}(t) + B {}^c D_b^\alpha \lambda(t), \quad (5.5)$$

$$\frac{\partial H}{\partial \lambda}(t, x, u, \lambda, \mu) = A \dot{x}(t) + B {}^c D_t^\alpha x(t), \quad (6.5)$$

$$\frac{\partial H}{\partial u}(t, x, u, \lambda, \mu) = 0, \quad (7.5)$$

$$\mu G(t, x, u) = 0, \quad (8.5)$$

$$\mu \geq 0, \quad (9.5)$$

$$G(t, x, u) \leq 0, \quad (10.5)$$

$$x(a) = x_a, \quad (11.5)$$

$$\lambda(b) = \frac{\partial \phi}{\partial x}|_b, \quad (12.5)$$

که H تابع همیلتونی است و به صورت زیر تعریف می‌شود

$$H(t, x, u, \lambda, \mu) = L(t, x, u) + \lambda F(t, x, u) + \mu G(t, x, u).$$

می‌توان رابطه‌ای بین جواب‌های (۸.۵)–(۱۰.۵) و جواب یک تابع مسئله مکمل غیرخطی^۱ (NCP) پیدا کرد [۸۵]. رده‌ای از توابع NCP در پایین تعریف شده است

تعریف ۱.۱.۵. • یک تابع $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ یک تابع NCP نامیده می‌شود اگر در شرط زیر صدق کند

$$\phi(a, b) = 0 \iff a \geq 0, b \geq 0, ab = 0.$$

• یک تابع NCP که عمومیت بیشتری دارد تابع فیشر–برمیستر^۲ (FB) است که قویاً نیمه

¹Nonlinear Complementarity Problem

²Fischer-Burmeister

هموار است [۱۰۴] و به‌صورت زیر تعریف می‌شود

$$\phi_{\text{FB}}(a, b) = (a + b) - \sqrt{a^2 + b^2}.$$

تابع FB اغتشاش یافته^۳ نیز به‌صورت زیر تعریف می‌شود

$$\phi_{\text{FB}}^{\varepsilon}(a, b) = (a + b) - \sqrt{a^2 + b^2 + \varepsilon}, \quad \varepsilon \rightarrow 0_+,$$

که ε عددی کوچک و مثبت است. در نتیجه زیر یک خاصیت مهم $\phi_{\text{FB}}^{\varepsilon}$ را می‌توان مشاهده کرد.

گزاره ۱.۱.۵. [۱۰۵] برای هر $\varepsilon \in \mathbb{R}$ داریم

$$\phi_{\text{FB}}^{\varepsilon}(a, b) = 0 \Leftrightarrow a > 0, b > 0, ab = \frac{\varepsilon}{4}.$$

گزاره ۱.۱.۵ واضح است و اثبات آن حذف شده است. باید توجه شود که $\phi_{\text{FB}}^{\varepsilon}(a, b)$ با توجه به a, b برای $\varepsilon > 0$ هموار است.

با استفاده از تابع FB اغتشاش یافته، می‌توان مسئله NCP (۸.۵)–(۱۰.۵) را به محدودیت‌های معادل زیر تبدیل کرد

$$\phi_{\text{FB}}^{\varepsilon}(\mu, -G(t, x, u)) = 0, \quad \varepsilon \rightarrow 0_+.$$

بنابراین شرایط بهینگی (۶.۵)–(۱۲.۵) به فرم زیر بازنویسی می‌شود

$$\begin{cases} \frac{\partial H}{\partial x}(t, x, u, \lambda, \mu) = -A \dot{\lambda}(t) + B {}_t D_b^{\alpha} \lambda(t), \\ \frac{\partial H}{\partial \lambda}(t, x, u, \lambda, \mu) = A \dot{x}(t) + B {}_a^c D_t^{\alpha} x(t), \\ \frac{\partial H}{\partial u}(t, x, u, \lambda, \mu) = 0, \\ \phi_{\text{FB}}^{\varepsilon}(\mu, -G(t, x, u)) = 0, \\ x(a) = x_a, \\ \lambda(b) = \frac{\partial \phi}{\partial x}|_b. \end{cases} \quad (۱۳.۵)$$

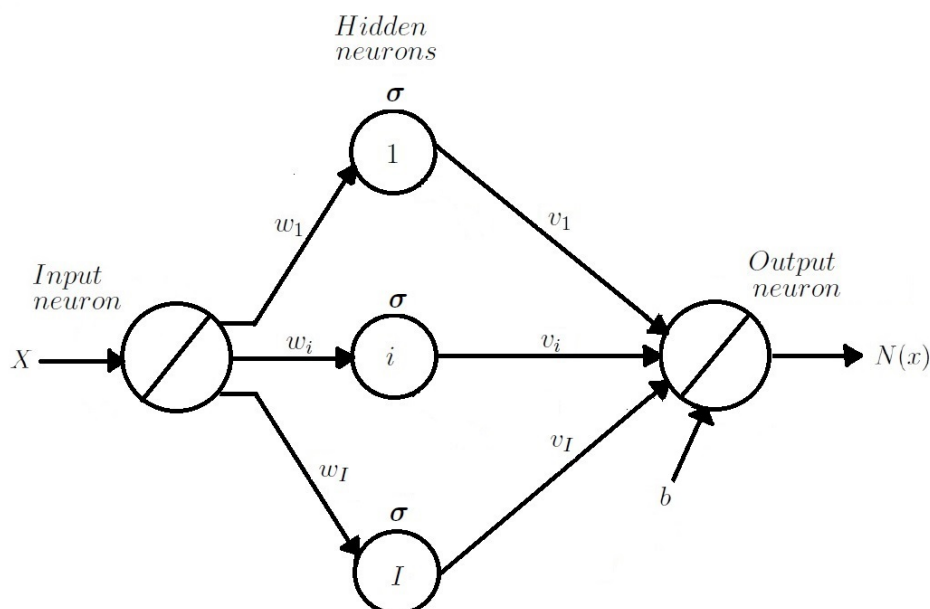
۲.۵ بهینه‌سازی بر پایه شبکه عصبی سری توانی کسری

ابتدا ساختار شبکه عصبی سری توانی معرفی می‌شود و سپس گسسته‌سازی و بهینه‌سازی (۱۳.۵) مورد بررسی قرار می‌گیرد.

³perturbed

۱.۲.۵ ساختار شبکه عصبی سری توانی کسری

روش‌های شبکه عصبی قابلیت زیادی دارند برای اینکه یک طرح قوی و قابل انعطاف برای مدل‌بندی ریاضی و حل پدیده‌های طبیعی باشند. شکل ۳.۶ ساختار یک شبکه عصبی پرسپترون



شکل ۱.۵: نمودار بلوکی ساختار عصبی ارائه شده

با یک ورودی، n نرون پنهان و یک خروجی را نشان می‌دهد [۹۵]. فرض کنید w و v بردار وزن‌های $1 \times n$ و b وزن بایاس باشد. در ساختار معرفی شده، تک ورودی t در وزن‌های v_i ضرب می‌شود و یک ورودی وزن‌دار به دست می‌آید. سپس تابع فعال‌ساز f روی این حاصل‌ضرب اثر می‌کند تا یک نتیجه مناسب حاصل شود. خروجی $N(t)$ با ضرب کردن i -امین نرون لایه پنهان در پارامتر وزن w_i و جمع با بایاس b به دست می‌آید. رابطه بین ورودی و خروجی هر واحد را می‌توان به صورت خلاصه شده در ساختار زیر مشاهده کرد

$$\begin{aligned} O^1 &= t, & \text{واحد ورودی} \\ O_i^* &= f(net_i), \quad net_i = t.v_i, \quad i = 1, \dots, n, & \text{واحد پنهان} \\ N(t) &= \sum_{i=1}^I (O_i^*.w_i) + b. & \text{واحد خروجی} \end{aligned} \quad (14.5)$$

یک پرسپترون سری توانی به صورت زیر بیان می‌شود

$$\sum_{i=0}^{\infty} a_i(t-t_0)^{i\alpha} = a_0 + a_1(t-t_0)^{\alpha} + a_2(t-t_0)^{2\alpha} + \dots, \quad (15.5)$$

که $0 \leq \alpha < n-1$ و $t \geq t_0$. در حالت خاص اگر $t_0 = 0$ باشد، بسط $\sum_{i=0}^{\infty} a_i t^{i\alpha}$ یک سری مک لورن کسری نامیده می‌شود. این روش، جواب $x(t)$ را به یک سری همگرا از مؤلفه‌های

جواب تجزیه می‌کند. بنابراین می‌توان جواب $x(t)$ را به فرم یک سری توانی کسری نوشت

$$x(t) = \sum_{i=0}^{\infty} a_i t^{i\alpha}. \quad (16.5)$$

قرار دهید $a_0 = b$. بنابراین معادله (۱۶.۵) به صورت زیر نوشته می‌شود

$$x(t) = b + \sum_{i=1}^{\infty} a_i t^{i\alpha}. \quad (17.5)$$

اگر تابع جواب به صورت یک سری توانی توسعه یافته قطع شده^۴ بیان شود، تابع جواب می‌تواند به آسانی با پیدا کردن ضرایب مجهول a_i ، برای $i = 1, 2, \dots, I$ تقریب زده شود. بنابراین جواب $x(t)$ توسط سری قطع شده

$$x(t) = b + \sum_{i=1}^I a_i t^{i\alpha}, \quad (18.5)$$

تقریب زده می‌شود. با جایگزین کردن $w_i = a_i$ ، $v_i = t^{i-1}$ و تابع فعال ساز $f(t) = t^\alpha$ در (۱۸.۵)، می‌توان استدلال کرد که سری توانی توسعه یافته قطع شده معادل با خروجی شبکه عصبی طراحی شده به صورت زیر است

$$N(t) = b + \sum_{i=1}^I w_i f(z_i), \quad z_i = v_i t, \quad v_i = t^{i-1}. \quad (19.5)$$

۲.۲.۵ بهینه‌سازی و گسسته‌سازی

در این بخش به دنبال جواب عددی سیستم (۱۳.۵) هستیم. چهار شبکه عصبی برای هر یک از توابع وضعیت (شبکه عصبی آن N_x)، ضرب‌گرهای لاگرانژ (شبکه‌های عصبی آن‌ها N_λ و N_μ) و کنترل (شبکه عصبی آن N_u) برای $i = 1, 2, \dots, n$ به صورت زیر تعریف می‌کنیم

$$\begin{cases} N_x(t) = \sum_{i=1}^I w_x^i f(z_x^i) + b_x, & z_x^i = v_x^i t, & v_x^i = t^{i-1}, \\ N_\lambda(t) = \sum_{i=1}^I w_\lambda^i f(z_\lambda^i) + b_\lambda, & z_\lambda^i = v_\lambda^i t, & v_\lambda^i = t^{i-1}, \\ N_\mu(t) = \sum_{i=1}^I w_\mu^i f(z_\mu^i) + b_\mu, & z_\mu^i = v_\mu^i t, & v_\mu^i = t^{i-1}, \\ N_u(t) = \sum_{i=1}^I w_u^i f(z_u^i) + b_u, & z_u^i = v_u^i t, & v_u^i = t^{i-1}, \end{cases} \quad (20.5)$$

که n تعداد نرون‌ها و $f(z) = z^\alpha$ ، $\alpha \in [0, 1]$ یک تابع فعال‌ساز است. با توجه به مطالبی که قبلاً گفته شد می‌توان به جای جواب‌های (۱۳.۵) از جواب‌های آزمایشی استفاده کرد. این جواب‌های آزمایشی برای توابع وضعیت، ضرب‌گرهای لاگرانژ و کنترل به صورت زیر پیشنهاد

⁴ truncated

$$\begin{cases} x_T(t) = A_1(t) + A_2(t)N_x(t), \\ \lambda_T(t) = A_3(t) + A_4(t)N_\lambda(t), \\ \mu_T(t) = A_5(t) + A_6(t)N_\mu(t), \\ u_T(t) = A_7(t) + A_8(t)N_u(t), \end{cases} \quad (21.5)$$

که $A_1(t), A_3(t), A_5(t), A_7(t)$ توابعی هستند که جواب‌های آزمایشی x_T, λ_T, μ_T و u_T باید شرایط اولیه و انتهایی صدق کنند و فاقد پارامترهای قابل تنظیم هستند. $A_2(t), A_4(t), A_6(t), A_8(t)$ باید طوری مشخص شوند که در شرایط اولیه و انتهایی صفر شوند. با استفاده از شبکه‌های عصبی پیشنهاد شده در (۲۰.۵) جواب‌های آزمایشی را می‌توان با ساختار زیر تعریف کرد

$$\begin{cases} x_T = x(a) + (t-a)N_x, \\ \lambda_T = \frac{\partial \Phi}{\partial x_T}|_b + (t-b)N_\lambda, \\ \mu_T = N_\mu, \\ u_T = N_u. \end{cases} \quad (22.5)$$

جواب‌های آزمایشی (۲۲.۵) باید در شرایط (۱۳.۵) صدق کنند. بنابراین

$$\begin{cases} \frac{\partial H}{\partial x_T}(t, x_T, u_T, \lambda_T, \mu_T) = -A \dot{\lambda}_T(t) + B {}_tD_b^\alpha \lambda_T(t), \\ \frac{\partial H}{\partial \lambda_T}(t, x_T, u_T, \lambda_T, \mu_T) = A \dot{x}_T(t) + B {}_aD_t^\alpha x_T(t), \\ \frac{\partial H}{\partial u_T}(t, x_T, u_T, \lambda_T, \mu_T) = 0, \\ \phi_{FB}^\varepsilon(\mu_T, -G(t, x_T, u_T)) = 0. \end{cases} \quad (23.5)$$

حال برای بیان (۲۳.۵) به صورت یک مسئله بهینه‌سازی نامقید ابتدا بازه $[a, b]$ را با تغییر متغیر

$$t = \frac{b-a}{L}\theta + a,$$

به بازه‌ی $[0, L]$ تبدیل می‌کنیم. در این حالت یک انتخاب مناسب برای نقاط هم‌محلی η_i به صورت

$$\eta_i = \theta_i^{\frac{1}{\alpha}}, i = 0, 1, \dots, m,$$

است که η_i ها نقاط انتخاب شده در بازه $[0, L]$ هستند. یعنی

$$\eta_i = \frac{L}{\varphi} - \frac{L}{\varphi} \cos \frac{\pi i}{m}, \quad i = 0, 1, \dots, m.$$

متناظر با

$$\begin{aligned} {}_aD_t^\alpha (t-a)^{\beta-1}(x) &= \frac{\Gamma(\beta)}{\Gamma(\beta-\alpha)}(x-a)^{(\beta-\alpha-1)}, \\ {}_tD_b^\alpha (b-t)^{\beta-1}(x) &= \frac{\Gamma(\beta)}{\Gamma(\beta-\alpha)}(b-x)^{(\beta-\alpha-1)}, \end{aligned} \quad (24.5)$$

و ساختار شبکه عصبی (۱۹.۵) توصیف شده در فرم کسری، یک تقریب از مشتقات راست ${}_t D_b^\alpha \lambda_T(t)$ و ${}_a D_t^\alpha x_T(t)$ کاپاتو به آسانی محاسبه می‌شود. حال یک مسئله بهبودسازی به صورت

$$\text{minimize}_y E(y) = \frac{1}{\Upsilon} \sum_{j=1}^{\Upsilon} \sum_{k=\circ}^m E_j(\eta_k, y), \quad (25.5)$$

تعریف کنید که $y = (w_x, w_\lambda, w_\mu, w_u, b_x, b_\lambda, b_\mu, b_u, v_x, v_\lambda, v_\mu, v_u)^T \in \mathbb{R}^{\Upsilon I(\Upsilon p+q)}$ و

$$\begin{cases} E_\lambda(\eta_k, y) = \left[\frac{\partial H}{\partial x_T}(\eta_k, x_T, u_T, \lambda_T, \mu_T) + A \dot{\lambda}_T(\eta_k) - B {}_c D_L^\alpha \lambda_T(\eta_k) \right]^\Upsilon, \\ E_\Upsilon(\eta_k, y) = \left[\frac{\partial H}{\partial \lambda_T}(\eta_k, x_T, u_T, \lambda_T, \mu_T) - A \dot{x}_T(\eta_k) - B {}_c D_L^\alpha x_T(\eta_k) \right]^\Upsilon, \\ E_\Upsilon(\eta_k, y) = \left[\frac{\partial H}{\partial u_T}(\eta_k, x_T, u_T, \lambda_T, \mu_T) \right]^\Upsilon, \\ E_\Upsilon(\eta_k, y) = [\phi_{\text{FB}}^\Upsilon(\mu_T, -G(\eta_k, x_T, u_T))]^\Upsilon, \end{cases} \quad (26.5)$$

که $k = \circ, 1, 2, \dots, m$

لم ۱.۲.۵. اگر $y = (w_x^*, w_\lambda^*, w_\mu^*, w_u^*, b_x^*, b_\lambda^*, b_\mu^*, b_u^*, v_x^*, v_\lambda^*, v_\mu^*, v_u^*)^T$ در

$$\gamma(y) = \left[E_\lambda(\eta_1, y), \dots, E_\lambda(\eta_m, y), E_\Upsilon(\eta_1, y), \dots, E_\Upsilon(\eta_m, y), \right. \\ \left. E_\Upsilon(\eta_1, y), \dots, E_\Upsilon(\eta_m, y), E_\Upsilon(\eta_1, y), \dots, E_\Upsilon(\eta_m, y) \right]^T = \circ, \quad (27.5)$$

صدق کند آن گاه y^* جواب بهینه (۲۵.۵) است.

پس مسئله کمینه‌سازی (۲۵.۵) معادل با مسئله زیر است

$$\text{minimize}_y E(y) = \frac{1}{\Upsilon} \|\gamma(y)\|^\Upsilon. \quad (28.5)$$

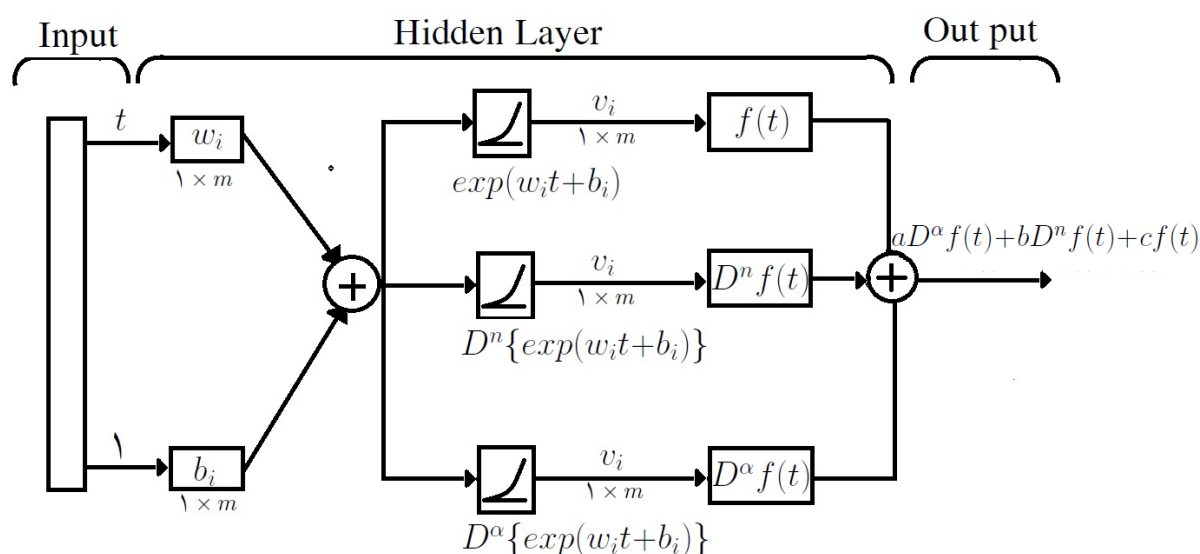
حال می‌توان این مسئله را با هر روش بهبودسازی دلخواه حل کرد. از طرفی با انتخاب تابع $E(y)$ به عنوان یک تابع لیپانف مطابق آنچه در فصل ۴ بررسی شد می‌توان همگرایی و پایداری روش را اثبات کرد.

۳.۵ بهینه‌سازی با کمک شبکه عصبی بر پایه تابع میتاگ-لفلر

در این روش جواب‌های آزمایشی برای توابع وضعیت، ضرب‌گرهای لاگرانژ و کنترل به صورت زیر انتخاب می‌شوند

$$\left\{ \begin{array}{ll} x_T = \sum_{i=1}^I v_i^x \sigma(z_i^x), & z_i^x = w_i^x t + b_i^x, \\ \lambda_T = \sum_{i=1}^I v_i^\lambda \sigma(z_i^\lambda), & z_i^\lambda = w_i^\lambda t + b_i^\lambda, \\ \mu_T = \sum_{i=1}^I v_i^\mu \sigma(z_i^\mu), & z_i^\mu = w_i^\mu t + b_i^\mu, \\ u_T = \sum_{i=1}^I v_i^u \sigma(z_i^u), & z_i^u = w_i^u t + b_i^u, \end{array} \right. \quad (29.5)$$

که I تعداد نرون، v و w پارامترهای وزن و b پارامتر بایاس هستند. همچنین σ تابع نمایی e^x است که به عنوان یک جایگزین از تابع سیگموئید در مدل بندی شبکه عصبی استفاده می‌شود. یک فرم کلی از ساختار شبکه عصبی برای مسائل کنترل بهینه کسری در شکل ۳.۵ نشان داده شد. با جایگذاری جواب‌های آزمایشی (۲۹.۵) در (۱۳.۵) داریم



شکل ۳.۵: نمودار بلوکی ساختار شبکه عصبی ارائه شده.

$$\begin{cases} \frac{\partial H}{\partial x_T}(t, x_T, u_T, \lambda_T, \mu_T) = -A \dot{\lambda}_T(t) + B {}_t D_b^\alpha \lambda_T(t), \\ \frac{\partial H}{\partial \lambda_T}(t, x_T, u_T, \lambda_T, \mu_T) = A \dot{x}_T(t) + B {}_a^c D_t^\alpha x_T(t), \\ \frac{\partial H}{\partial u_T}(t, x_T, u_T, \lambda_T, \mu_T) = 0, \\ \phi_{\text{FB}}^\varepsilon(\mu_T, -G(t, x_T, u_T)) = 0, \quad \varepsilon \rightarrow 0_+, \\ x_T(a) = x_a, \quad \lambda_T(b) = \frac{\partial \phi}{\partial x}|_b. \end{cases} \quad (30.5)$$

مشتق ریمن-لیوویل و کاپاتو از تابع نمایی $f(t) = e^x$ به صورت زیر محاسبه می‌شوند

$$\begin{aligned} {}_a D_t^\alpha(e^{ct}) &= t^{-\alpha} E_{1, 1-\alpha}(ct), \\ {}_a^c D_t^\alpha(e^{ct}) &= ct^{1-\alpha} E_{1, 2-\alpha}(ct), \end{aligned}$$

که $E_{\beta, \gamma}$ تابع میتاگ لفلر دو پارامتری به صورت

$$E_{\alpha, \beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + \beta)}, \quad \alpha, \beta > 0, \quad z \in \mathbb{C}, \quad (31.5)$$

است. با توجه به اینکه در ساختار شبکه عصبی از تابع فعال ساز $f(x) = e^{w_i x + b_i}$ استفاده شده است، مشتق چپ ${}_a^c D_t^\alpha x(t)$ توسط رابطه زیر تعریف می‌شود

$${}_a^c D_t^\alpha x_T(t) = \sum_{i=0}^I w_i e^{b_i} t^{1-\alpha} E_{1, 2-\alpha}(w_i t). \quad (32.5)$$

مشتق راست ${}_t D_b^\alpha \lambda(t)$ با کمک روش انتگرال گیری عددی محاسبه می‌شود. یک مسئله بهینه‌سازی به صورت

$$\text{minimize}_y E(y) = \frac{1}{r} \sum_{j=1}^r \sum_{k=0}^m E_j(\eta_k, y), \quad (33.5)$$

تعریف می‌کنیم که $y = (w_x, w_\lambda, w_\mu, w_u, b_x, b_\lambda, b_\mu, b_u, v_x, v_\lambda, v_\mu, v_u)^T \in \mathbb{R}^{rI(rp+q)}$ و

$$\begin{cases} E_1(\eta_k, y) = \left[\frac{\partial H}{\partial x_T}(\eta_k, x_T, u_T, \lambda_T, \mu_T) + A \dot{\lambda}_T(\eta_k) - B {}_k^c D_L^\alpha \lambda_T(\eta_k) \right]^2, \\ E_2(\eta_k, y) = \left[\frac{\partial H}{\partial \lambda_T}(\eta_k, x_T, u_T, \lambda_T, \mu_T) - A \dot{x}_T(\eta_k) - B {}_a^c D_L^\alpha x_T(\eta_k) \right]^2, \\ E_3(\eta_k, y) = \left[\frac{\partial H}{\partial u_T}(\eta_k, x_T, u_T, \lambda_T, \mu_T) \right]^2, \\ E_4(\eta_k, y) = [\phi_{\text{FB}}^\varepsilon(\mu_T, -G(\eta_k, x_T, u_T))]^2, \\ E_5(\eta_k, y) = [x_T(a) - x_a]^2, \\ E_6(\eta_k, y) = \left[\lambda_T(b) - \frac{\partial \phi}{\partial x} \right]^2, \end{cases} \quad (34.5)$$

که $k = 0, 1, 2, \dots, m$. با کمک لم ۱.۲.۵ مسئله کمینه‌سازی (۳۳.۵) معادل مسئله زیر است

$$\text{minimize}_y E(y) = \frac{1}{r} \|\gamma(y)\|^2. \quad (35.5)$$

حال می‌توان این مسئله را با هر روش بهینه‌سازی دلخواه حل کرد.

۴.۵ مثال‌های عددی

در این فصل چند مثال عددی مورد بررسی قرار می‌گیرد و جواب‌های به‌دست آمده از روش پیشنهاد شده با جواب‌های تحلیلی مقایسه می‌شود. برای بعضی از مثال‌ها یک جدول خطا برای مقادیر مختلف $0 < \alpha < 1$ ارائه شده است. دو مثال اول با هر دو روش عددی ارائه شده حل شده است و بقیه مثال‌ها با کمک روش اول حل شده‌اند. همچنین در مثال اول در دو جدول مجزا برای توابع وضعیت و کنترل نتایج به‌دست آمده از روش اول پیشنهاد شده در این فصل با نتایج به‌دست آمده از چند روش دیگر مقایسه شده‌اند. برای هر مسئله ۵ پارامتر برای وزن‌های ورودی، خروجی و بایاس استفاده می‌کنیم. در همه مثال‌های روش اول از نرم‌افزار GAMS و مثال‌های روش دوم از نرم‌افزار MATLAB برای شبیه‌سازی استفاده شده است. در انتهای فصل یک مسئله کنترل بهینه کسری با مشتق مرتبه متغیر مورد بررسی قرار می‌گیرد و روش اول ارائه شده در این فصل روی آن پیاده‌سازی می‌شود. برای حل این مثال از نرم‌افزار MAPLE استفاده می‌کنیم.

مثال ۱.۴.۵. [۵۶] مسئله زیر را در نظر بگیرید

$$\begin{aligned} \text{minimize} \quad & J = \frac{1}{\tau} \int_0^1 (x^2(t) + u^2(t)) dt, \\ \text{subject to} \quad & {}^C D_t^\alpha x(t) = -x(t) + u(t), \end{aligned}$$

با شرایط اولیه

$$x(0) = 1.$$

جواب‌های تحلیلی مسئله برای $\alpha = 1$ به‌صورت زیر است

$$\begin{aligned} x(t) &= \cosh(\sqrt{2}t) + \beta \sinh(\sqrt{2}t), \\ u(t) &= (1 + \sqrt{2}\beta) \cosh(\sqrt{2}t) + (\sqrt{2} + \beta) \sinh(\sqrt{2}t), \end{aligned}$$

که

$$\beta = -\frac{\cosh(\sqrt{2}t) + \sqrt{2} \sinh(\sqrt{2}t)}{\sqrt{2} \cosh(\sqrt{2}t) + \sinh(\sqrt{2}t)} \approx -0.98.$$

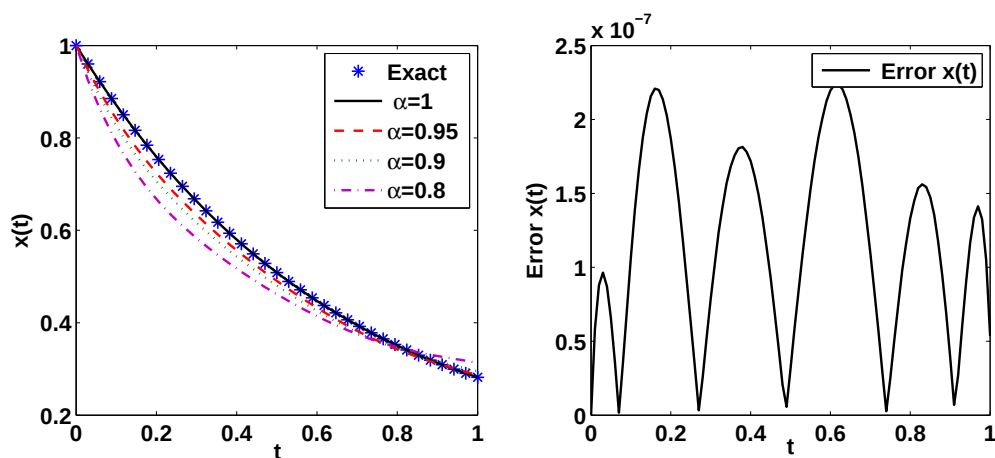
متناظر با (۱۳.۵) داریم

$$\begin{cases} {}_t D_1^\alpha \lambda(t) = x(t) - \lambda(t), \\ {}^C D_t^\alpha x(t) = -x(t) + u(t), \\ u(t) + \lambda(t) = 0, \\ x(0) = 1, \quad \lambda(1) = 0. \end{cases}$$

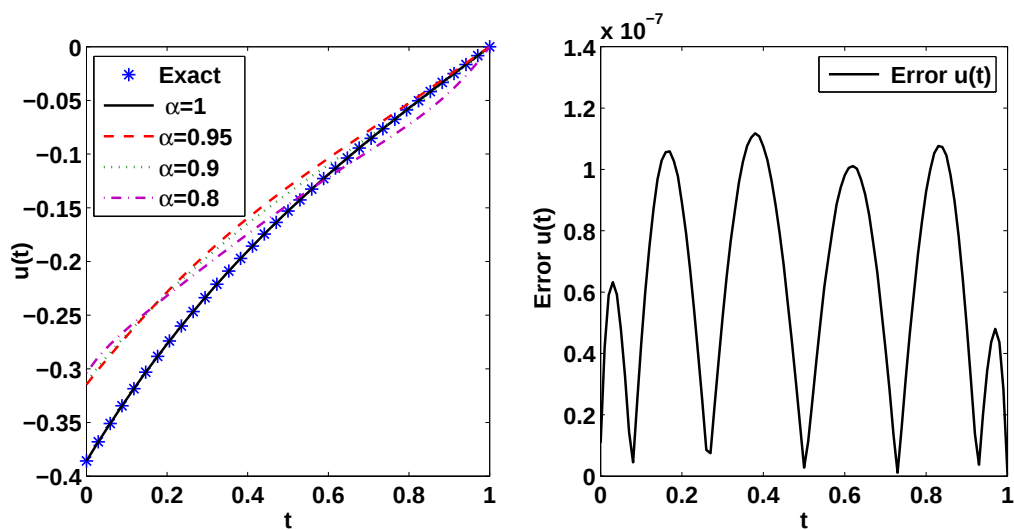
جواب‌های آزمایشی برای روش اول (بر پایه شبکه عصبی سری توانی کسری) با توجه به شرط اولیه به صورت

$$\begin{cases} x_T = 1 + tN_x, \\ \lambda_T = (t - 1)N_\lambda, \\ u_T = N_u, \end{cases}$$

است. با استفاده از روش ارائه شده جواب تقریبی توابع وضعیت و کنترل در حالت $\alpha = 1$ و چند مقدار α به دست آمده است. در شکل‌های ۳.۵ و ۴.۵ مسیر و کنترل بهینه برای $\alpha = 1$ و چند مقدار α و همچنین مسیر خطا برای $\alpha = 1$ رسم شده است. جواب‌های آزمایشی برای



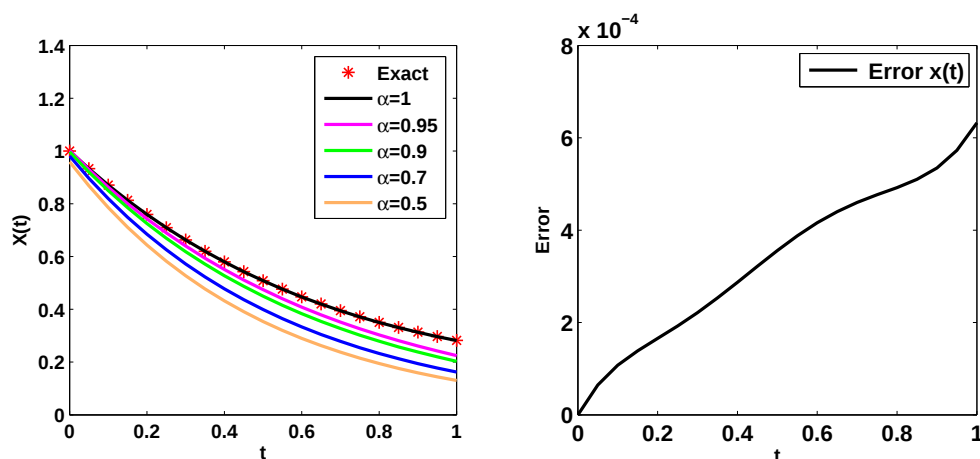
شکل ۳.۵: مسیر بهینه $x(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$.



شکل ۴.۵: کنترل بهینه $u(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$.

$$\begin{cases} x_T = \sum_{i=1}^I v_i^x e^{w_i^x t + b_i^x}, \\ \lambda_T = \sum_{i=1}^I v_i^\lambda e^{w_i^\lambda t + b_i^\lambda}, \\ u_T = \sum_{i=1}^I v_i^u e^{w_i^u t + b_i^u}, \\ x_T(\circ) = \circ, \quad \lambda_T(1) = \circ, \end{cases} \quad (36.5)$$

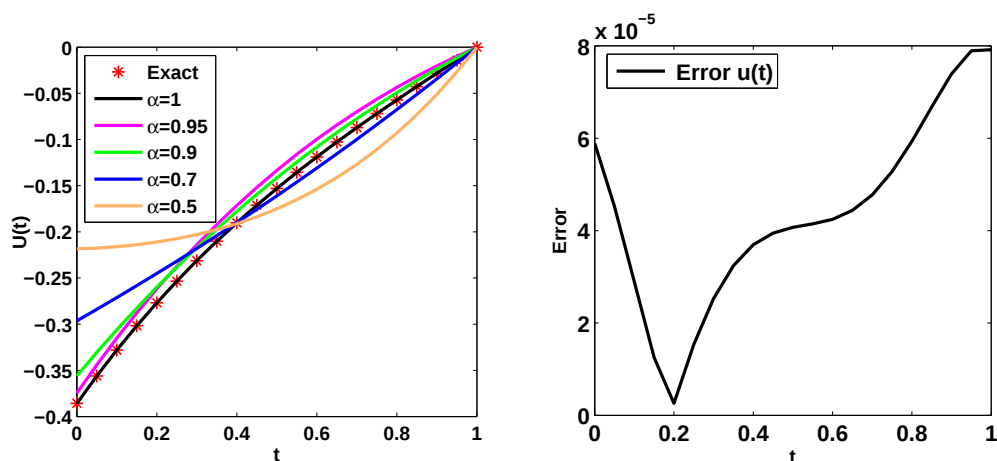
است. در شکل‌های ۵.۵ و ۶.۵ جواب‌های تحلیلی توابع مسیر و کنترل با جواب تقریبی به دست آمده از روش دوم برای $\alpha = 1$ مقایسه شده و خطاهای متناظر با آن‌ها رسم شده است. برای چند مقدار α نیز مسیرهای بهینه و کنترل رسم شده‌اند. با توجه به اینکه روش سری توانی کسری روش قوی‌تری از روش بر پایه تابع میتاگ-لفلر است، جدول‌های ۱.۵ و ۲.۵ به ترتیب مقایسه‌ای بین مقادیر متغیرهای وضعیت و کنترل به دست آمده از چند نوع از چندجمله‌ای‌های مختلف ارائه شده در [۵۶] و روش شبکه عصبی سری توانی کسری ارائه شده در این فصل را نشان می‌دهند.



شکل ۵.۵: مسیر بهینه $x(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8, 0.7, 0.5$ و خطای مطلق برای $\alpha = 1$.

مثال ۲.۴.۵. یک مسئله بهینه‌سازی با محدودیت‌های نامساوی را به صورت زیر در نظر بگیرید

$$\begin{aligned} & \text{minimize} && \int_0^1 u(t) dt, \\ & \text{subject to} && {}^c D_t^\alpha x(t) = u(t), \\ & && x(\circ) = 1, \\ & && u(t) \geq \circ, \\ & && x(t) - u(t) \geq \circ. \end{aligned}$$



شکل ۶.۵: کنترل بهینه $u(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8, 0.7, 0.5$ و خطای مطلق برای $\alpha = 1$.

Time (s)	Analytical Solution	Shifted Legendre	Power Series	Shifted Chebyshev 1st kind	Neural Network
0	۱/۰۰۰۰۰۰۰	۱/۰۰۰۰۰۰۰	۱/۰۰۰۰۰۰۰	۰/۹۹۹۹۹۹۹	۱/۰۰۰۰۰۰۰
۰/۱	۰/۸۷۰۹۷۲۷	۰/۸۷۰۹۷۳۱	۰/۸۷۰۸۶۰۰	۰/۸۷۰۹۷۲۸	۰/۸۷۰۹۷۲۰
۰/۲	۰/۷۵۹۳۹۳۳	۰/۷۵۹۳۹۳۶	۰/۷۵۹۳۱۴۴	۰/۷۵۹۳۹۳۳	۰/۷۵۹۳۹۳۰
۰/۳	۰/۶۶۳۰۲۷۵	۰/۶۶۳۰۲۷۸	۰/۶۶۲۹۸۱۹	۰/۶۶۳۰۲۷۵	۰/۶۶۳۰۲۷۴
۰/۴	۰/۵۷۹۹۴۴۳	۰/۵۷۹۹۴۴۷	۰/۵۷۹۹۰۰۴	۰/۵۷۹۹۴۴۶	۰/۵۷۹۹۴۴۳
۰/۵	۰/۵۰۸۴۷۹۳	۰/۵۰۸۴۷۹۷	۰/۵۰۸۴۲۳۷	۰/۵۰۸۴۷۹۵	۰/۵۰۸۴۷۹۱
۰/۶	۰/۴۴۷۲۰۰۸	۰/۴۴۷۲۰۱۰	۰/۴۴۷۱۴۳۴	۰/۴۴۷۲۰۰۸	۰/۴۴۷۲۰۰۳
۰/۷	۰/۳۹۴۸۸۱۲	۰/۳۹۴۸۸۱۴	۰/۳۹۴۸۳۷۶	۰/۳۹۴۸۸۱۲	۰/۳۹۴۸۸۱۰
۰/۸	۰/۳۵۰۴۷۲۵	۰/۳۵۰۴۷۲۸	۰/۳۵۰۴۴۲۱	۰/۳۵۰۴۷۲۷	۰/۳۵۰۴۷۲۶
۰/۹	۰/۳۰۳۰۸۵۱	۰/۳۱۳۰۸۵۳	۰/۳۱۳۰۵۰۱	۰/۳۱۳۰۸۵۱	۰/۳۱۳۰۸۵۰
۱	۰/۲۸۱۹۶۹۶	۰/۲۸۱۹۶۹۸	۰/۲۸۱۹۳۴۹	۰/۲۸۱۹۶۹۶	۰/۲۸۱۹۶۹۱

جدول ۱.۵: مقایسه بین جواب تحلیلی و چند روش تقریبی برای مسیر $x(t)$.

جواب تحلیلی توابع وضعیت و کنترل برای $\alpha = 1$ ، $x(t) = \exp(t)$ و $u(t) = \exp(t)$ هستند [۸۵]. چون $x(1)$ آزاد است داریم $\lambda(1) = 0$. با توجه به این شرط و شرط اولیه $x(0) = 1$ ،

۸۴ حل رده‌ای از مسائل کنترل بهینه کسری با محدودیت‌های مساوی و نامساوی با دو رویکرد متفاوت از شبکه‌های عصبی

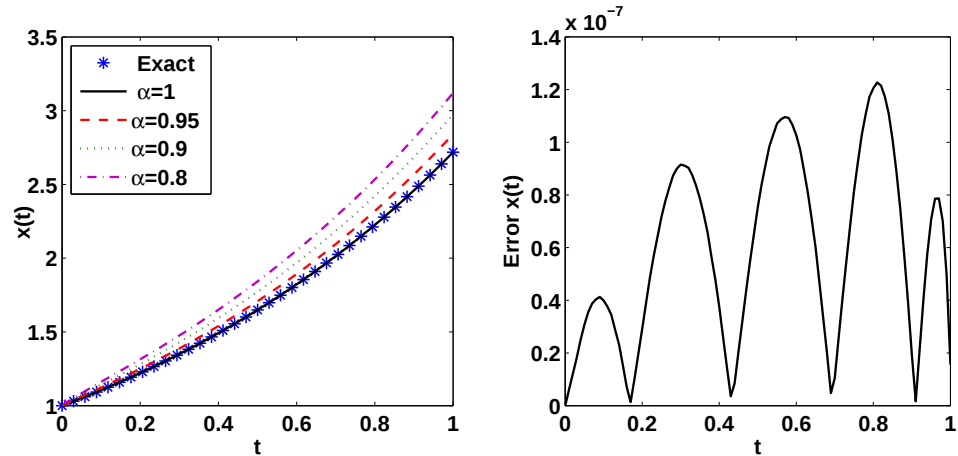
Time (s)	Analytical Solution	Shifted Legendre	Power Series	Shifted Chebyshev 1st kind	Neural Network
۰	-۰/۳۸۵۸۱۸۵	-۰/۳۸۵۸۱۸۷	-۰/۳۷۵۷۸۹۵	-۰/۳۸۵۸۱۸۷	-۰/۳۸۵۸۱۸۷
۰/۱	-۰/۳۲۸۰۶۰۱	-۰/۳۲۸۰۶۰۲	-۰/۳۲۸۰۳۸۸	-۰/۳۲۸۰۶۰۲	-۰/۳۲۸۰۶۰۰
۰/۲	-۰/۲۷۶۸۷۳۸	-۰/۲۷۶۸۷۳۹	-۰/۲۷۶۸۴۸۵	-۰/۲۷۶۸۷۳۹	-۰/۲۷۶۸۷۳۷
۰/۳	-۰/۲۳۱۲۳۴۲	-۰/۲۳۱۲۳۴۳	-۰/۲۳۱۲۱۷۷	-۰/۲۳۱۲۳۴۳	-۰/۲۳۱۲۳۴۳
۰/۴	-۰/۱۹۰۲۲۷۰	-۰/۱۹۰۲۲۷۰	-۰/۱۹۰۲۲۴۷	-۰/۱۹۰۲۲۷۱	-۰/۱۹۰۲۲۷۱
۰/۵	-۰/۱۵۳۰۳۰۷	-۰/۱۵۳۰۳۰۶	-۰/۱۵۳۰۳۵۷	-۰/۱۵۳۰۳۰۶	-۰/۱۵۳۰۳۰۷
۰/۶	-۰/۱۱۸۹۰۰۱	-۰/۱۱۸۹۰۰۰	-۰/۱۱۸۹۰۲۵	-۰/۱۱۸۹۰۰۰	-۰/۱۱۸۹۰۰۱
۰/۷	-۰/۰۸۷۱۵۱۴	-۰/۰۸۷۱۵۱۴	-۰/۰۸۷۱۵۰۹	-۰/۰۸۷۱۵۱۵	-۰/۰۸۷۱۵۱۵
۰/۸	-۰/۰۵۷۱۴۸۷	-۰/۰۵۷۱۴۸۷	-۰/۰۵۷۱۵۷۶	-۰/۰۵۷۱۴۸۹	-۰/۰۵۷۱۴۹۰
۰/۹	-۰/۰۲۸۲۹۱۰	-۰/۰۲۸۲۹۰۷	-۰/۰۲۸۳۱۷۴	-۰/۰۲۸۲۹۰۹	-۰/۰۲۸۲۹۱۲
۱	۰/۰۰۰۰۰۰۰	-۴/۵۵E - ۰۹	۷/۵۴۱E - ۰۹	-۲/۰۹E - ۰۹	۰

جدول ۲.۵: مقایسه بین جواب تحلیلی و چند روش تقریبی برای کنترل $u(t)$.

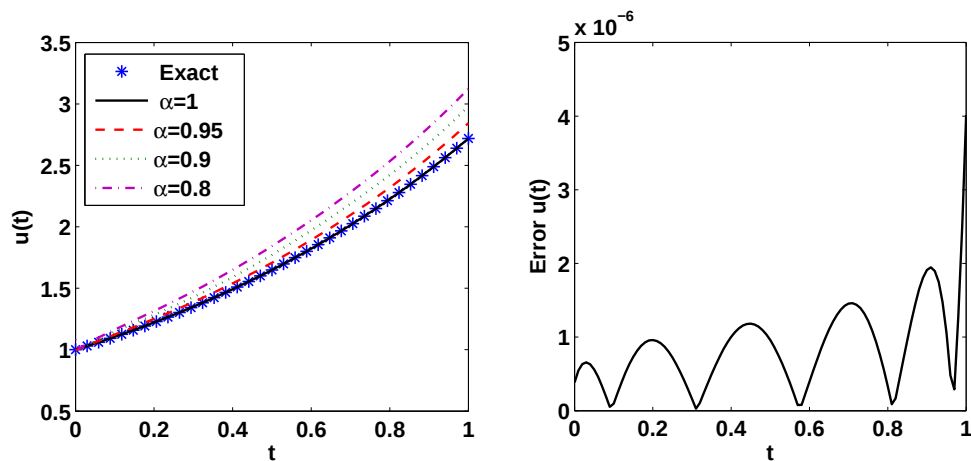
جواب‌های آزمایشی برای روش سری توانی کسری به صورت زیر انتخاب می‌شود

$$\begin{cases} x_T = 1 + tN_x, \\ \lambda_T = (t - 1)N_\lambda, \\ u_T = N_u, \\ \mu_{1T} = N_{\mu_1}, \\ \mu_{2T} = N_{\mu_2}. \end{cases}$$

نتایج تقریبی به دست آمده برای $x(t)$ و $u(t)$ برای چند مقدار α و خطای مطلق برای $\alpha = 1$ به ترتیب در شکل‌های ۷.۵ و ۸.۵ نشان داده می‌شوند. با توجه به شرط اولیه، جواب‌های



شکل ۷.۵: مسیر بهینه $x(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$.



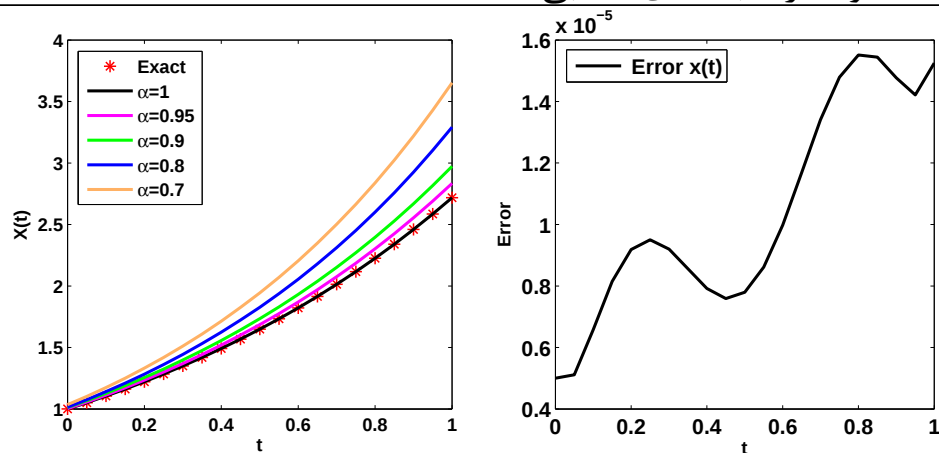
شکل ۸.۵: کنترل بهینه $u(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$.

آزمایشی بر پایه میتاگ-لفلر به صورت

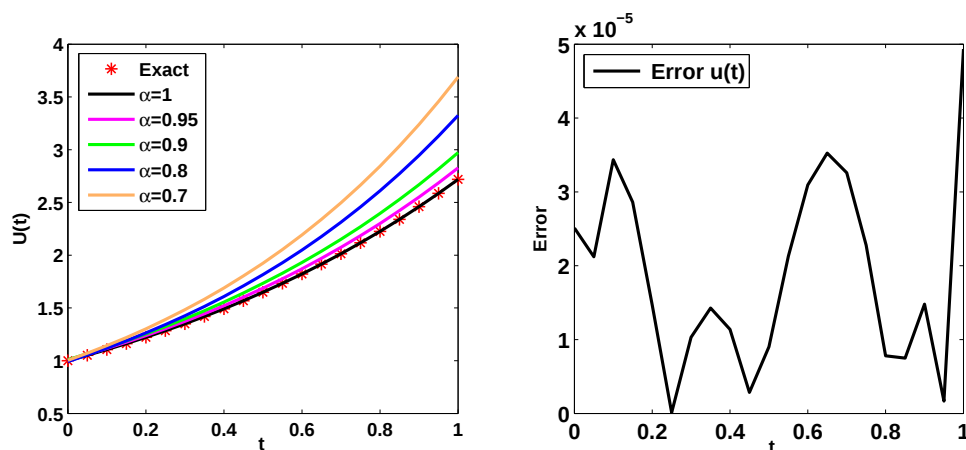
$$\begin{cases} x_T = \sum_{i=1}^I v_i^x e^{w_i^x t + b_i^x}, \\ \lambda_T = \sum_{i=1}^I v_i^\lambda e^{w_i^\lambda t + b_i^\lambda}, \\ u_T = \sum_{i=1}^I v_i^u e^{w_i^u t + b_i^u}, \\ \mu_{\lambda TT} = \sum_{i=1}^I v_i^{\mu_\lambda} e^{w_i^{\mu_\lambda} t + b_i^{\mu_\lambda}}, \\ \mu_{u TT} = \sum_{i=1}^I v_i^{\mu_u} e^{w_i^{\mu_u} t + b_i^{\mu_u}}, \\ x_T(\circ) = \circ, \quad \lambda_T(1) = \circ, \end{cases} \quad (37.5)$$

است. نتایج تقریبی به دست آمده برای $x(t)$ و $u(t)$ برای چند مقدار α و خطای مطلق برای $\alpha = 1$ به ترتیب در شکل‌های ۹.۵ و ۱۰.۵ نشان داده می‌شود. همچنین مقادیر تابع خطا در $E(y)$ در (۲۰.۶) در جدول ۳.۵ برای چند مقدار $1 > \alpha > 0$ برای هر دو روش نشان داده

۸۶ حل رده‌ای از مسائل کنترل بهینه کسری با محدودیت‌های مساوی و نامساوی با دو رویکرد متفاوت از شبکه‌های عصبی



شکل ۹.۵: مسیر بهینه $x(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8, 0.7, 0.5$ و خطای مطلق برای $\alpha = 1$.



شکل ۱۰.۵: کنترل بهینه $u(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8, 0.7, 0.5$ و خطای مطلق برای $\alpha = 1$.

می‌شوند.

	α	۰/۹۵	۰/۹	۰/۸	۰/۷	۰/۵
روش اول	E	$1/1 \times 10^{-2}$	$1/8 \times 10^{-2}$	$2/8 \times 10^{-2}$	$3/2 \times 10^{-2}$	$3/1 \times 10^{-2}$
روش دوم	E	$2/4 \times 10^{-2}$	$3/9 \times 10^{-2}$	$4/8 \times 10^{-2}$	$4/2 \times 10^{-2}$	$7/2 \times 10^{-2}$

جدول ۳.۵: مقدار تابع خطا $E(y)$ برای چند مقدار $0 < \alpha < 1$.

با توجه به قوی‌تر بودن روش اول، مثال‌های بعدی با کمک این روش حل می‌شوند.

مثال ۳.۴.۵. [۸۵] مسئله بهینه‌سازی زیر را در نظر بگیرید

$$\begin{aligned} \text{minimize} \quad & -\ln(2) \int_0^1 x(t) dt, \\ \text{subject to} \quad & {}^C D_t^\alpha x(t) = \ln(2)(x(t) - u(t)), \\ & x(0) = 0, \\ & |u(t)| \leq 1, \\ & x(t) + u(t) \leq 2. \end{aligned}$$

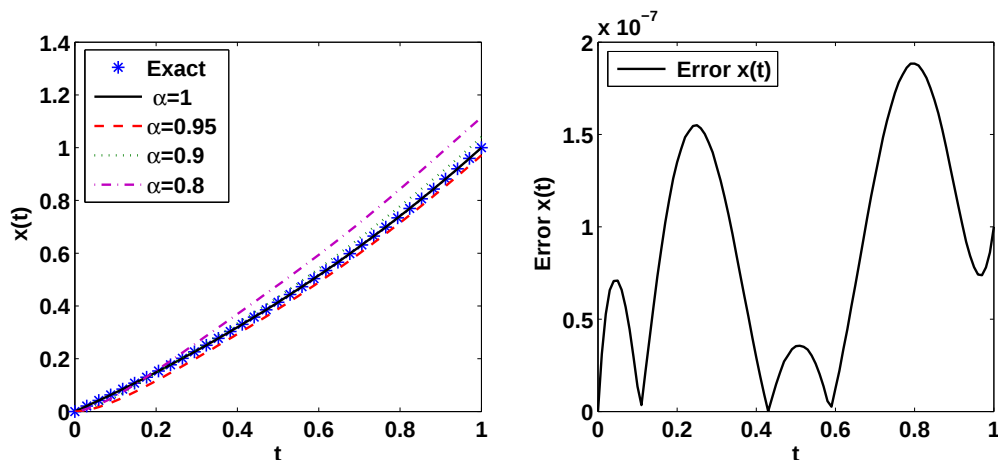
جواب‌های دقیق برای توابع کنترل و وضعیت به صورت زیر هستند

$$\begin{cases} x(t) = \exp(\ln 2(t)) - 1, \\ u(t) = 1. \end{cases}$$

جواب‌های آزمایشی به صورت زیر انتخاب می‌شوند

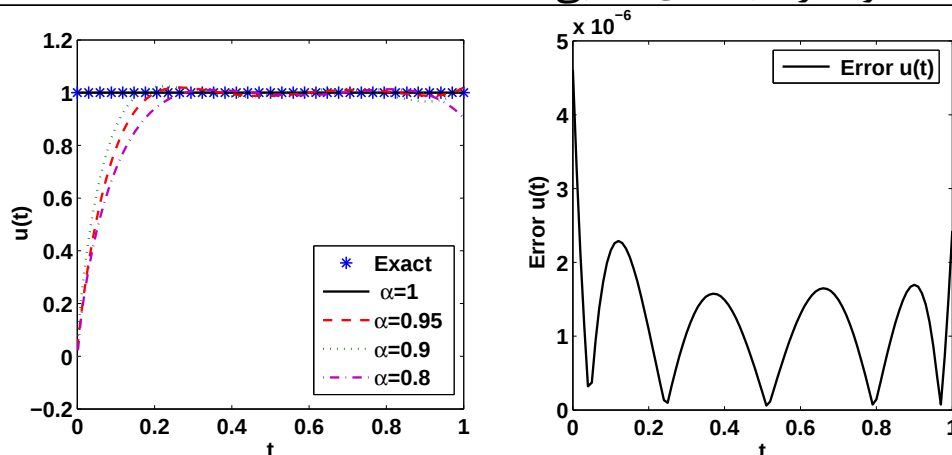
$$\begin{cases} x_T = tN_x, \\ \lambda_T = (t-1)N_\lambda, \\ u_T = N_u, \\ \mu_{1T} = N_{\mu_1}, \\ \mu_{2T} = N_{\mu_2}, \\ \mu_{3T} = N_{\mu_3}. \end{cases}$$

جواب‌های تقریبی برای $x(t)$ و $u(t)$ برای چند مقدار α و خطاهای مطلق متناظر با آن‌ها برای $\alpha = 1$ را می‌توان در شکل‌های ۱۱.۵ و ۱۲.۵ مشاهده کرد. همچنین مقادیر تابع خطای $E(y)$ برای چند مقدار $0 < \alpha < 1$ در جدول ۴.۵ ارائه شده است.



شکل ۱۱.۵: مسیر بهینه $x(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$.

۸۸ حل رده‌ای از مسائل کنترل بهینه کسری با محدودیت‌های مساوی و نامساوی با دو رویکرد متفاوت از شبکه‌های عصبی



شکل ۱۲.۵: کنترل بهینه $u(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$.

α	۰/۹۵	۰/۹	۰/۸	۰/۷	۰/۵
E	$2/1 \times 10^{-3}$	$2/6 \times 10^{-3}$	$2/6 \times 10^{-3}$	$4/4 \times 10^{-3}$	$4/8 \times 10^{-2}$

جدول ۴.۵: مقدار تابع خطا $E(y)$ برای چند مقدار $0 < \alpha < 1$.

مثال ۴.۴.۵. [۸۵] در این مثال می‌خواهیم کمینه تابعی

$$\int_0^3 2x_1(t)dt,$$

با محدودیت‌ها و شرایط مرزی

$$\begin{aligned} {}^c D_t^\alpha x_1(t) &= x_2(t), \\ {}^c D_t^\alpha x_2(t) &= u(t), \\ |u(t)| &\leq 2, \\ -7 &\leq x_1(t), \\ x_1(0) &= 2, \quad x_2(0) = 0, \end{aligned}$$

را به‌دست آوریم. جواب‌های دقیق برای توابع کنترل و وضعیت برای $\alpha = 1$ به‌صورت زیر هستند

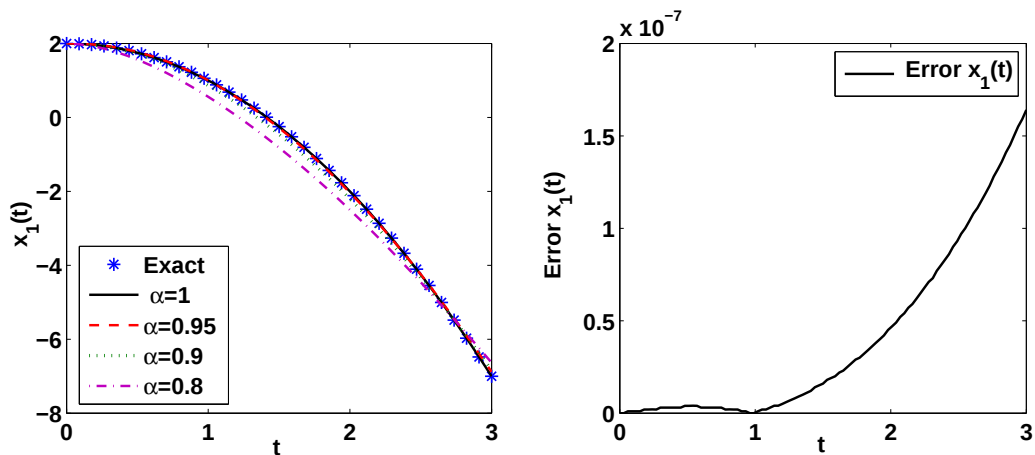
$$\begin{cases} x_1(t) = 2 - t^2, \\ x_2(t) = -2t, \\ u(t) = -2. \end{cases}$$

چون $x_1(3)$ و $x_2(3)$ آزاد هستند، داریم $\lambda_1(3) = \lambda_2(3) = 0$. با توجه به این شرایط و شرایط

اولیه $x_1(0) = 2, x_2(0) = 0$ ، جواب‌های آزمایشی به صورت

$$\begin{cases} x_{1T} = 2 + tN_{x_1}, \\ x_{2T} = tN_{x_2}, \\ \lambda_{1T} = (t-3)N_{\lambda_1}, \\ \lambda_{2T} = (t-3)N_{\lambda_2}, \\ u_T = N_u, \\ \mu_{1T} = N_{\mu_1}, \\ \mu_{2T} = N_{\mu_2}, \\ \mu_{3T} = N_{\mu_3}, \end{cases}$$

هستند. نتایج روش پیشنهادی برای توابع $x_1(t)$ ، $x_2(t)$ و $u(t)$ برای چند مقدار α و خطای مطلق متناظر با آن‌ها برای $\alpha = 1$ را می‌توان به ترتیب در شکل‌های ۱۳.۵-۱۵.۵ مشاهده کرد. همچنین مقادیر تابع خطا $E(y)$ برای چند مقدار $0 < \alpha < 1$ در جدول ۵.۵ مشخص شده است.

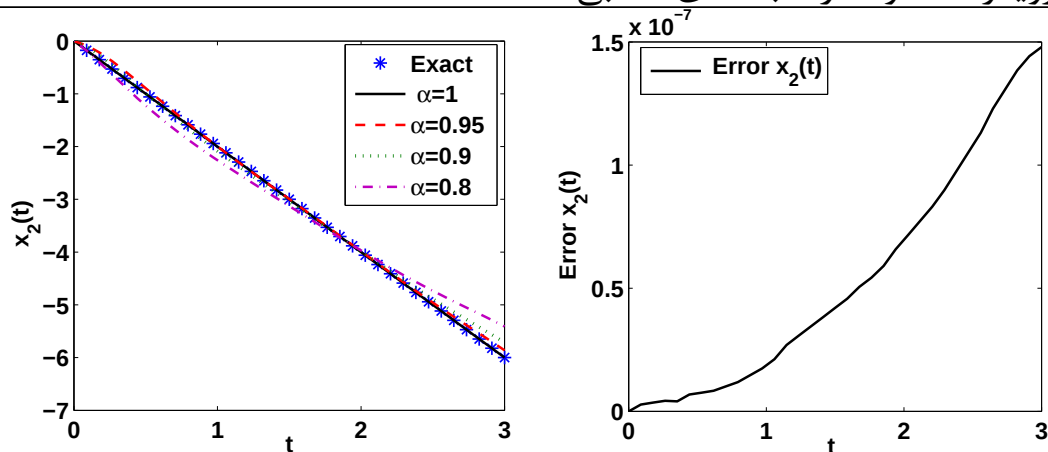


شکل ۱۳.۵: مسیر بهینه $x_1(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$.

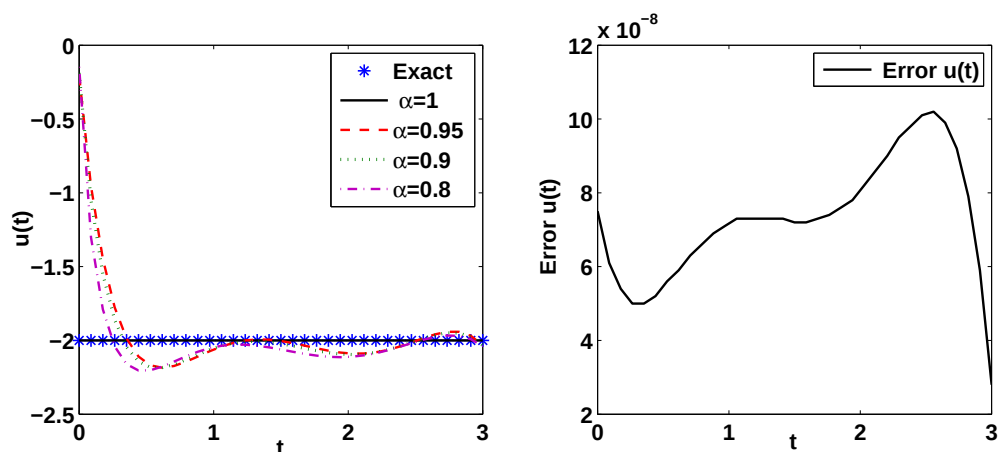
α	۰/۹۵	۰/۹	۰/۸	۰/۷	۰/۵
E	$1/2 \times 10^{-3}$	$4/3 \times 10^{-3}$	$1/3 \times 10^{-3}$	$3/8 \times 10^{-3}$	$4/5 \times 10^{-3}$

جدول ۵.۵: مقدار تابع خطای $E(y)$ برای چند مقدار $0 < \alpha < 1$.

۹۰ حل رده‌ای از مسائل کنترل بهینه کسری با محدودیت‌های مساوی و نامساوی با دو رویکرد متفاوت از شبکه‌های عصبی



شکل ۱۴.۵: مسیر بهینه $x_2(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$.



شکل ۱۵.۵: کنترل بهینه $u(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$.

مثال ۵.۴.۵. مسئله بولزای^۵ [۸۵]

$$\begin{aligned} & \text{minimize} && \frac{H}{2} x^2(T) + \int_0^T \frac{1}{2} u^2(t) dt, \\ & \text{subject to} && {}^c D_t^\alpha x(t) = -ax(t) + u(t), \\ & && x(0) = 5, \end{aligned}$$

را در نظر بگیرید. با فرض $H = 5$ ، $a = -0.2$ و $T = 15$ ، جواب دقیق برای توابع کنترل و

⁵Bolza

وضعیت برای $\alpha = 1$ به صورت زیر است

$$\begin{cases} x(t) = B \exp\left(\frac{-t}{\delta}\right) - \delta A \exp\left(\frac{t}{\delta}\right), \\ u(t) = -2A \exp\left(\frac{t}{\delta}\right), \end{cases}$$

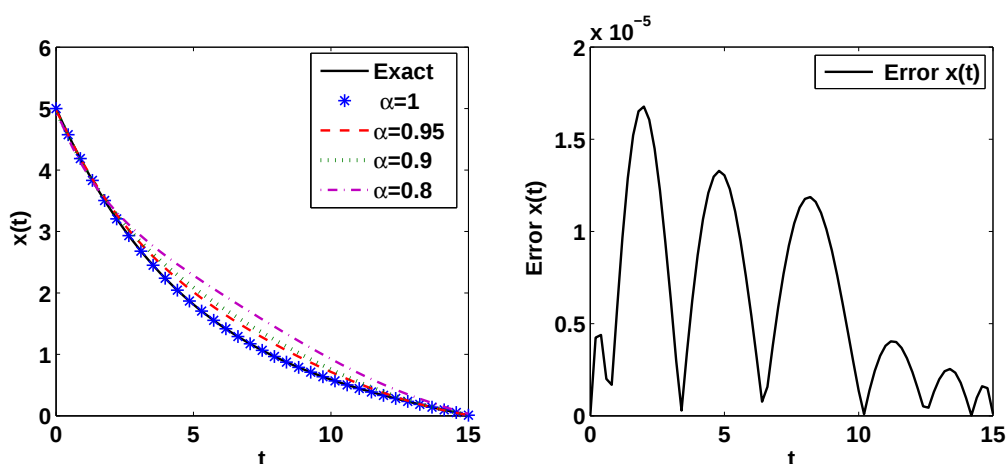
که

$$A = \frac{25}{26 \exp(6) - 25}, \quad B = \frac{26 A \exp(6)}{\delta}.$$

چون $x(15)$ آزاد است داریم $\lambda(15) = \delta x(15)$. جواب‌های آزمایشی به صورت زیر هستند

$$\begin{cases} x_T = \delta + t N_x, \\ \lambda_T = (t - 15) N_\lambda + \delta x(15), \\ u_T = N_u. \end{cases}$$

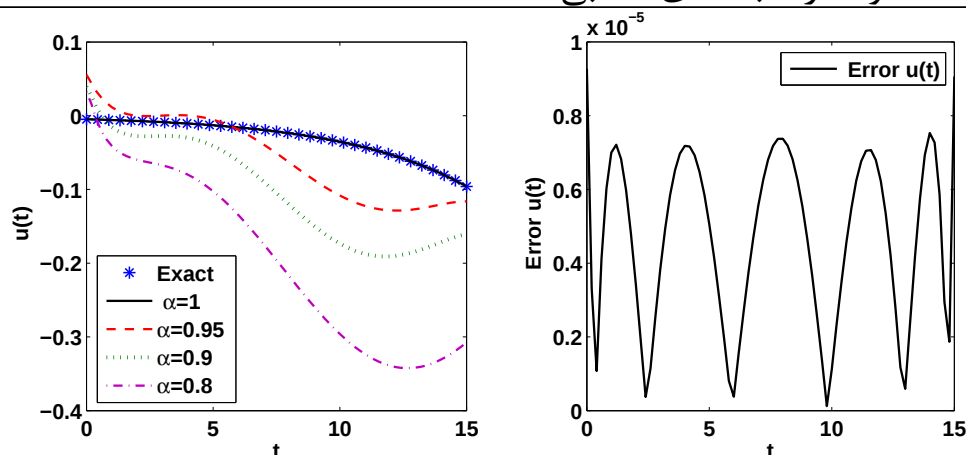
مسیر بهینه و کنترل بهینه برای چند مقدار α و تابع خطای متناظر با آن‌ها در شکل‌های ۱۶.۵ و ۱۷.۵ برای $\alpha = 1$ رسم شده‌اند. همچنین جدول ۶.۵ مقدار تابع خطا $E(y)$ را نشان می‌دهد.



شکل ۱۶.۵: مسیر بهینه $x(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$.

α	۰/۹۵	۰/۹	۰/۸	۰/۷	۰/۵	۰/۳
E	$3/4 \times 10^{-4}$	$7/1 \times 10^{-4}$	$3/7 \times 10^{-3}$	$2/8 \times 10^{-3}$	$2/5 \times 10^{-2}$	$3/1 \times 10^{-2}$

جدول ۶.۵: مقدار تابع خطا $E(y)$ برای چند مقدار $0 < \alpha < 1$.



شکل ۱۷.۵: کنترل بهینه $u(\cdot)$ برای $\alpha = 1, 0.95, 0.9, 0.8$ و خطای مطلق برای $\alpha = 1$.

مثال ۶.۴.۵. در این مثال و مثال بعدی یک مسئله بهینه‌سازی با مشتقات مرتبه صحیح و کسری با نقطه انتهایی ثابت بررسی می‌شود. فرض کنید که می‌خواهید تابعی زیر را کمینه کنید

$$J(x, u) = \int_0^1 (u^2(t) - f(x(t)))^2 dt,$$

که محدودیت‌های دینامیکی آن به صورت

$$\dot{x}(t) + {}^C D_t^{2/5} x(t) = u(t) + \frac{2}{\Gamma(2/5)} t^{1/5}, \quad t \in [0, 1],$$

با شرایط مرزی

$$x(0) = 0, \quad x(1) = 1,$$

است. جواب‌های دقیق برای $\alpha = 0.5$ جفت

$$(x(t), u(t)) = (t^2, 2t),$$

است [۷۰]. متناظر با تابع همیلتونی

$$H = (u^2(t) - f(x(t)))^2 + \lambda(t)(u(t) + \frac{2}{\Gamma(2/5)} t^{1/5}),$$

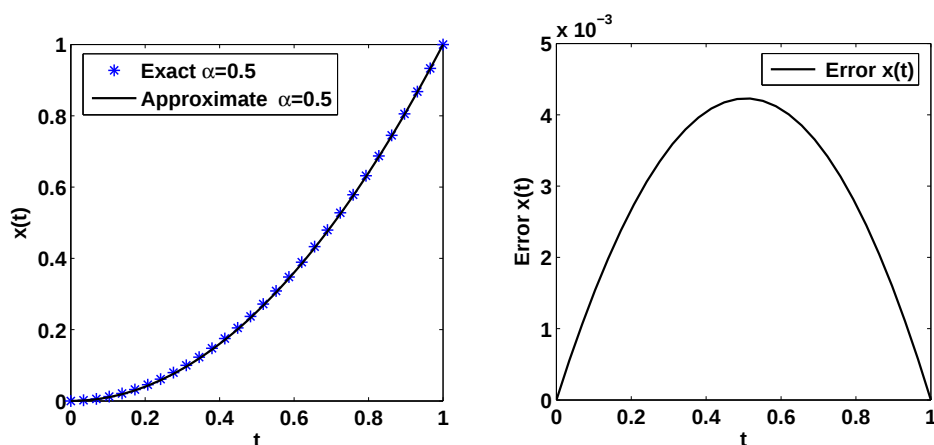
شرایط بهینگی عبارت است از

$$\begin{cases} -\frac{\partial H}{\partial x} = \dot{\lambda}(t) - {}_t D_t^\alpha \lambda(t) = (u^2(t) - f(x(t))), \\ \frac{\partial H}{\partial \lambda} = \dot{x}(t) + {}^C D_t^\alpha x(t) = u(t) + \frac{2}{\Gamma(2/5)} t^{1/5}, \\ \frac{\partial H}{\partial u} = f_u(t)(u^2(t) - f(x(t))) + \lambda(t) = 0. \end{cases}$$

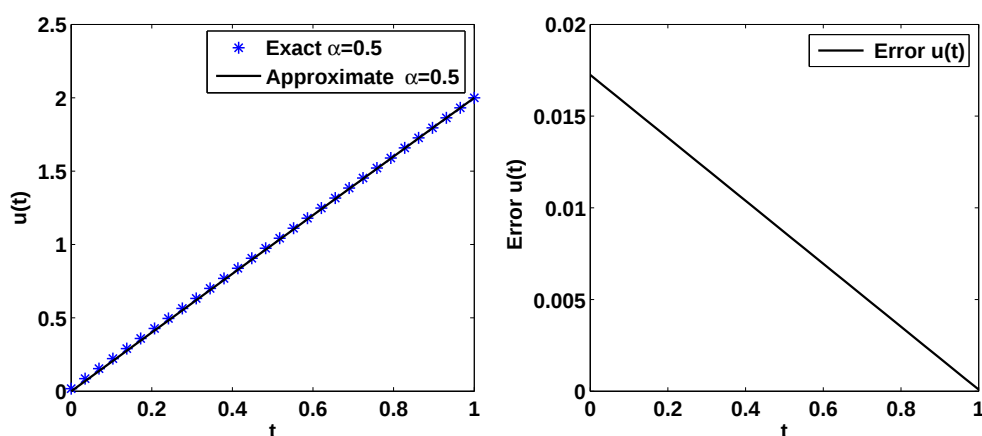
با توجه به شرایط اولیه جواب‌های آزمایشی به صورت

$$\begin{cases} x_T = t + t(1-t)N_x, \\ \lambda_T = N_\lambda, \\ u_T = N_u, \end{cases}$$

هستند. با کمک رویکرد ارائه شده در این فصل مسیر بهینه، کنترل بهینه و همچنین خطای مطلق آن‌ها برای $\alpha = 0.5$ در شکل‌های ۱۸.۵ و ۱۹.۵ رسم شده است.



شکل ۱۸.۵: مسیر بهینه $x(\cdot)$ خطای مطلق برای $\alpha = 0.5$.



شکل ۱۹.۵: کنترل بهینه $u(\cdot)$ و خطای مطلق برای $\alpha = 0.5$.

مثال ۷.۴.۵. [۷۲] مسئله کنترل بهینه کسری زیر را در نظر بگیرید

$$\text{minimize } J(x, u) = \int_0^1 (tu(t) - (\alpha + 2)x(t))^2 dx,$$

$$\dot{x}(t) + {}^C D_t^\alpha x(t) = u(t) + t^2,$$

$$x(0) = 0, \quad x(1) = \frac{2}{\Gamma(3 + \alpha)}.$$

جواب دقیق برای $\alpha = 0.5$ به صورت زوج زیر است

$$(x(t), u(t)) = \left(\frac{2t^{(\alpha+2)}}{\Gamma(\alpha+3)}, \frac{2t^{\alpha+1}}{\Gamma(\alpha+2)} \right).$$

متناظر با تابع همیلتونی

$$H = (tu(t) - (\alpha + 2)x(t))^2 + \lambda(t)(u(t) + t^2),$$

شرایط بهینگی را می‌توان به صورت زیر نوشت

$$\begin{cases} -\frac{\partial H}{\partial x} = \dot{\lambda}(t) - {}_t D_t^\alpha \lambda(t) = 2(\alpha + 2)(tu(t) - (\alpha + 2)x(t)), \\ \frac{\partial H}{\partial \lambda} = \dot{x}(t) + {}^C D_t^\alpha x(t) = u(t) + t^2, \\ \frac{\partial H}{\partial u} = 2t(tu(t) - (\alpha + 2)x(t)) + \lambda(t) = 0. \end{cases}$$

همچنین جواب‌های آزمایشی عبارت است از

$$\begin{cases} x_T = \frac{2}{\Gamma(3 + \alpha)}t + t(1 - t)N_x, \\ \lambda_T = N_\lambda, \\ u_T = N_u. \end{cases}$$

مسیرهای بهینه و کنترل بهینه را می‌توان در شکل‌های ۲۰.۵ و ۲۱.۵ مشاهده کرد.

در مثال آخر یک مسئله کنترل بهینه کسری با مشتق مرتبه متغیر بررسی می‌شود.

مثال ۸.۴.۵. [۱۰۶] مسئله کنترل بهینه کسری از مرتبه متغیر به صورت

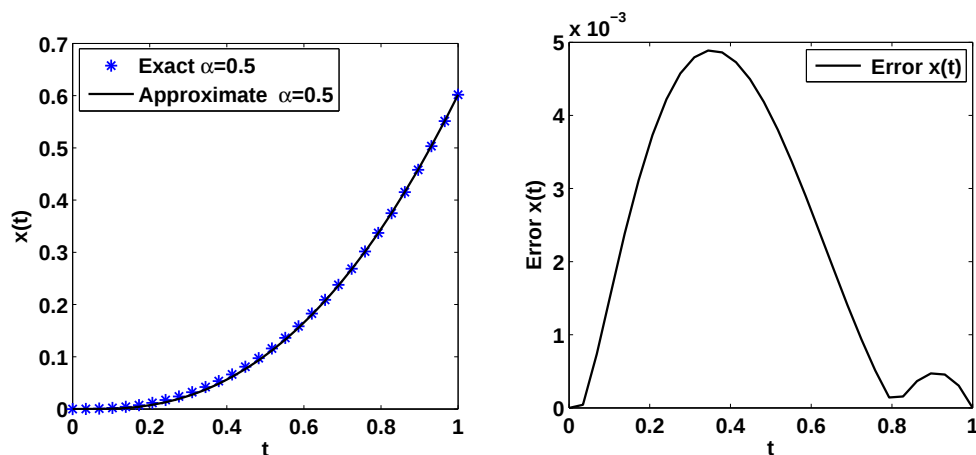
$$\text{minimize } J(x(t), u(t)) = \int_0^1 [(x(t) - t^{\frac{1}{2}})^2 + (1 + t^2)(u(t) + t^2 - \frac{15\sqrt{\pi}}{8\Gamma(\frac{3}{2} - \alpha(t))} t^{\frac{1}{2} - \alpha(t)})^2] dt,$$

با سیستم دینامیکی

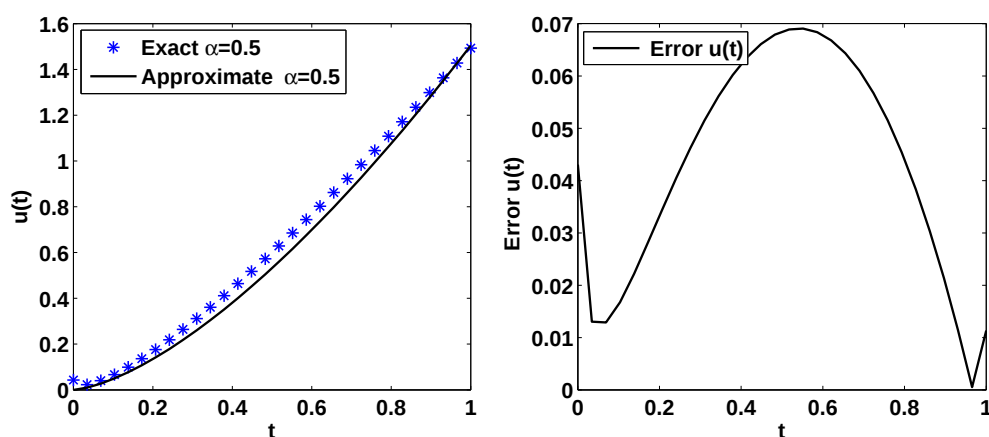
$${}^C D_t^{\alpha(t)} x(t) = tx(t)^2 + u(t), \quad 1 < \alpha(t) \leq 2,$$

و شرایط اولیه

$$x(0) = \dot{x}(0) = 0,$$



شکل ۲۰.۵: مسیر بهینه $x(\cdot)$ و خطای مطلق برای $\alpha = 0.5$.



شکل ۲۱.۵: کنترل بهینه $u(\cdot)$ و خطای مطلق برای $\alpha = 0.5$.

را در نظر بگیرید. جواب‌های تحلیلی برای متغیرهای وضعیت و کنترل به صورت

$$x(t) = t^{\frac{5}{2}}, \quad u(t) = \frac{15\sqrt{\pi}}{\lambda\Gamma(\frac{5}{2} - \alpha(t))} t^{\frac{5}{2} - \alpha(t)} - t^6,$$

هستند. تابع همیلتونی عبارت است از

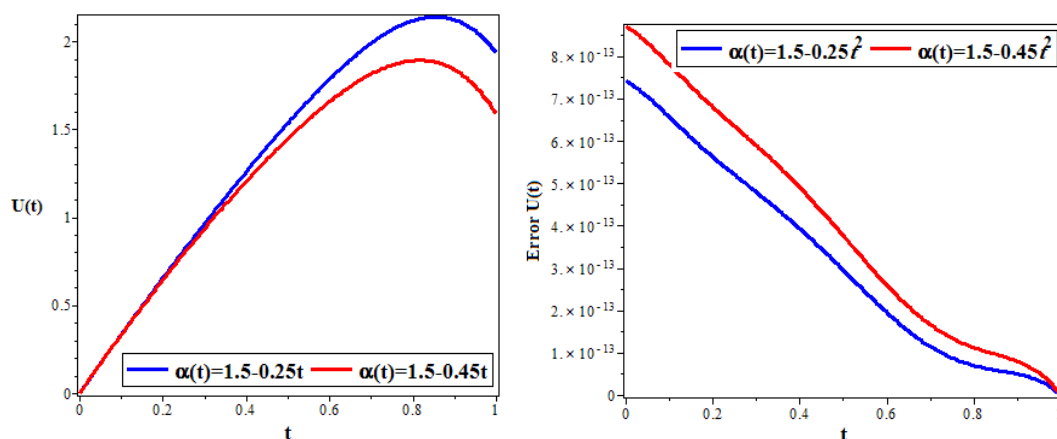
$$H = (x(t) - t^{\frac{5}{2}})^4 + (1 + t^2)(u(t) + t^6 - \frac{15\sqrt{\pi}}{\lambda\Gamma(\frac{5}{2} - \alpha(t))} t^{\frac{5}{2} - \alpha(t)})^2 + \lambda(t)(tx(t)^2 + u(t)),$$

که $\lambda(t)$ ضرب‌گر لاگرانژ است. شرایط بهینگی به صورت زیر نوشته می‌شود

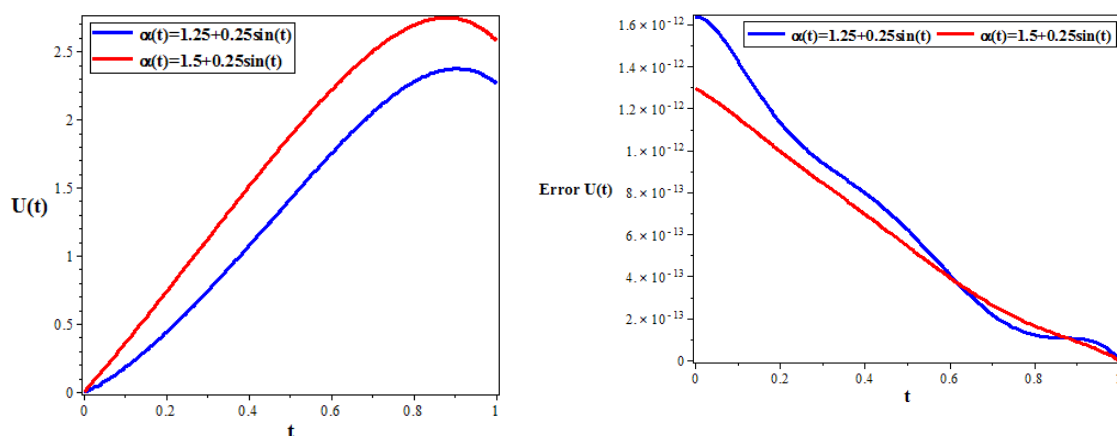
$$\begin{cases} {}^C D_t^{\alpha(t)} x(t) = tx(t)^2 + u(t), \\ {}_t D_1^{\alpha(t)} \lambda(t) = 4(x(t) - t^{\frac{5}{2}})^3 + 2t\lambda(t)x(t), \\ 2(1 + t^2)(u(t) + t^6 - \frac{15\sqrt{\pi}}{\lambda\Gamma(\frac{5}{2} - \alpha(t))} t^{\frac{5}{2} - \alpha(t)}) + \lambda(t) = 0. \end{cases}$$

$$\begin{cases} x_T = t^{\gamma} N_x, \\ \lambda_T = (1-t) N_{\lambda}, \\ u_T = N_u. \end{cases} \quad (38.5)$$

این مسئله با $N = 8$ نقطه گسسته‌سازی شده در بازه $[0, 1]$ برای مرتبه‌های کسری $\alpha(t) = 1/5 + 0.25 \sin(t)$ و $\alpha(t) = 1/5 - 0.45t^2$ ، $\alpha(t) = 1/5 + 0.25 \sin(t)$ ، $\alpha(t) = 1/5 - 0.45t^2$ حل شده است. برای شبیه‌سازی از نرم‌افزار MAPLE استفاده شده است. شکل‌های ۲۲.۵ و ۲۳.۵ رفتار جواب‌های عددی متغیر کنترل $u(t)$ برای چند $\alpha(t)$ را نشان می‌دهند. همچنین مقایسه‌ای بین خطاهای مطلق روش اول ارائه شده در این فصل و روش ارائه شده در [۱۰۶] را می‌توان در جدول‌های ۷.۵ و ۸.۵ مشاهده کرد.



شکل ۲۲.۵: مسیر جواب‌های عددی متغیر کنترل $u(t)$ برای چند $\alpha(t)$.



شکل ۲۳.۵: مسیر جواب‌های عددی متغیر کنترل $u(t)$ برای چند $\alpha(t)$.

$$\alpha(t) = 1/5 + 0/25 \sin(t)$$

t	Method of [۱۰۶]		Neural Network Method	
	$e_u(t)$	$e_x(t)$	$e_u(t)$	$e_x(t)$
۰/۱	$1/79E - 6$	$5/33E - 5$	$8/91E - 14$	$1/10E - 4$
۰/۳	$1/06E - 6$	$1/00E - 4$	$3/95E - 13$	$8/90E - 5$
۰/۵	$6/99E - 7$	$1/36E - 4$	$1/15E - 12$	$8/10E - 5$
۰/۷	$4/87E - 7$	$1/69E - 4$	$2/64E - 13$	$6/06E - 5$
۰/۹	$2/58E - 7$	$2/08E - 4$	$8/14E - 13$	$8/09E - 6$

جدول ۷.۵: مقایسه بین روش ارائه شده در [۱۰۶] با $N = 15$ و روش اول ارائه شده در این فصل با $N = 8$.

$$\alpha(t) = 1/5 - 0/45t^2$$

t	Method of [۱۰۶]		Neural Network Method	
	$e_u(t)$	$e_x(t)$	$e_u(t)$	$e_x(t)$
۰/۱	$1/20E - 7$	$5/06E - 5$	$7/99E - 14$	$9/15E - 5$
۰/۳	$7/90E - 8$	$9/12E - 5$	$1/66E - 13$	$7/70E - 5$
۰/۵	$5/35E - 8$	$1/17E - 4$	$3/76E - 13$	$4/50E - 5$
۰/۷	$3/78E - 8$	$1/38E - 4$	$5/89E - 13$	$1/03E - 5$
۰/۹	$2/01E - 8$	$1/68E - 4$	$7/80E - 13$	$1/12E - 5$

جدول ۸.۵: مقایسه بین روش ارائه شده در [۱۰۶] با $N = 15$ و روش اول ارائه شده در این فصل با $N = 8$.

فصل ۶

مسئله کنترل بهینه کسری با یک کاربرد در سیستم‌های آشوبناک کسری

مسائل کسری برای اینکه بتوانند مسئله را به صورت عددی حل کنند از تقریب زدن عملگرها استفاده می‌کنند. برای مثال مشتق کسری ریمن-لیوویل قابل بسط دادن به یک سری توانی شامل تنها مشتق مرتبه صحیح است [۲۰]. اگر x یک تابع تحلیلی باشد آن‌گاه

$${}_a D_t^\alpha (x(t)) = \sum_{k=1}^{\infty} \binom{\alpha}{k} \frac{(t-a)^k - \alpha}{\Gamma(k+1-\alpha)} x^{(k)}(t), \quad (1.6)$$

که

$$\binom{\alpha}{k} = \frac{(-1)^{k-1} \alpha \Gamma(k-\alpha)}{\Gamma(1-\alpha) \Gamma(1+k)}.$$

از معایب آشکار استفاده (۱.۶) در محاسبات عددی این است، که برای داشتن یک خطای کوچک، باید تعداد زیادی از جملات را جمع کرد. بنابراین تابع باید دارای مشتقات مرتبه بالاتر باشد که برای کنترل بهینه مناسب نیست. برای مقابله با این مشکل روش دیگری در [۶۸] ارائه شده است که تقریب خوبی برای مشتق کسری توابع بدون نیاز به وجود مشتقات مراتب بالاتر است. این روش برای مشتقات کسری ریمن-لیوویل چپ تعریف شده است.

استفاده از این تقریب باعث می‌شود که یک مسئله کنترل بهینه کسری به یک مسئله کنترل بهینه غیرکسری تبدیل شود. شرط لازم بهینگی برای مسئله کنترل بهینه غیرکسری در این مسئله به دست آمده اعمال می‌شود و با یک ساختار شبکه عصبی که در ادامه معرفی می‌شود می‌توان مسئله را حل کرد.

۱.۶ تقریب مشتق کسری ریمن-لیوویل

مشتق کسری ریمن-لیوویل می‌تواند به صورت زیر بسط داده شود

$${}_a D_t^\alpha x(t) = A(\alpha)(t-a)^{-\alpha}x(t) + B(\alpha)(t-a)^{1-\alpha}\dot{x}(t) - \sum_{p=2}^{\infty} C(\alpha, p)(t-a)^{1-p-\alpha}y_p(t), \quad (2.6)$$

که $\alpha \in (0, 1)$ ، $x \in C^2[a, b]$ و $y_p(t)$ جواب سیستم زیر است

$$\begin{cases} \dot{y}_p(t) = (1-p)(t-a)^{p-2}y(t), \\ y_p(t) = 0. \end{cases} \quad (3.6)$$

همچنین A ، B و C به صورت زیر تعریف می‌شوند

$$\begin{aligned} A(\alpha) &= \frac{1}{\Gamma(1-\alpha)} \left[1 + \sum_{p=2}^{\infty} \frac{\Gamma(p-1+\alpha)}{\Gamma(\alpha)(p-1)!} \right], \\ B(\alpha) &= \frac{1}{\Gamma(2-\alpha)} \left[1 + \sum_{p=1}^{\infty} \frac{\Gamma(p-1+\alpha)}{\Gamma(\alpha-1)p!} \right], \end{aligned} \quad (4.6)$$

$$C(\alpha, p) = \frac{\Gamma(p-1+\alpha)}{\Gamma(2-\alpha)\Gamma(\alpha-1)(p-1)!}, \quad p = 2, 3, \dots$$

با توجه به رابطه‌ای که بین مشتق کسری ریمن-لیوویل و کاپاتو وجود دارد و در فصل اول بیان شد، می‌توان تقریبی مشابه برای مشتق کاپاتو در نظر گرفت. در تقریب‌ها تنها از یک مجموع متناهی به صورت

(5.6)

$${}_a D_t^\alpha x(t) \approx A(\alpha, k)(t-a)^{-\alpha}x(t) + B(\alpha, k)(t-a)^{1-\alpha}\dot{x}(t) - \sum_{p=2}^k C(\alpha, p)(t-a)^{1-p-\alpha}y_p(t),$$

استفاده می‌شود که $k \geq 2$ و

$$A(\alpha, k) = \frac{1}{\Gamma(1-\alpha)} \left[1 + \sum_{p=2}^k \frac{\Gamma(p-1+\alpha)}{\Gamma(\alpha)(p-1)!} \right], \quad (6.6)$$

$$B(\alpha, k) = \frac{1}{\Gamma(2-\alpha)} \left[1 + \sum_{p=1}^k \frac{\Gamma(p-1+\alpha)}{\Gamma(\alpha-1)p!} \right].$$

خطای این تقریب کراندار است و به صورت زیر است

$$|E_{tr}(t)| \leq \max_{\tau \in [a, t]} |\ddot{x}(\tau)| \frac{\exp((1-\alpha)^2 + 1 - \alpha)}{\Gamma(2-\alpha)(1-\alpha)k^{1-\alpha}} (t-a)^{2-\alpha}. \quad (7.6)$$

برای اثبات و جزئیات بیشتر [۷۱] را ببینید.

۲.۶ مدل بندی ریاضی مسئله

این فصل یک رده‌ای از مسائل کنترل بهینه کسری با مشتق کاپاتو را به صورت زیر بررسی می‌کند

$$\begin{aligned} \text{minimize} \quad & J(x, u) = \int_a^b L(t, x, u) dt, \\ \text{subject to} \quad & M \dot{x}(t) + N {}^c D_t^\alpha x(t) = f(t, x, u), \\ & x(a) = x_a, \quad x(b) = x_b, \end{aligned} \quad (8.6)$$

که $x(t) \in \mathbb{R}^p$ ، $(M, N) \neq (0, 0)$ و $0 < \alpha < 1$. ابتدا عملگر ${}^c D_t^\alpha x(t)$ با کمک رابطه (۱۰.۱) و تقریب (۵.۶) جایگزین می‌شود. برای سادگی در بحث فرض کنید $A = A(\alpha, k)$ ، $B = B(\alpha, k)$ و $p = 2, 3, \dots$ ، $C_p = C(\alpha, p)$ بنابراین

$$\dot{x}(t) = \frac{f(t, x(t), u(t)) - NA(t-a)^{-\alpha} x(t) + \sum_{p=2}^k NC_p(t-a)^{1-p-\alpha} y_p(t) + \frac{Nx_a(t-a)^{-\alpha}}{\Gamma(1-\alpha)}}{M + NB(t-a)^{1-\alpha}}. \quad (9.6)$$

همچنین تعریف کنید $Y(t) = (y_2(t), y_3(t), \dots, y_k(t))$ و

$$F(t, x(t), Y(t), u(t)) = \frac{f(t, x(t), u(t)) - NA(t-a)^{-\alpha} x(t) + \sum_{p=2}^k NC_p(t-a)^{1-p-\alpha} y_p(t) + \frac{Nx_a(t-a)^{-\alpha}}{\Gamma(1-\alpha)}}{M + NB(t-a)^{1-\alpha}}.$$

بنابراین مسئله (۸.۶) به مسئله بهینه‌سازی زیر تبدیل می‌شود

$$\begin{aligned} \text{minimize} \quad & J(x, u) = \int_a^b L(t, x, u) dt, \\ \text{subject to} \quad & \dot{x}(t) = F(t, x(t), Y(t), u(t)), \\ & \dot{y}_p(t) = (1-p)(t-a)^{p-2} x(t), \quad p = 2, \dots, k, \\ & x(a) = x_a, \\ & x(b) = x_b, \\ & y_p(a) = 0, \quad p = 2, \dots, k. \end{aligned} \quad (10.6)$$

با توجه به [۸۵]، اگر (x, u) یک جواب بهینه مسئله (۱۰.۶) باشد آن‌گاه $\lambda(t)$ و $\Upsilon = (\gamma_1, \dots, \gamma_k)$ وجود دارند که $(x, Y, u, \lambda, \Upsilon)$ در شرایط زیر صدق می‌کند

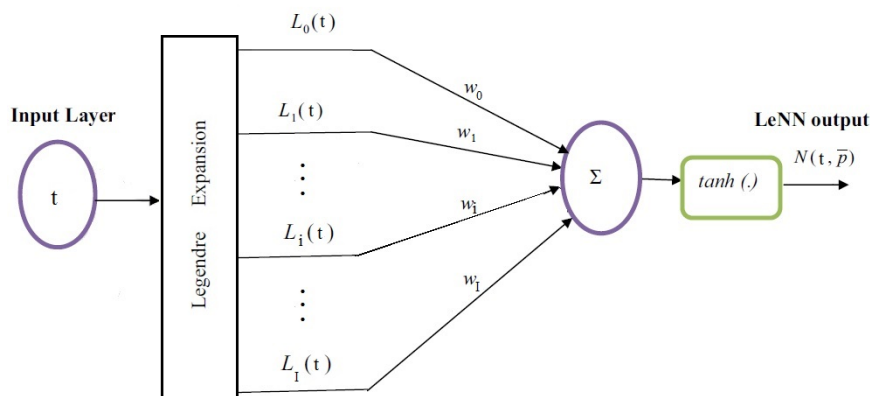
$$\left\{ \begin{array}{l} \frac{\partial H}{\partial x}(t, x, Y, u, \lambda, \Upsilon) = -\dot{\lambda}(t), \\ \frac{\partial H}{\partial \lambda}(t, x, Y, u, \lambda, \Upsilon) = \dot{x}(t), \\ \frac{\partial H}{\partial Y}(t, x, Y, u, \lambda, \Upsilon) = -\dot{\Upsilon}(t), \\ \frac{\partial H}{\partial \Upsilon}(t, x, Y, u, \lambda, \Upsilon) = \dot{Y}(t), \\ \frac{\partial H}{\partial u}(t, x, Y, u, \lambda, \Upsilon) = 0, \\ x(a) = x_a, \\ x(b) = x_b, \\ y_p(a) = 0, \quad p = 1, \dots, k, \end{array} \right. \quad (11.6)$$

که λ و γ_p ، $p = 1, \dots, k$ ضرب‌گرهای لاگرانژ هستند و H تابع همیلتونی به صورت زیر است

$$H(t, x, Y, u, \lambda, \Upsilon) = L(t, x, u) + \lambda F(t, x, Y, u) + \sum_{p=1}^k \gamma_p (1-p)(t-a)^{p-1} x(t).$$

۳.۶ ساختار شبکه عصبی و بهینه‌سازی

شکل ۳.۶ ساختار یک شبکه عصبی لژاندر تک لایه را نشان می‌دهد که شامل یک ورودی تنها، یک لایه خروجی و بلوک بسط تابعی بر پایه چندجمله‌ای‌های لژاندر است [۱۰۷]. در ساختار



شکل ۱.۶: نمودار بلوکی شبکه عصبی لژاندر.

نمایش داده شده مقدار خروجی با ورودی t و بردار وزن $\bar{P}$ به صورت زیر تعریف می‌شود

$$\begin{cases} N(t, \bar{P}) = \sigma(z), \\ z = \sum_{i=0}^I w_i L_i(t), \end{cases} \quad (12.6)$$

که $L_0(t), L_1(t), \dots, L_I(t)$ عناصر پردازش (چندجمله‌ای‌های لژاندر) و w_i و $i = 0, 1, 2, \dots, I$ وزن‌های متناظر با آن‌ها هستند. همچنین σ یک تابع فعال‌ساز به صورت زیر است

$$\sigma(z) = \tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}. \quad (13.6)$$

برای حل عددی (۱۱.۶) شبکه‌های زیر را برای توابع وضعیت، ضرب‌گرهای لاگرانژ و کنترل می‌توان در نظر گرفت

$$\begin{cases} N_x(t) = \sigma(z^x), & z^x = \sum_{i=1}^I w_i^x L_i(t), \\ N_Y(t) = \sigma(z^Y), & z^Y = \sum_{i=1}^I w_i^Y L_i(t), \\ N_\lambda(t) = \sigma(z^\lambda), & z^\lambda = \sum_{i=1}^I w_i^\lambda L_i(t), \\ N_\Upsilon(t) = \sigma(z^\Upsilon), & z^\Upsilon = \sum_{i=1}^I w_i^\Upsilon L_i(t), \\ N_u(t) = \sigma(z^u), & z^u = \sum_{i=1}^I w_i^u L_i(t), \end{cases} \quad (14.6)$$

که σ یک تابع فعال‌ساز به صورت (۱۳.۶) است. با توجه به رابطه‌ای که بین جواب‌های به دست آمده از شرایط بهینگی (۱۱.۶) و جواب‌های آزمایشی که در فصل‌های قبل بحث شد وجود دارد می‌توان جواب‌های آزمایشی برای توابع وضعیت، کنترل و ضرب‌گرهای لاگرانژ را به صورت زیر تعریف کرد

$$\begin{cases} x_T(t) = x(a) + (t-a)N_x(t), \\ Y_T(t) = tN_Y(t), \\ \lambda_T(t) = N_\lambda(t), \\ \Upsilon_T(t) = (t-b)N_\Upsilon(t), \\ u_T(t) = N_u(t). \end{cases} \quad (15.6)$$

با توجه به اینکه برای مسئله با نقطه انتهایی آزاد $\lambda(\cdot) = 0$ ، می‌توان λ_T در (۱۵.۶) را به صورت $\lambda_T = (t-b)N_\lambda$ تعریف کرد. جواب‌های آزمایشی (۱۵.۶) باید در شرایط (۱۱.۶) صدق کنند.

بنابراین

$$\begin{cases} \frac{\partial H}{\partial x_T}(t, x_T, Y_T, u_T, \lambda_T, \Upsilon_T) = -\dot{\lambda}_T(t), \\ \frac{\partial H}{\partial \lambda_T}(t, x_T, Y_T, u_T, \lambda_T, \Upsilon_T) = \dot{x}_T(t), \\ \frac{\partial H}{\partial Y_T}(t, x_T, Y_T, u_T, \lambda_T, \Upsilon_T) = -\dot{\Upsilon}_T(t), \\ \frac{\partial H}{\partial \Upsilon_T}(t, x_T, Y_T, u_T, \lambda_T, \Upsilon_T) = \dot{Y}_T(t), \\ \frac{\partial H}{\partial u_T}(t, x_T, Y_T, u_T, \lambda_T, \Upsilon_T) = 0. \end{cases} \quad (۱۶.۶)$$

از طرفی برای تبدیل (۱۶.۶) به یک مسئله کمینه‌سازی نامقید، ابتدا بازه $[a, b]$ را با یک تغییر متغیر مناسب به بازه $[-1, 1]$ تبدیل می‌کنیم. در این حالت یک انتخاب مناسب برای نقاط هم‌محلی ریشه‌های جندجمله‌ای لژاندر در بازه $[-1, 1]$ هستند. صورت کلی متناظر با تابع خطا برای دستگاه معادلات دیفرانسیل معمولی (۱۶.۶) می‌تواند به صورت زیر فرمول‌بندی شود

$$\text{minimize}_P E(P) = \frac{1}{\Upsilon} \sum_{j=1}^{\Delta} \sum_{l=1}^m E_j(t_l, P), \quad (۱۷.۶)$$

که $P = (w_x, w_Y, w_\lambda, w_\Upsilon, w_u, v_x, v_Y, v_\lambda, v_\Upsilon, v_u)^T$ و

$$\begin{cases} E_1(t_l, P) = \left[\frac{\partial H}{\partial x_T}(t, x_T, Y_T, u_T, \lambda_T, \Upsilon_T) + \dot{\lambda}_T(t) \right]^2, \\ E_2(t_l, P) = \left[\frac{\partial H}{\partial \lambda_T}(t, x_T, Y_T, u_T, \lambda_T, \Upsilon_T) - \dot{x}_T(t) \right]^2, \\ E_3(t_l, P) = \left[\frac{\partial H}{\partial Y_T}(t, x_T, Y_T, u_T, \lambda_T, \Upsilon_T) + \dot{\Upsilon}_T(t) \right]^2, \\ E_4(t_l, P) = \left[\frac{\partial H}{\partial \Upsilon_T}(t, x_T, Y_T, u_T, \lambda_T, \Upsilon_T) - \dot{Y}_T(t) \right]^2, \\ E_5(t_l, P) = \left[\frac{\partial H}{\partial u_T}(t, x_T, Y_T, u_T, \lambda_T, \Upsilon_T) \right]^2. \end{cases} \quad (۱۸.۶)$$

لم ۱.۳.۶. اگر $P = (w_x, w_Y, w_\lambda, w_\Upsilon, w_u, v_x, v_Y, v_\lambda, v_\Upsilon, v_u)^T$ در معادله زیر صدق کند

$$\eta(P) = \left[E_1(t_1, P), \dots, E_1(t_m, P), E_2(t_1, P), \dots, E_2(t_m, P), E_3(t_1, P), \dots, E_3(t_m, P), \right. \\ \left. E_4(t_1, P), \dots, E_4(t_m, P), E_5(t_1, P), \dots, E_5(t_m, P) \right]^T = 0,$$

آن‌گاه P^* یک جواب بهینه از (۱۷.۶) است.

با استفاده از لم ۱.۳.۶، می‌توان به راحتی بررسی کرد که مسئله (۱۷.۶) معادل با مسئله زیر است

$$\text{minimize}_P E(P) = \frac{1}{\Upsilon} \|\eta(P)\|^2. \quad (۲۰.۶)$$

این مسئله بهینه‌سازی را می‌توان با هر روش بهینه‌سازی مسائل بهینه‌سازی نامقید حل کرد.

۴.۶ کنترل بهینه سیستم‌های آشوبناک کسری

آشوب به عنوان یک پدیده بسیار جالب رفتارهای پیچیده دینامیکی دارد، مثلاً به تغییرات کوچک در شرایط اولیه حساس بودن و غیره. در سال‌های اخیر به کنترل آشوب با توجه به برنامه‌های کاربردی بالقوه خود در زمینه ارتباطات امن، سیستم‌های نوری، شیمیایی، فیزیکی، بیولوژیکی، شبکه‌های عصبی و غیره توجه بسیار زیادی شده است [۱۰۸، ۱۰۹]. تاکنون برای کنترل سیستم‌های آشوبناک، طیف وسیعی از روش‌ها پیشنهاد شده است. مانند روش کنترل تطبیقی^۱ [۱۱۰]، کنترل کمینه آنتروپی^۲ [۱۱۱]، روش کنترل مد لغزشی^۳ [۱۱۲] و غیره. در اینجا یک رده جدید از سیستم‌های آشوبناک که سیستم‌های آشوبناک مرتبه کسری نامیده می‌شود را معرفی می‌کنیم. هارتلی در^۴ [۱۱۳] یک سیستم آشوبناک مرتبه کسری را معرفی کرد. در [۱۱۴] شبکه عصبی بافت سلولی^۵ از مرتبه کسری در نظر گرفته شد. همچنین سیستم‌های آشوبناک مرتبه کسری در بسیاری از کارهای دیگر نیز مورد بررسی قرار گرفت. در این بخش قصد داریم یک مسئله کنترل بهینه کسری از سیستم‌های آشوبناک را حول نقاط تعادل آن‌ها مورد بررسی قرار دهیم. یک سیستم دینامیکی مدل‌بندی شده توسط معادلات وضعیت زیر را در نظر بگیرید

$$\begin{cases} {}^C D_t^\alpha x_i(t) = f_i(x_1(t), x_2(t), \dots, x_N(t)), & i = 1, 2, \dots, s, \\ x_i(0) = x_{i,0}, \end{cases} \quad (21.6)$$

که $x = (x_1, x_2, \dots, x_s) : [0, \infty) \rightarrow A \subseteq \mathbb{R}^s$ متغیر وضعیت و f_i یک تابع پیوسته غیرخطی است. از [۱۱۵]، می‌دانیم که برای کنترل آشوب در سیستم‌های آشوبناک (۲۱.۶) برای $i = 1, \dots, s$ باید تابع کنترل u_i تعیین شود به‌طوری‌که

$$\begin{cases} {}^C D_t^\alpha x_i(t) = f_i(x_1(t), x_2(t), \dots, x_N(t)) + u_i(t), & i = 1, 2, \dots, s, \\ \lim_{t \rightarrow \infty} x_i(t) = \bar{x}_i, \end{cases} \quad (22.6)$$

که برای $\bar{x}_i, i = 1, 2, \dots, s$ حالت تعادل است و $u = (u_1, u_2, \dots, u_s) : [0, \infty) \rightarrow B \subseteq \mathbb{R}^s$ ورودی کنترل است که از شرایط سیستم دینامیکی (۲۱.۶) در اطراف نقطه تعادلش به دست آمده است. یک مسئله کنترل بهینه افق نامتناهی کسری^۶ به‌صورت زیر پیشنهاد می‌شود

$$\text{minimize } J = \int_0^\infty \sum_{i=0}^n (x_i(t) - \bar{x}_i)^2 dt, \quad (23.6)$$

¹ Adaptive Control Method

² Minimum Entropy Control

³ Sliding Mode Control Method

⁴ Hartley

⁵ Cellular Neural Network

⁶ Fractional Infinite Horizon Optimal Control

(۲۴.۶)

$$\text{subject to } \begin{cases} {}^C D_t^\alpha x_i(t) = f_i(x_1(t), x_2(t), \dots, x_N(t)) + u_i(t), & i = 1, 2, \dots, s, \\ x_i(0) = x_{i,0}. \end{cases}$$

حال ثابت می‌کنیم که مسئله کنترل بهینه (۲۳.۶) و (۲۴.۶) می‌تواند سیستم آشوبناک کنترل شده (۲۲.۶) را حل کند.

گزاره ۱.۴.۶. برای هر تابع $y : [0, \infty) \rightarrow \mathbb{R}^n$ در $L_2([0, \infty))$ از نرم‌های زیر استفاده می‌کنیم

$$\begin{aligned} \|y(t)\|_2 &= (\sum_{i=1}^n |\eta_i(t)|^2)^{\frac{1}{2}}, \quad t \geq 0, \\ \|y(\cdot)\|_{L_2} &= (\int_0^\infty \|y(t)\|_2^2 dt)^{\frac{1}{2}}, \end{aligned}$$

که n عددی متناهی است و $y(t) = (y_1(t), y_2(t), \dots, y_n(t))$ و برای $i = 1, \dots, n$ ، $y_i : [0, \infty) \rightarrow \mathbb{R}$

تعریف ۱.۴.۶. تابع ورودی $u(\cdot) = \hat{u}(t)$ ، $t \geq 0$ کنترل کننده سیستم آشوبناک (۲۱.۶) به نقطه تعادل ناپایدار $\bar{x}$ است اگر خروجی متناظر $x(\cdot)$ از سیستم (۲۱.۶) در $\lim_{t \rightarrow \infty} x(t) = \bar{x}$ صدق کند، یعنی

$$\lim_{t \rightarrow \infty} \|x(t) - \bar{x}\|_2 = 0.$$

قضیه ۱.۴.۶. [۱۱۶] فرض کنید که یک جواب $u(\cdot) = \hat{u}(t)$ برای کنترل سیستم آشوبناک (۲۱.۶) وجود داشته باشد به طوری که خروجی $x(\cdot) = \hat{x}(t)$ متناظر با آن در $L_2([0, \infty))$ باشد. مسئله کنترل بهینه افق نامتناهی کسری (۲۳.۶) و (۲۴.۶) را در نظر بگیرید. اگر $(x^*(\cdot), u^*(\cdot))$ یک جواب بهینه مسئله کنترل بهینه افق نامتناهی کسری (۲۳.۶) و (۲۴.۶) باشد آن گاه کنترل بهینه $u^*(\cdot)$ ، سیستم آشوبناک (۲۱.۶) را به نقطه تعادل ناپایدار منحصر بفردش می‌رساند.

برهان. چون $u(\cdot) = \hat{u}(t)$ یک جواب برای کنترل سیستم آشوبناک (۲۱.۶) است، خروجی $x(\cdot) = \hat{x}(t)$ متناظرش در $\lim_{t \rightarrow \infty} \hat{x}(t) = \bar{x}$ صدق می‌کند. بنابراین

$$\lim_{t \rightarrow \infty} \|\hat{x}(t) - \bar{x}\|_2^2 = \lim_{t \rightarrow \infty} (\hat{x}(t) - \bar{x})^T (\hat{x}(t) - \bar{x}) = 0. \quad (25.6)$$

چون $\hat{x} \in L_2([0, \infty))$ داریم $J(\hat{x}(\cdot))$ متناهی است. از طرفی $(x^*(\cdot), u^*(\cdot))$ یک جواب بهینه مسئله کنترل بهینه افق نامتناهی کسری (۲۳.۶) و (۲۴.۶) است. بنابراین

$$0 \leq J(x^*(\cdot)) = \|x^*(\cdot) - \bar{x}\|_{L_2}^2 = \int_0^\infty \|x^*(\cdot) - \bar{x}\|_2^2 dt \leq J(\hat{x}(\cdot)). \quad (26.6)$$

در نتیجه $J(x^*(\cdot))$ متناهی است. حال فرض کنید $\bar{x} \neq K \in \mathbb{R}^n$ و $\lim_{t \rightarrow \infty} x^*(t) = K$. پس $\lim_{t \rightarrow \infty} \|x^*(t) - \bar{x}\|_2 = \|K - \bar{x}\|_2$ و با استفاده از (۲۶.۶)، باید یک عدد متناهی باشد. در نتیجه $n_1 \in \mathbb{N}$ و $t_1 > 0$ وجود دارند به طوری که برای هر $t \geq t_1$ خواهیم داشت

$$\|x^*(t) - \bar{x}\|_2 \geq \|K - \bar{x}\|_2 - \frac{1}{n_1} \quad \text{آن‌گاه}$$

$$\begin{aligned} J(x^*(.)) &= \int_0^\infty \|x^*(t) - \bar{x}\|_2^2 dt \\ &= \int_0^{t_1} \|x^*(t) - \bar{x}(t)\|_2^2 dt + \int_{t_1}^\infty \|x^*(t) - \bar{x}\|_2^2 dt \\ &\geq \int_0^{t_1} \|x^*(t) - \bar{x}\|_2^2 dt + (\|K - \bar{x}\|_2 - \frac{1}{n_1}) \int_{t_1}^\infty dt. \end{aligned} \quad (27.6)$$

با توجه به انتگرال آخر در رابطه (۲۷.۶)، $J(x^*(.))$ نمی‌تواند متناهی باشد و این یک تناقض است. بنابراین $\lim_{t \rightarrow \infty} x^*(t) = \bar{x}$.

□

۵.۶ مثال‌های عددی

در این بخش چند مثال عددی و یک کاربرد در سیستم‌های آشوبناک کسری ارائه می‌شود.

مثال ۱.۵.۶. [۷۰] یک مسئله کمینه‌سازی به‌صورت

$$\text{minimize } J(x, u) = \int_0^1 (u^2(t) - 4x(t))^2 dt,$$

با سیستم دینامیکی

$$\dot{x}(t) + {}^C D_t^{0.5} x(t) = u(t) + \frac{2}{\Gamma(2/5)} t^{1/5}, \quad t \in [0, 1],$$

و شرایط مرزی

$$x(0) = 0, \quad x(1) = 1,$$

را در نظر بگیرید. جواب‌های دقیق برای $\alpha = 0.5$ به‌صورت زوج $(t^2, 2t) = (x(t), u(t))$ هستند. با کمک تقریب (۹.۶) مسئله کمینه‌سازی اصلی به مسئله زیر

$$\text{minimize } J(x, u) = \int_0^1 (u^2(t) - 4x(t))^2 dt,$$

با سیستم‌های دینامیکی

$$\begin{cases} \dot{x}(t) = \frac{u(t) + \frac{2}{\Gamma(2/5)} t^{1/5} - A t^{-0.5} x(t) + \sum_{p=2}^k c_p t^{0.5-p} y_p(t)}{1 + B t^{0.5}}, \\ y_p(t) = (1-p)t^{p-2} x(t), \quad p = 2, \dots, k, \end{cases}$$

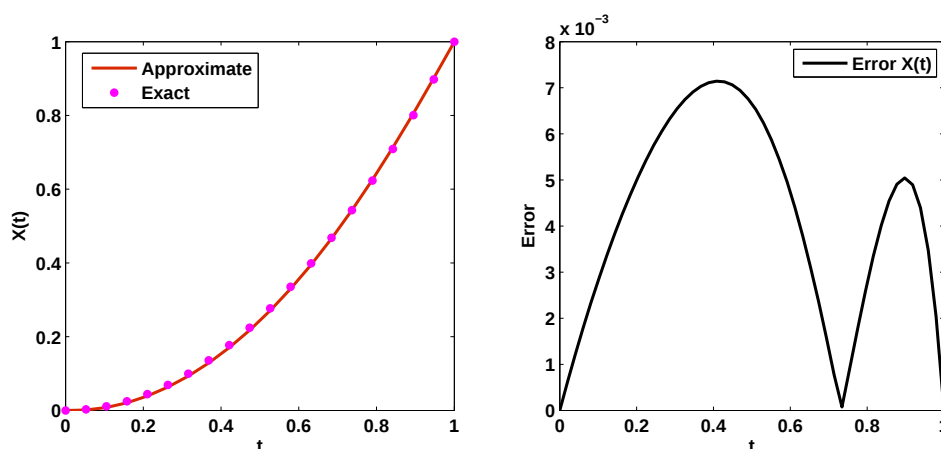
و شرایط مرزی

$$\begin{cases} x(0) = 0, & x(1) = 1, \\ y_p(0) = 0, & p = 2, \dots, k, \end{cases} \quad (28.6)$$

تبدیل می‌شود. متناظر با شرایط اولیه (۲۸.۶) با $k = 3$ ، جواب‌های آزمایشی به‌صورت زیر انتخاب می‌شوند

$$\begin{cases} x_T = t + t(1-t)N_x, \\ \lambda_T = N_\lambda, \\ y_{2_T} = tN_{y_2}, \\ y_{3_T} = tN_{y_3}, \\ \gamma_{2_T} = (1-t)N_{\gamma_2}, \\ \gamma_{3_T} = (1-t)N_{\gamma_3}, \\ u_T = N_u. \end{cases}$$

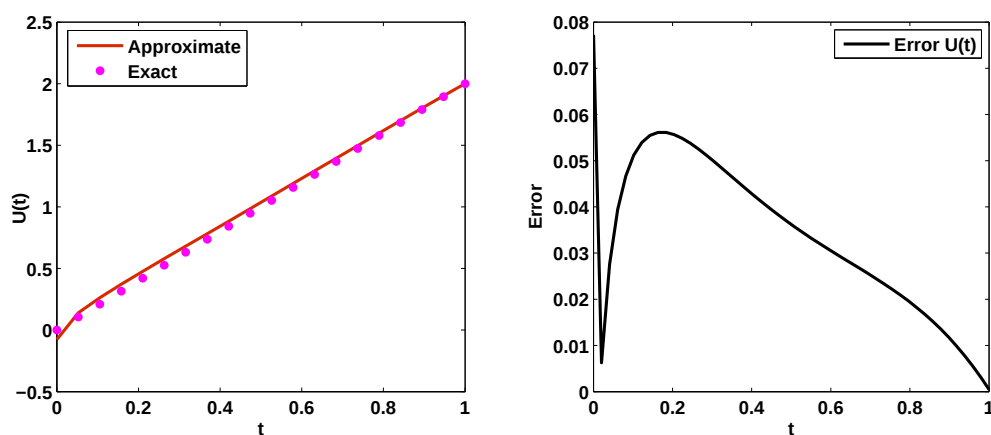
با استفاده از طرح پیشنهاد شده در این فصل مسیر بهینه و کنترل بهینه برای $\alpha = 0.5$ در شکل‌های ۲.۶ و ۳.۶ رسم شده‌اند. در شکل ۴.۶ جواب‌های تقریبی برای چند مقدار α و جواب دقیق رسم شده است. همچنین مقدار تابع خطا $E(y)$ برای چند مقدار $0 < \alpha < 1$ در جدول ۱.۶ نشان داده می‌شود.



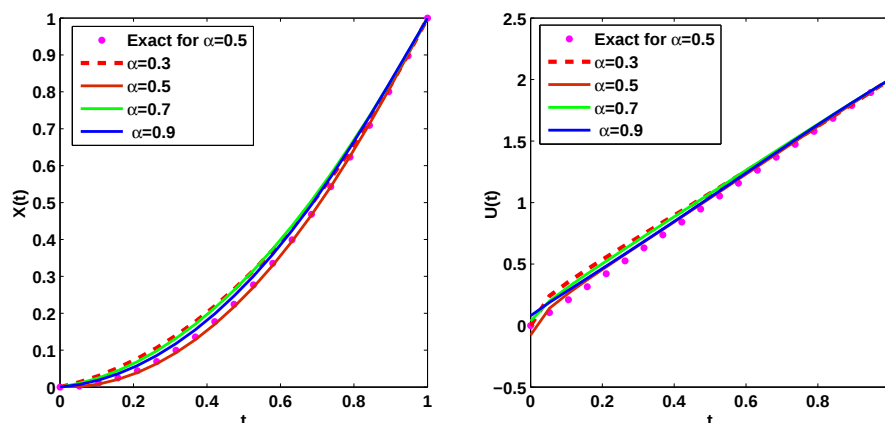
شکل ۲.۶: چپ (مسیر بهینه $x(t)$ برای $\alpha = 0.5$)، راست (خطای مطلق $x(t)$ برای $\alpha = 0.5$).

α	۰/۳	۰/۵	۰/۷	۰/۹
E	$7/33 \times 10^{-3}$	$8/75 \times 10^{-4}$	$1/94 \times 10^{-3}$	$1/56 \times 10^{-3}$

جدول ۱.۶: مقدار تابع خطای $E(y)$ برای چند مقدار $0 < \alpha < 1$.



شکل ۳.۶: چپ (کنترل بهینه $u(t)$ برای $\alpha = 0.5$)، راست (خطای مطلق $u(t)$ برای $\alpha = 0.5$).



شکل ۴.۶: مسیر بهینه $x(\cdot)$ و کنترل بهینه $u(\cdot)$ با مقادیر مختلف α .

مثال ۲.۵.۶ [۷۲] مسئله کنترل بهینه زیر را در نظر بگیرید

$$\text{minimize } J(x, u) = \int_0^1 (tu(t) - (\alpha + 2)x(t))^2 dt,$$

که سیستم دینامیکی و شرایط مرزی به صورت زیر هستند

$$\begin{cases} \dot{x}(t) + {}^C D_t^\alpha x(t) = u(t) + t^2, \\ x(0) = 0, \quad x(1) = \frac{2}{\Gamma(3 + \alpha)}. \end{cases}$$

جواب دقیق به صورت

$$(x(t), u(t)) = \left(\frac{t^{2(\alpha+2)}}{\Gamma(\alpha+3)}, \frac{t^{\alpha+1}}{\Gamma(\alpha+2)} \right),$$

است. با استفاده از (۹.۶) مسئله کمینه‌سازی اصلی به

$$\text{minimize } J(x, u) = \int_0^1 (tu(t) - (\alpha + 2)x(t))^2 dt,$$

با سیستم دینامیکی

$$\begin{cases} \dot{x}(t) = \frac{u(t) + t^\alpha - At^{-\alpha}x(t) + \sum_{p=2}^k c_p t^{1-p-\alpha} y_p(t)}{1 + Bt^{1-\alpha}}, \\ y_p(t) = (1-p)t^{p-2}x(t), \quad p = 2, \dots, k, \end{cases}$$

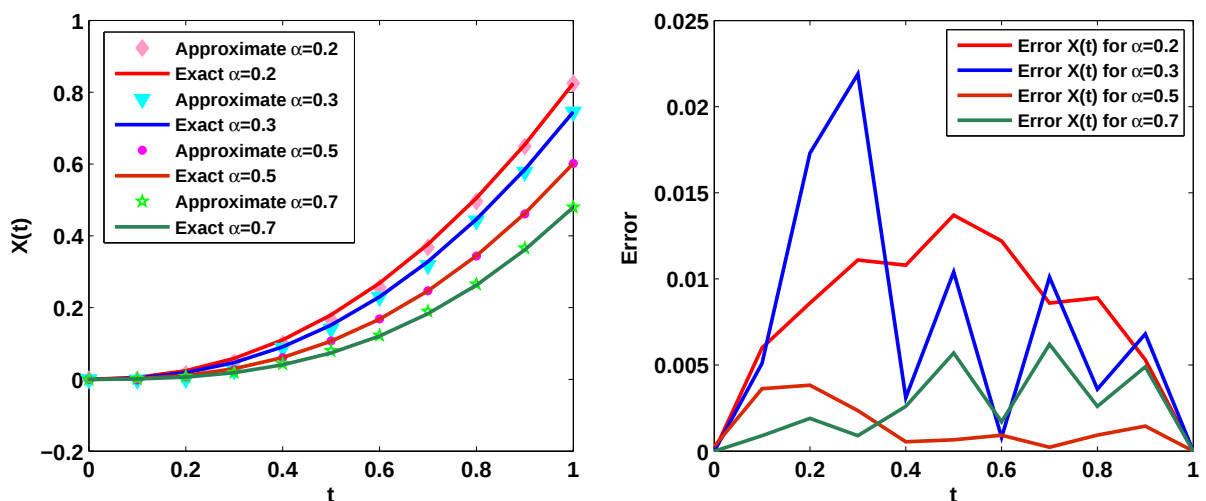
و شرایط مرزی

$$\begin{cases} x(0) = 0, & x(1) = \frac{2}{\Gamma(3+\alpha)}, \\ y_p(0) = 0, & p = 2, \dots, k, \end{cases} \quad (29.6)$$

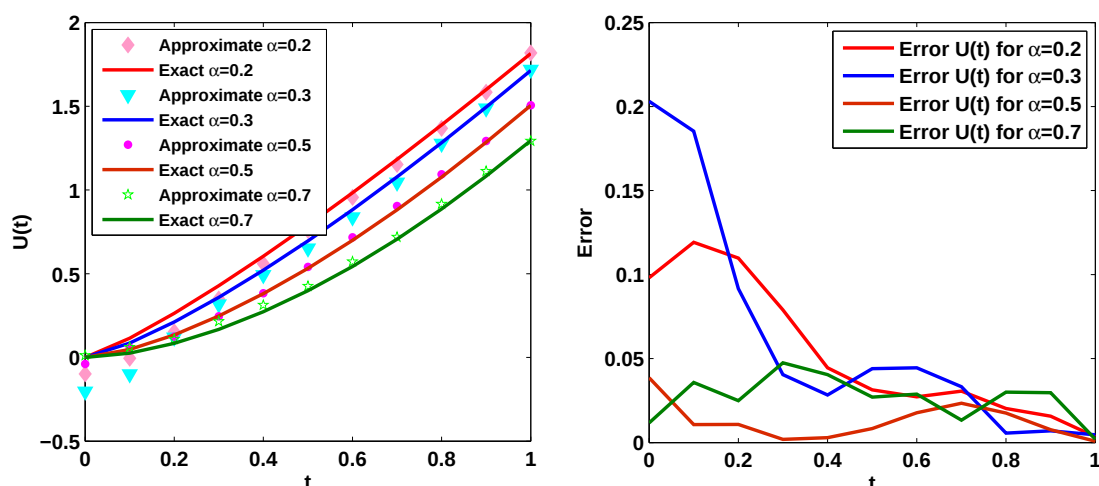
تبدیل می‌شود. مطابق شرایط اولیه (29.6) با $k = 3$ جواب‌های آزمایشی را می‌توان به صورت زیر انتخاب کرد

$$\begin{cases} x_T = \frac{2}{\Gamma(3+\alpha)}t + t(1-t)N_x, & \lambda_T = N_\lambda, \\ y_{2T} = tN_{y_2}, & y_{3T} = tN_{y_3}, \\ \gamma_{2T} = (1-t)N_{\gamma_2}, & \gamma_{3T} = (1-t)N_{\gamma_3}, \\ u_T = N_u. \end{cases}$$

با استفاده از روش ارائه شده، مسیر بهینه، کنترل بهینه و خطای مطلق برای $\alpha = 0.2, 0.3, 0.5, 0.7$ به ترتیب در شکل‌های 5.6 و 6.6 رسم شده‌اند.



شکل 5.6: چپ (مسیر بهینه $x(t)$ برای $\alpha = 0.2, 0.3, 0.5, 0.7$)، راست (خطای مطلق $x(t)$ برای $\alpha = 0.2, 0.3, 0.5, 0.7$).



شکل ۶.۶: چپ (کنترل بهینه $u(t)$ برای $\alpha = 0.2, 0.3, 0.5, 0.7$)، راست (خطای مطلق $u(t)$ برای $\alpha = 0.2, 0.3, 0.5, 0.7$).

مثال ۳.۵.۶. سیستم آشوبناک کسری زیر را در نظر بگیرید

$$\begin{cases} {}^C D_t^\alpha X_1(t) = -X_1(t) - X_2(t), \\ {}^C D_t^\alpha X_2(t) = 2/5 X_2(t) - 4 X_1(t) X_3(t), \\ {}^C D_t^\alpha X_3(t) = -5 X_3(t) + 4 X_1(t) X_2(t), \\ X_1(0) = 0.2, \quad X_2(0) = 0, \quad X_3(0) = 0.5. \end{cases} \quad (30.6)$$

به منظور کنترل آشوب در سیستم آشوبناک (۳۰.۶) باید کنترل‌های $U_1(t)$ ، $U_2(t)$ و $U_3(t)$ را طوری تعیین کرد که

$$\lim_{t \rightarrow \infty} X_i(t) = \bar{X}_i \quad i = 1, 2, 3. \quad (31.6)$$

این معادل با حل مسئله بهینه‌سازی زیر است

$$\begin{aligned} & \text{minimize} \quad \int_0^\infty [(X_1(t) - \bar{X}_1)^2 + (X_2(t) - \bar{X}_2)^2 + (X_3(t) - \bar{X}_3)^2] dt, \\ & \text{subject to} \quad \begin{cases} {}^C D_t^\alpha X_1(t) = -X_1(t) - X_2(t) + U_1(t), \\ {}^C D_t^\alpha X_2(t) = 2/5 X_2(t) - 4 X_1(t) X_3(t) + U_2(t), \\ {}^C D_t^\alpha X_3(t) = -5 X_3(t) + 4 X_1(t) X_2(t) + U_3(t), \\ X_1(0) = 0.2, \quad X_2(0) = 0, \quad X_3(0) = 0.5. \end{cases} \end{aligned} \quad (32.6)$$

با استفاده از (۵.۶)، تغییر متغیر $t = \frac{\tau}{(1-\tau)}$ و

$$\begin{cases} x_i(\tau) = X_i\left(\frac{\tau}{1-\tau}\right), & i = 1, 2, 3, \\ u_i(\tau) = U_i\left(\frac{\tau}{1-\tau}\right), & i = 1, 2, 3, \\ v_i(\tau) = V_i\left(\frac{\tau}{1-\tau}\right), & i = 1, 2, 3. \end{cases}$$

مسئله (۳۲.۶) را برای $(p = 2, 3)$ می‌توان به صورت زیر بازنویسی کرد
(۳۳.۶)

$$\begin{aligned} \text{minimize} \quad & \int_0^1 \frac{1}{(1-\tau)^2} [(x_1(\tau) - \bar{X}_1)^2 + (x_2(\tau) - \bar{X}_2)^2 + (x_3(\tau) - \bar{X}_3)^2] d\tau, \\ \text{subject to} \quad & \begin{cases} \dot{x}_1 = \frac{1}{(1-\tau)^2 B(\frac{\tau}{1-\tau})^{1-\alpha}} \left(C_2 \left(\frac{\tau}{1-\tau}\right)^{-1-\alpha} v_{21}(\tau) + C_3 \left(\frac{\tau}{1-\tau}\right)^{-2-\alpha} v_{31}(\tau) \right. \\ \quad \left. - (A \left(\frac{\tau}{1-\tau}\right)^{-\alpha} + 1) x_1(\tau) + x_2^2(\tau) + u_1(\tau) \right), \\ \dot{x}_2 = \frac{1}{(1-\tau)^2 B(\frac{\tau}{1-\tau})^{1-\alpha}} \left(C_2 \left(\frac{\tau}{1-\tau}\right)^{-1-\alpha} v_{21}(\tau) + C_3 \left(\frac{\tau}{1-\tau}\right)^{-2-\alpha} v_{31}(\tau) \right. \\ \quad \left. + (A \left(\frac{\tau}{1-\tau} - 2/5\right)^{-\alpha}) x_2(\tau) - 4 x_1(\tau) - x_1(\tau) x_3(\tau) + u_2(\tau) \right), \\ \dot{x}_3 = \frac{1}{(1-\tau)^2 B(\frac{\tau}{1-\tau})^{1-\alpha}} \left(C_2 \left(\frac{\tau}{1-\tau}\right)^{-1-\alpha} v_{21}(\tau) + C_3 \left(\frac{\tau}{1-\tau}\right)^{-2-\alpha} v_{31}(\tau) \right. \\ \quad \left. - (A \left(\frac{\tau}{1-\tau}\right)^{-\alpha} + 5) x_3 + 4 x_1(\tau) x_2(\tau) + u_3(\tau) \right), \\ \dot{v}_{21}(t) = -\frac{1}{(1-\tau)^2} x_1(\tau), \quad \dot{v}_{22}(t) = -\frac{1}{(1-\tau)^2} x_2(\tau), \\ \dot{v}_{23}(t) = -\frac{1}{(1-\tau)^2} x_3(\tau), \quad \dot{v}_{31}(t) = -\frac{2\tau}{(1-\tau)^2} x_1(\tau), \\ \dot{v}_{32}(t) = -\frac{2\tau}{(1-\tau)^2} x_1(\tau), \quad \dot{v}_{33}(t) = -\frac{2\tau}{(1-\tau)^2} x_1(\tau), \\ x_1(0) = 0/2, \quad x_2(0) = 0, \quad x_3(0) = 0/5, \\ v_{21}(0) = 0, \quad v_{22}(0) = 0, \quad v_{23}(0) = 0, \\ v_{31}(0) = 0, \quad v_{32}(0) = 0, \quad v_{33}(0) = 0. \end{cases} \end{aligned}$$

فرض کنید $h = \left(\frac{1}{1-\tau}\right)^2$ و $K = \frac{1}{(1-\tau)^2 B(\frac{\tau}{1-\tau})^{1-\alpha}}$. تابع همیلتونی (۳۳.۶) عبارت است از

$$\begin{aligned} H = & h \left[(x_1(\tau) - \bar{X}_1)^2 + (x_2(\tau) - \bar{X}_2)^2 + (x_3(\tau) - \bar{X}_3)^2 \right] + \lambda_1 K \\ & \left[C_2 \left(\frac{\tau}{1-\tau}\right)^{-1-\alpha} v_{21}(\tau) + C_3 \left(\frac{\tau}{1-\tau}\right)^{-2-\alpha} v_{31}(\tau) - (A \left(\frac{\tau}{1-\tau}\right)^{-\alpha} + 1) \right. \\ & \left. x_1(\tau) + x_2^2(\tau) + u_1(\tau) \right] + \lambda_2 K \left[C_2 \left(\frac{\tau}{1-\tau}\right)^{-1-\alpha} v_{21}(\tau) + C_3 \right. \\ & \left(\frac{\tau}{1-\tau} \right)^{-2-\alpha} v_{31}(\tau) + (A \left(\frac{\tau}{1-\tau}\right)^{-\alpha} - 2/5) x_2(\tau) - 4 x_1(\tau) x_3(\tau) \\ & \left. + u_2(\tau) \right] + \lambda_3 K \left[C_2 \left(\frac{\tau}{1-\tau}\right)^{-1-\alpha} v_{21}(\tau) + C_3 \left(\frac{\tau}{1-\tau}\right)^{-2-\alpha} v_{31}(\tau) \right. \\ & \left. - (A \left(\frac{\tau}{1-\tau}\right)^{-\alpha} + 5) x_3 + 4 x_1(\tau) x_2(\tau) + u_3(\tau) \right]. \end{aligned} \quad (34.6)$$

با توجه به تابع همیلتونی شرایط بهینگی را می‌توان به صورت زیر نوشت
(۳۵.۶)

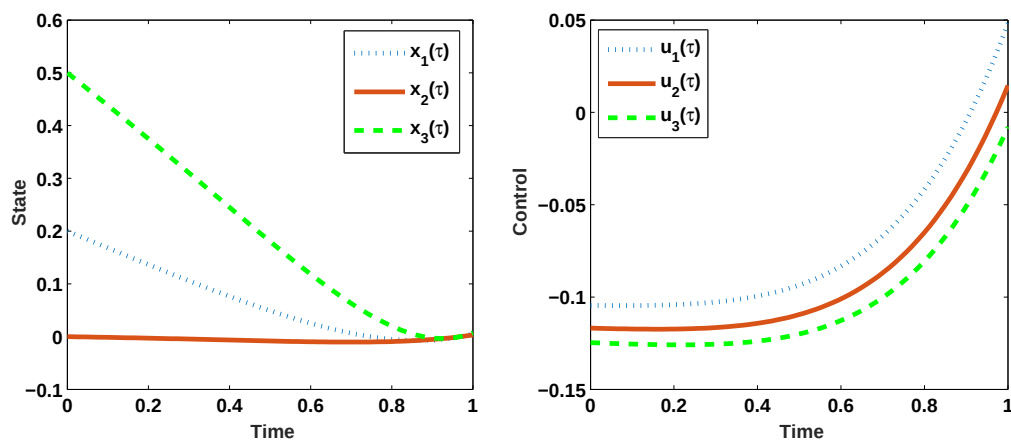
$$\left\{ \begin{array}{l} \Upsilon h(x_1(\tau) - \bar{x}_1)^2 - \lambda_1 K(A(\frac{\tau}{1-\tau})^{-\alpha} + 1) - \lambda_2 Kx_2(\tau) + \Upsilon \lambda_3 Kx_2(\tau) + \lambda_3 Kx_2(\tau) - h\Upsilon_1 \\ - \frac{\Upsilon\tau}{1-\tau} \Upsilon_\epsilon = -\dot{\lambda}_1, \\ \Upsilon h(x_2(\tau) - \bar{x}_2)^2 + \Upsilon \lambda_1 Kx_2(\tau) - \lambda_2 K(A(\frac{\tau}{1-\tau})^{-\alpha} - \Upsilon/\Delta)^{-\alpha} + \Upsilon \lambda_3 Kx_1(\tau) - h\Upsilon_2 - \frac{\Upsilon\tau}{1-\tau} \Upsilon_\Delta \\ = -\dot{\lambda}_2, \\ \Upsilon h(x_3(\tau) - \bar{x}_3)^2 - \Upsilon \lambda_2 Kx_1(\tau) - \lambda_3 K(A(\frac{\tau}{1-\tau})^{-\alpha} + \Delta) - h\Upsilon_3 - \frac{\Upsilon\tau}{1-\tau} \Upsilon_\epsilon = -\dot{\lambda}_3, \\ h[C_2(\frac{\tau}{1-\tau})^{-1-\alpha} v_{\Upsilon 1}(\tau) + C_3(\frac{\tau}{1-\tau})^{-2-\alpha} v_{\Upsilon 1}(\tau) - (A(\frac{\tau}{1-\tau})^{-\alpha} + 1)x_1(\tau) - x_2^2(\tau) + u_1(\tau)] \\ = \dot{x}_1, \\ h[C_2(\frac{\tau}{1-\tau})^{-1-\alpha} v_{\Upsilon 1}(\tau) + C_3(\frac{\tau}{1-\tau})^{-2-\alpha} v_{\Upsilon 1}(\tau) + (A(\frac{\tau}{1-\tau})^{-\alpha} - \Upsilon/\Delta)x_2(\tau) - \Upsilon x_1(\tau)x_3(\tau) \\ + u_2(\tau)] = \dot{x}_2, \\ h[C_2(\frac{\tau}{1-\tau})^{-1-\alpha} v_{\Upsilon 1}(\tau) + C_3(\frac{\tau}{1-\tau})^{-2-\alpha} v_{\Upsilon 1}(\tau) - (A(\frac{\tau}{1-\tau})^{-\alpha} + \Delta)x_3 + \Upsilon x_1(\tau)x_2(\tau) + u_3(\tau)] \\ = \dot{x}_3, \\ \lambda_1 KC_2(\frac{\tau}{1-\tau})^{-1-\alpha} = -\dot{\Upsilon}_1, \quad \lambda_2 KC_2(\frac{\tau}{1-\tau})^{-1-\alpha} = -\dot{\Upsilon}_2, \quad \lambda_3 KC_2(\frac{\tau}{1-\tau})^{-1-\alpha} = -\dot{\Upsilon}_3, \\ \lambda_1 KC_3(\frac{\tau}{1-\tau})^{-1-\alpha} = -\dot{\Upsilon}_\epsilon, \quad \lambda_2 KC_3(\frac{\tau}{1-\tau})^{-1-\alpha} = -\dot{\Upsilon}_\Delta, \quad \lambda_3 KC_3(\frac{\tau}{1-\tau})^{-1-\alpha} = -\dot{\Upsilon}_\epsilon, \\ -hx_1 = \dot{v}_{\Upsilon 1}, \quad -hx_1 = \dot{v}_{\Upsilon 2}, \quad -hx_1 = \dot{v}_{\Upsilon 3}, \\ \frac{-\Upsilon\tau}{1-\tau} \Upsilon^3 = \dot{v}_{\Upsilon 1}, \quad \frac{-\Upsilon\tau}{1-\tau} \Upsilon^3 = \dot{v}_{\Upsilon 2}, \quad \frac{-\Upsilon\tau}{1-\tau} \Upsilon^3 = \dot{v}_{\Upsilon 3}, \\ \frac{\lambda_1}{(1-\tau)^\Upsilon} = 0, \quad \frac{\lambda_2}{(1-\tau)^\Upsilon} = 0, \quad \frac{\lambda_3}{(1-\tau)^\Upsilon} = 0. \end{array} \right.$$

با در نظر گرفتن شرایط اولیه، می‌توان جواب‌های آزمایشی را به صورت زیر انتخاب کرد
(۳۶.۶)

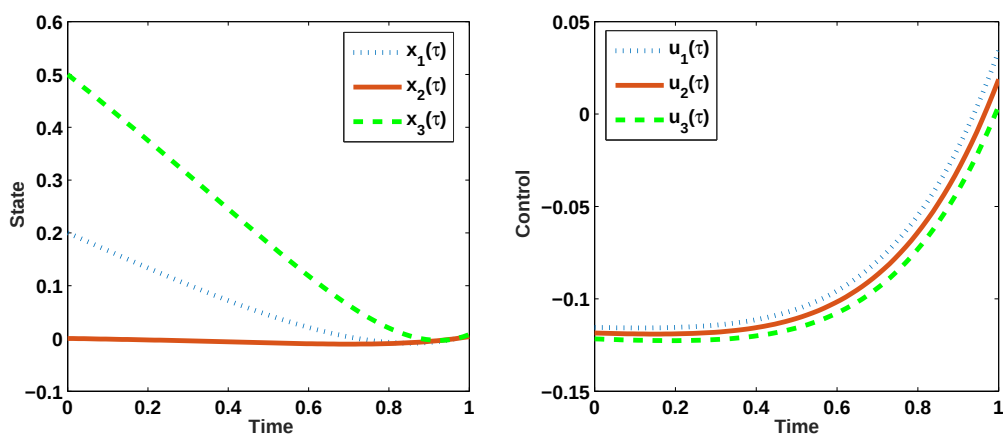
$$\left\{ \begin{array}{lll} x_{1T} = 0/\Upsilon + \tau N_{x_1} & x_{2T} = \tau N_{x_2} & x_{3T} = 0/\Delta + \tau N_{x_3} \\ v_{\Upsilon 1T} = \tau N_{v_{\Upsilon 1}}, & v_{\Upsilon 2T} = \tau N_{v_{\Upsilon 2}}, & v_{\Upsilon 3T} = \tau N_{v_{\Upsilon 3}}, \\ v_{\Upsilon 1T} = \tau N_{v_{\Upsilon 1}}, & v_{\Upsilon 2T} = \tau N_{v_{\Upsilon 2}}, & v_{\Upsilon 3T} = \tau N_{v_{\Upsilon 3}}, \\ \lambda_{1T} = (1-\tau)N_{\lambda_1}, & \lambda_{2T} = (1-\tau)N_{\lambda_2}, & \lambda_{3T} = (1-\tau)N_{\lambda_3}, \\ \Upsilon_{1T} = (1-\tau)N_{\Upsilon_1}, & \Upsilon_{2T} = (1-\tau)N_{\Upsilon_2}, & \Upsilon_{3T} = (1-\tau)N_{\Upsilon_3}, \\ u_{1T} = N_{u_1}, & u_{2T} = N_{u_2}, & u_{3T} = N_{u_3}. \end{array} \right.$$

جواب‌های تقریبی برای توابع وضعیت و کنترل برای $\alpha = 0/5, 0/8, 0/9$ به ترتیب در شکل‌های

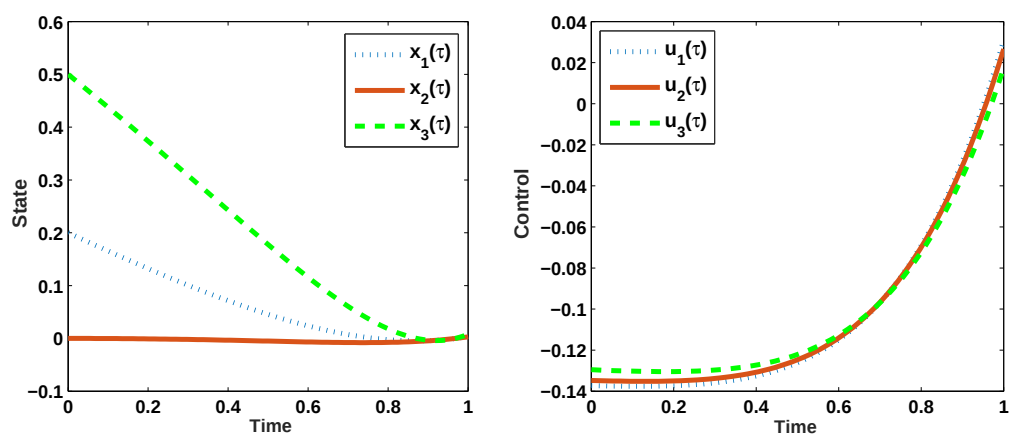
۷.۶-۹.۶ رسم شده‌اند.



شکل ۷.۶: رفتار پایدار تابع کنترل و وضعیت برای $\alpha = \pi/9$.



شکل ۸.۶: رفتار پایدار تابع کنترل و وضعیت برای $\alpha = \pi/8$.



شکل ۹.۶: رفتار پایدار تابع کنترل و وضعیت برای $\alpha = 0.05$.

نتیجه‌گیری و پیشنهادات

در این پایان‌نامه رده‌ای از مسائل کنترل بهینه کسری بر پایه شبکه عصبی با چهار مدل مختلف را مورد بررسی قرار دادیم. در فصل ۴ برای تقریب مشتق کسری، از تقریب مشتق گراناوالد-لتنیکف استفاده شد و ساختار شبکه عصبی پرسپترون روی آن پیاده‌سازی شد. در فصل ۵ از دو روش برای حل مسائل کنترل بهینه کسری با محدودیت‌های مساوی و نامساوی استفاده شد. روش اول بر پایه شبکه عصبی سری توانی بود. با توجه به اینکه مشتق این سری به راحتی قابل محاسبه است پیچیدگی محاسباتی به‌طور قابل توجهی کاهش یافت. در روش دوم با جانشینی تابع نمایی به جای تابع سیگموئید در ساختار شبکه عصبی، مشتق کسری از تابع نمایی به یک تابع میتاگ-لفلر تبدیل شد که این تابع در حسابان کسری اهمیت و کاربرد فراوان دارد. در فصل ۶ مسئله کنترل بهینه کسری ابتدا با یک تقریب به یک مسئله کنترل بهینه غیرکسری تبدیل کردیم و سپس شبکه عصبی لژاندر روی آن پیاده‌سازی شد.

در یک مسئله کنترل بهینه کسری، مشتق کسری می‌تواند هم از مرتبه ثابت و هم از مرتبه متغیر باشد. روش‌های ارائه شده در این پایان‌نامه ابتدا روی مسائلی با مشتق کسری از مرتبه ثابت پیاده‌سازی شد. اما با توجه به اینکه اخیراً یکی از بحث‌های جالب و مورد توجه محققان حل مسئله کنترل بهینه با مشتق مرتبه متغیر است این مسئله را نیز در پایان‌نامه معرفی کردیم و سپس با دو روش ارائه شده در فصل‌های ۴ و ۵ با همان ساختار شبکه‌های معرفی شده مسئله را حل کردیم. تمام روش‌هایی که در این پایان‌نامه استفاده شد روش‌های غیرمستقیم هستند و در تمام این روش‌ها در نهایت به یک تابع خطا رسیدیم. با انتخاب این تابع خطا به‌عنوان یک تابع لیاپانف در ساختار شبکه عصبی، به‌طور کامل همگرایی و پایداری روش‌ها در فصل ۴ بررسی شد.

اکنون یک سوال منطقی پیش می‌آید. مزایای استفاده از شبکه‌های عصبی در مقایسه با دیگر روش‌های موجود برای حل مسائل کنترل بهینه چیست؟ برای پاسخ به این سوال بعضی از مزیت‌های روش شبکه‌های عصبی پیشنهادی به‌طور خلاصه بیان می‌شود.

(۱) از شبکه‌های عصبی برای حل مسائلی که جواب‌های تحلیلی ندارند یا جواب‌های تحلیلی آن‌ها به راحتی قابل محاسبه نیست استفاده می‌شود.

(۲) جواب مسئله حول تمام دامنه مسئله پیوسته است. در حالی که اکثر روش‌های عددی جواب‌ها را تنها در نقاط گسسته‌سازی شده به‌دست می‌آورند و جواب بین این نقاط باید از طریق درونیابی به‌دست آید.

(۳) روش‌ها بر پایه شبکه عصبی از ماتریس عملگر استفاده نمی‌کنند. برخی از روش‌های عددی از ماتریس عملگر برای حل استفاده می‌کنند. پیدا کردن ماتریس عملگر معمولاً سخت است و این ماتریس‌های عملگر در مرتبه‌های بالاتر باعث پیچیدگی محاسباتی نیز خواهند شد.

(۴) جواب تقریبی مسئله نزدیک به جواب تحلیلی است.

(۵) به راحتی می‌توان از روش شبکه عصبی برای حل مسائل غیرخطی یا مسائلی با مرتبه‌های

بالا تر استفاده کرد.

۶) جواب مسئله را می‌توان به طور سریع و با تعداد کمی از پارامترها در کاربردهای واقعی تخمین زد.

مشتق کاپاتو و همدیس‌پذیر این مشکل را دارند که مشتق دو تابع در صفر، صفر است که این اشکال بزرگی است. برای رهایی از این مشکل پیشنهاد می‌شود که در ادامه از مشتق کاپاتو-فابریسیو استفاده شود. برخی از مسائل کنترل بهینه ارائه شده در این پایان‌نامه و حتی مسئله کنترل بهینه با مشتق کاپاتو-فابریسیو را بر پایه شبکه عصبی به صورت مستقیم حل کردیم و نتایج بسیار خوبی به دست آمد. برای کارهای آتی پیشنهاد می‌شود که این مسائل را بر پایه شبکه عصبی به صورت مستقیم نیز حل شود و پایداری و همگرایی روش بررسی شود. همچنین می‌توان برای کارهای آتی پیشنهاد کرد که این روش‌ها را روی مسائلی با سیستم‌های تاخیری، مسائلی با مشتقات جزئی کسری، مسائل افق نامتناهی و ... پیاده‌سازی کرد.

مراجع

- [۱] پتکین ف، (۱۳۸۷) ”شبکه‌های عصبی“ ترجمه مبرصالحی م.م. تقی‌زاده کاخکی ح، دانشگاه فردوسی مشهد.
- [۲] منهاج م.ب، (۱۳۷۷) ”هوش محاسباتی“، جلد اول، مرکز نشر پرفسور حسابی.
- [3] Magin R.L. (2006), ”**Fractional Calculus in Bioengineering**”, Begell House Publishers.
- [4] Hilfer R. (2000), ”**Applications of Fractional Calculus in Physics**”, Word Scientific, Singapore.
- [5] Torvik P. J. and Bagley R. L. (1984), ”On the appearance of the fractional derivative in the behavior of real materials”, **Journal of Applied Mechanics**, 51, 2, 294-298.
- [6] Chen S., Liu F. and Burrage K. (2014), ”Numerical simulation of a new two-dimensional variable-order fractional percolation equation in non-homogeneous porous media”, **Elsevier Computers & Mathematics with Applications**, 68, 2133-2141.
- [7] Odziejewicz T. , Malinowska A. B. and Torres D. F. M. (2013), ”Fractional variational calculus of variable order”, **Advances in harmonic analysis and operator theory**, 229, 291–301.
- [8] Chechkin A. V. , Gorenflo R. and Sokolov I. M. (2005), ”Fractional diffusion in inhomogeneous media”, **Journal of Physics A: Mathematical and General** , 38 ,42, L679-L684.
- [9] Santamaria F. , Wils S. , de Schutter E. and Augustine G. J. (2006), ”Anomalous diffusion in Purkinje cell dendrites caused by spines”, **Neuron**. 52, 635-648.
- [10] Sun H. G. , Chen W. and Chen Y. Q. (2009), ”Variable order fractional differential operators in anomalous diffusion modeling”, **Physica A**. 388, 4586-4592.

-
- [11] Lorenzo C.F. and Hartley T.T. (1998), "Initialization, conceptualization and application in the generalized fractional calculus", **NASA/TP-1998-208-208415**.
 - [12] Ingman D., Suzdalnitsky J. and Zeifman M. (2000), "Constitutive dynamic-order model for nonlinear contact phenomena", **ASME Journal of Applied Mechanics**, 67, 383-390.
 - [13] Lorenzo C.F. and Hartley T.T. (2002), "Variable order and distributed order fractional operators", **Nonlinear Dynamics**, 29, 57-98.
 - [14] Coimbra C.F.M. (2003), "Mechanics with variable-order differential operators", **Annalen der Physik**, 12, 692-703.
 - [15] Sun H.G., Chen W., Wei H. and Chen Y.Q. (2011), "A comparative study of constant-order and variable-order fractional models in characterizing memory property of systems", **The European Physical Journal Special Topics**, 193, 185-192.
 - [16] Sierociuk D., Malesza W. and Macias M. (2013), "Switching scheme, equivalence, and analog validation of the alternative fractional variable-order derivative definition", 52nd IEEE Conference on Decision and Control December 10-13, Florence, Italy.
 - [17] Podlubny I. (1999), "**Fractional differential equations**", Mathematics in Science and Engineering, Academic Press Inc, San Diego, CA, 198.
 - [18] Mittag-Leffler G. M. (1904), "Sopra la funzione $E_\alpha(x)$ ", **Rend Accad Lincei**, ser. , 5 (13), 3-5.
 - [19] Mittag-Leffler G. M. (1905), "Sur la representation analytique d'une branche uniforme d'une fonction monogene", **Acta Mathematica**, 29, 101-182.
 - [20] Kilbas, A. A. and H. M. Srivastava, and J. J. Trujillo (2006), "**Theory and Applications of Fractional Differential Equations**". North-Holland Mathematics Studies, Elsevier Science B.V., Amsterdam.
 - [21] Diethelm K. (2010), "The Analysis of Fractional Differential Equations", **Springer**.
 - [22] Jumarie G. (2008), " Modeling fractional stochastic systems as non-random fractional dynamics driven by Brownian motions", **Applied Mathematical Modelling**, 32, 836-859.

- [23] Baleanu D., Defterli O. and Agrawal O. P. (2009), " A Central Difference Numerical Scheme for Fractional Optimal Control Problems", **Jurnal of Vibration and Control**, 15, 4, 583-597.
- [24] Jumarie G. (2006), "Modified Riemann-Liouville Derivative and Fractional Taylor Series of Nondifferentiable Functions Further Results", **Computers and Mathematics with Applications**, 51, 1367-1376.
- [25] Jumarie G. (2007), "Lagrangian mechanics of fractional order, Hamilton- Jacobi fractional PDE and Taylors series of non differentiable functions", **Chaos Solitons Fractals**, 32(3), 969-987.
- [26] Jumarie G. (2009), "Table of some basic fractional calculus formulae derived from a modified Riemann - Liouville derivative for non-differentiable funtions", **Applied Mathematics Letters**, 22, 378-385.
- [27] Caputo M. and Fabrizio M. (2015), " A New Definition of Fractional Derivative without Singular Kernel", **Progress in Fractional Differentiation and Applications**, 1-13.
- [28] Khalil R., Al Horani M. , Yousef A. and Sababheh M.(2014) , "A new definition of fractional derivative", **Journal of Computational and Applied Mathematics**, 264, 65-70.
- [29] Benkhettou N. , Hassani S. and Torres F. M.D.(2016) , "A conformable fractional calculus on arbitrary time scales", **Journal of king saud university-science**, 28, 93-98.
- [30] Tavares D., Almeida R. and Torres D. F. M. (2016), " Caputo derivatives of fractional variable order: numerical approximations", **Communications in Nonlinear Science and Numerical Simulation**, 35, 69-87.
- [31] Ejlali N. and Hosseini S. M. (2017), "A Pseudospectral Method for Fractional Optimal Control Problems", **Journal of Optimization Theory and Applications**, 174, 83-107.
- [32] Zayernouri M. , Karniadakis G.E. (2014), " Fractional spectral collocation method", **SIAM**, 36, A40-A62.

- [33] Badalyan G.V. (1956), " Generalization of Legendre polynomials and some of their applications", **Russian, Armenian summary**, 9, 3–22.
- [34] Rida S.Z. and Yousef A.M. (2011), " On the fractional order Rodrigues formula for the Legendre polynomials", **Advances in Applied Mathematics**, 10, 509–518.
- [35] Kazem S. , Abbasbandy S. and Kumar S. (2013), " Fractional-order Legendre functions for solving fractional-order differential equations", **Applied Mathematical Modelling**, 37, 5498–5510.
- [36] Khader M. M. and Hendy A. S. (2012), "An Efficient Numerical Scheme for Solving Fractional Optimal Control Problems", **International Journal of Nonlinear Science**, 14 (3), 287-296.
- [37] Hosseinpour S. , Nazemi A. and Yavari M. (2016), "A numerical method based on fractional-order Chebyshev polynomials for solving delay fractional optimal control problems", **9th International Iranian Operation Research conference, Shiraz , Iran**.
- [38] Taslakyan A. K. (1984), " Some properties of Legendre quasipolynomials with respect to a Müntz system", **Russian, Armenian summary, Érevan University**, 2, 179–189.
- [39] Borwein P. , Erdélyi T. and Zhang J. (1994), " Müntz systems and orthogonal Müntz–Legendre polynomials", **Transactions of the American Mathematical Society**, 342, 523–542.
- [40] Esmaeili Sh. , Shamsi M. and Luchko Y. (2011), " Numerical solution of fractional differential equations with a collocation method based on Müntz polynomials", **Computers and Mathematics with Applications**, 62, 918–929.
- [41] Hebb, D. O. (1949), " **The Organization of Behavior: A Neuropsychological Theory**", John Wiley, Sons, Inc., New York.
- [42] Rosenblatt F. (1962), " **Principles of neurodynamics: perceptrons and the theory of brain mechanisms**", Report (Cornell Aeronautical Laboratory). Spartan Books.

- [43] Widrow B. and Lehr M. A. (1990), "30 years of adaptive neural networks: perceptron, madaline, and backpropagation", **Proceedings of the IEEE**, 78(9), 1415-1442.
- [44] Widrow B. and Winter R. (1988), "Neural nets for adaptive filtering and adaptive pattern recognition", **IEEE Computer Magazine**, 21(3), 25-39.
- [45] Hopfield J. J. (1982), "Neural networks and physical systems with emergent collective computational abilities", **Proceedings of the National Academy of Science**, 79, 8, 2554-2558.
- [46] Hopfield J. J. and Tank D. W. (1985), "neural computation of decisions in optimization problems", **Biological Cybernetics**, 52(3), 141-152.
- [47] Hornik K., Stinchcombe M. and White H. (1989), "Multilayer feedforward networks are universal approximators", **Neural Networks**, 2, 5, 359-366.
- [48] Perko, L. (2001), "**Differential Equations and Dynamical Systems**", Springer, 3rd ed..
- [49] Ortega, J. M. and Rheinboldt, W. C. (2000), "**Iterative Solution of Nonlinear Equations in Several Variables**", Classics in Applied Mathematics. Society for Industrial and Applied Mathematics.
- [50] Zabczyk, J. (1992), "**Mathematical control theory: an introduction**" Systems , control. Birkhäuser.
- [51] Khalil H. K. (2002), "**Nonlinear Systems**", Prentice Hall PTR, 3rd ed..
- [52] Pinch E.R. (1995), "**Optimal control and the calculus of variations**", Oxford University Press.
- [53] Kirk D.E. (1970), "**Optimal Control Theory**",.: An Introduction, Englewood Cliffs, New York,
- [54] Agrawal O. P. (2008), "A formulation and numerical scheme for fractional optimal control problems", **Journal of Vibration and Control**, 14, 1291-1299.
- [55] Guo, T. L. (2013), "The necessary conditions of fractional optimal control in the sense of Caputo", **Journal of Optimization Theory and Applications** , 156, 115-126.

- [56] Agrawal O. M. P. (2004), "A general formulation and solution scheme for fractional optimal control problem", **Nonlinear Dynamics** , 38, 323-337.
- [57] Riewe F. (1996), "Nonconservative Lagrangian and Hamiltonian mechanics", **Physical Review E** , 53 , 1890–1899.
- [58] Samko S. G., Kilbas A. A., Marichev O. I. (1993), "**Fractional Integrals and Derivatives -Theory and Applications**, Gordon and Breach, Longhorne, Pennsylvania.
- [59] Baleanu D., Defterli O. and Agrawal O.P. (2009), "A central difference numerical scheme for fractional optimal control problems", **Journal of Vibration and Control**, 15, 583-597.
- [60] Sabouri J. , Effati S. and Pakdaman M. (2017), "A Neural Network Approach for Solving a Class of Fractional Optimal Control Problems", **Neural Processing Letters**, 45, 59-74.
- [61] Rakhshan S. A. , Effati S. and Vahidian Kamyad A. (2016), "Solving a class of fractional optimal control problems by the Hamilton-Jacobi-Bellman equation", **Journal of Vibration and Control**, 1-16.
- [62] Singha N. and Nahak Ch. (2017), "An Efficient Approximation Technique for Solving a Class of Fractional Optimal Control Problems", **Journal of Optimization Theory and Applications**, 174, 785-802.
- [63] Alizadeh A. and Effati S. (2016), "An iterative approach for solving fractional optimal control problems", **Journal of Vibration and Control**, DOI 10.1177/1077546316633391.
- [64] Pooseh S. , Almeida R. and Delfim F. M. Torres (2014), "Fractional Order Optimal Control Problems with Free Terminal Time", **J. Ind. Manag. Optim.** , 10, 363-381.
- [65] Hosseinpour S. , and Nazemi A. R. (2016), "Solving fractional optimal control problems with fixed or free final states by Haar wavelet collocation method", **IMA Journal of Mathematical Control and Information**, 33, 2, 543-561.

- [66] Hosseinpour S. and Nazemi A.R. (2016), "A collocation method via block-pulse functions for solving delay fractional optimal", **IMA Journal of Mathematical Control and Information**, 1-23, 10.1093/imamci/dnw020.
- [67] Bhrawy A.H. , Ezz-Eldien S.S., DohaE.H., Abdelkawy M.A. and BaleanuD. (2017), "Solving Fractional Optimal Control Problems within a Chebyshev-Legendre Operational Technique", **International Journal of Control**, doi.org/10.1080/00207179.2016.1278267.
- [68] Pooseh, S. and Almeida, R. and Torres, D. F. M. (2013), "A numerical scheme to solve fractional optimal control problems", **Conference papers in mathematics**, Volume , Article ID 165298, 10 Pages.
- [69] Rabiei K. , Ordokhani Y. and Babolian E. (2016), "The Boubaker polynomials and their application to solve fractional optimal control problems", **Nonlinear Dynamic**, DOI 10.1007/s11071-016-3291-2.
- [70] Almeida and Delfim R. and Torres F. M. (2015), "A discrete method to solve fractional optimal control problems", **Nonlinear Dynamic** , 80, 1811-1816.
- [71] Pooseh S., Almeida R. and Torres, D. F. M. (2014), "Fractional order optimal control problems with free terminal time", **J. Ind. Manag. Optim.** 10, 363-381.
- [72] Sweilam N. H. , Al-Ajami T. M. (2015), "Legendre Spectral-Collocation Method for Solving Some Types of Fractional Optimal Control Problems", **Journal of Advanced Research**, 6, pp. 393-403.
- [73] Tanga X. , Shi Y. and Li-Lian Wangc (2017), "A new framework for solving fractional optimal control problems using fractional pseudospectral methods", **Automatica**, 78, 333-340, .
- [74] Lotfi A. , Yousefi S.A. and Dehghan M. (2013), "Numerical solution of a class of fractional optimal control problems via the Legendre orthonormal basis combined with the operational matrix and the Gauss quadrature rule", **Journal of Computational and Applied Mathematics**, 250, 143-160.
- [75] Atangana A. and Clout A. H. (2013), "Stability and convergence of the space fractional variable-order Schrödinger equation", **Advances in Difference Equations** , 80, 1-10.

- [76] Zhuang P. , Liu F. , Anh V. and Turner I. (2009), "Numerical methods for the variable-order fractional advection-diffusion equation with a nonlinear source term", **SIAM Journal on Numerical Analysis** 47 , 3, 1760-1781.
- [77] Haykin S. (2007), "**Neural Networks :A Comprehensive Foundation**, third edition, Prentice-Hall, Upper Saddle River, NJ.
- [78] Tang H. , TanK. C. ,Yi Z. (2007), "**Neural Networks: Computational Models and Applications**, Springer, Berlin.
- [79] Muller B. , Reinhardt J. and Strickland M. T. (2002), " **Neural Networks: An Introduction**", second edition, Springer-Verlag, Berlin.
- [80] Shekari Beidokhti R. and Malek A. (2009), "Solving initial-boundary value problems for systems of partial differential equations using neural networks and optimization techniques", **Journal of the Franklin Institute**, 346, 898- 913.
- [81] Kumar M. and Yadav N. (2011), "Multilayer perceptrons and radial basis function neural network methods for the solution of differential equations", **A survey, Computers and Mathematics with Applications**, 62, 3796- 3811.
- [82] Shirvany Y. , Hayati M. and Moradian R. (2009), "Multilayer perceptron neural networks with novel unsupervised training method for numerical solution of the partial differential equations", **Applied Soft Computing**, 9, 20-29.
- [83] Vrabie D. and Lewis F. (2009), "Neural network approach to continuous-time direct adaptive optimal control for partially unknown nonlinear systems", **Neural Networks**, 22, 237- 246.
- [84] Cheng T. , Lewis F. L. , Abu-Khalaf M. (2007), "Fixed-final-time-constrained optimal control of nonlinear systems using neural network HJB approach, **IEEE Transactions on Neural Networks**, 18, 1725-1737.
- [85] Nazemi A. R. and Karami R. (2017), "A Neural Network Approach for Solving Optimal Control Problems with Inequality Constraints and Some Applications", **Journal of Neural Processing Letters**, 45, 995-1023, .
- [86] Effati S. and Pakdaman M. (2013), "Optimal control problem via neural networks", **Neural Computing and Applications**, 23, 2093-2100.

- [87] Nazemi A. R. (2012), "A dynamic system model for solving convex nonlinear optimization problems", **Communications in Nonlinear Science and Numerical Simulation**, 17, 1696-1705.
- [88] Nazemi A. R. (2013), "Solving general convex nonlinear optimization problems by an efficient neurodynamic model", **Engineering Applications of Artificial Intelligence**, 26, 685-696.
- [89] Nazemi A. R. and Omid F. (2012), "A capable neural network model for solving the maximum flow problem", **Journal of Computational and Applied Mathematics**, 236, 3498-3513.
- [90] Nazemi A. R. (2014), "A neural network model for solving convex quadratic programming problems with some applications", **Engineering Applications of Artificial Intelligence**, 32, 54-62.
- [91] Nazemi A. R. , Omid F. (2013), "An efficient dynamic model for solving the shortest path problem", **Transportation Research Part C**, 26, 1-19.
- [92] Nazemi A. R. , Sharifi E. (2013), "Solving a class of geometric programming problems by an efficient dynamic model", **Communications in Nonlinear Science and Numerical Simulation**, 18, 692-709.
- [93] Nazemi A. R. , Effati S. (2013), "An application of a merit function for solving convex programming problems", **Computers and Industrial Engineering**, 66, 212 -221.
- [94] Nazemi A. R. , Nazemi M. (2014), "A gradient-based neural network method for solving strictly convex quadratic programming problems", **Cognitive Computation**, 6, 484-495.
- [95] Jafarian A. , Mokhtarpour M. and Baleanu D. (2017), "Artificial neural network approach for a class of fractional ordinary differential equation", **Neural Computing and Applications**, 28, 765-773.
- [96] Zuniga-Aguilar C. J. , Coronel-Escamilla A. , J, Gomes-Aguilar F. , Alvarodo-Martinez V. M. and Romero-Ugalde H. M. (2018), "New numerical approximation for solving fractional delay differential equations of variable order using artificial neural networks", **The European Physical Journal PLUS**, DOI 10.1140/epjp/i2018-11917-0.

- [97] Malek A. and Shekari Beidokhti R. (2006), "Numerical solution for high order differential equations using a hybrid neural network optimization method", **Appl. Math. Comput.**, 183, 260- 271.
- [98] Bazaraa M. S., Sherali H. D. and Shetty C. M. (2006), " **Nonlinear Programming-Theory and Algorithms** Third edition, John Wiley & Sons, Hoboken, New Jersey.
- [99] Zhang X-S. (2000), " **Neural Networks in Optimization**", Kluwer Academic Publishers, Dordrecht, Boston, London.
- [100] Lee K. Y. and El-Sharkawi M. A. (2008), " **Modern Heuristic Optimization Techniques: Theory and Applications to Power Systems**", IEEE Press Series on Power Engineering.
- [101] Miller R. K. and Michel A. N. (1982), " **Ordinary Differential Equations**", NewYork: Academic Press, 16.
- [102] Sun J., Chen J. S. and Ko C.-H. (2012), "Neural networks for solving second-order cone constrained variational inequality problem", **Computational Optimization and Applications**, 51, 623-648.
- [103] Saberi Nik H., Effati S. and Yildirim A. (2013), "Solution of linear optimal control systems by differential transform method", **Neural Computing and Applications**, 23, 1311- 1317.
- [104] Pan S.-H. and Chen J.-S. (2010), "A semismooth Newton method for the SOCCP based on a one-parametric class of SOC complementarity functions", **Computational Optimization and Applications**, 45, 59-88.
- [105] Facchinei F., Jiang H. and Qi L. (1999), "A smoothing method for mathematical programs with equilibrium constraints", **Mathematical Programming**, 35, 107-134.
- [106] Heydari M. H. (2018), " A new direct method based on the Chebyshev cardinal functions for variable-order fractional optimal control problems", **Journal of the Franklin Institute**, 1-26.
- [107] Mall S. , Chakraverty S. (2016), "Application of Legendre Neural Network for solving ordinary differential equations", **Applied Soft Computing**, 43, 347-356.

- [108] Chen G. , Dong X. (1998), "**From Chaos to Order: Methodologies, Perspectives and Applications**", World Scientific, Singapore.
- [109] Das A. , Roy A.B. and Das P. (2000), "Chaos in a three dimensional neural network", **Applied Mathematical Modelling**, 24 (7) ,511–522.
- [110] Yassen M.T. (2003), "Adaptive control and synchronization of a modified Chua's circuit system", **Applied Mathematics and Computation**, 135, 113-128.
- [111] Sadeghpour M., Salarieh H. and Alasty A. (2013), "Controlling chaos in tapping mode atomic force microscopes using improved minimum entropy control", **Applied Mathematical Modelling**, 37, 3, 1599-1606.
- [112] Yin C. , Dadras S. , Zhong S.M. and Chen Y.Q. (2013), "Control of a novel class of fractional-order chaotic systems via adaptive sliding mode control approach", **Applied Mathematical Modelling**, 37, 4, 2469-2483.
- [113] Hartley T. T. , Lorenzo C. F. and Qammer H. K. (1995), "Chaos in a fractional Chua's system", **IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications** , 42, 485-490.
- [114] Arena P. , Caponetto R., Fortuna L. and Porto D. (1998), "Bifurcation and chaos in noninteger order cellular neural networks", **International Journal of Bifurcation and Chaos**, 8, 1527-1539.
- [115] Petras I. (2011), "**Fractional order nonlinear systems: Modeling, Analysis and Simulation**", Springer Heidelberg Dordrecht, London, New York, .
- [116] Noori Skandari M. H. and Nazemi A. (2018), "A new approach to design asymptotically stabilizing control and adaptive control", **Optimal Control, Applications and Methods**, <https://doi.org/10.1002/oca.2458>.

Aabstract

Fractional optimal control problems are problems that their associated dynamical systems have derivatives and integrals of fractional order. These derivatives and integrals provide more accurate models for existing systems in nature. Due to the necessity of solving problems and the high capability of neural network models, which are one of the active and useful fields in the area of artificial intelligence, in this thesis we study a fractional optimal control problem with equality and inequality constraints. In order to solve these problems, we use numerical methods based on the perceptron neural network and Grunwald–Letnikov approximation, the neural network based on the Mittag-Leffler function, the neural network based on the fractional power series and also the neural network based on Legendre polynomials. On the other hand, some of the fractional optimal control problems have the time-dependent fractional derivatives which in this thesis, we will also solve these kinds of problems.

keywords: Fractional optimal control problem with fixed and variable order derivatives, Optimal condition, Artificial neural network, Unconstrained optimization, Stability and convergence.



Faculty Of Mathematical Sciences

PhD Thesis in Optimal Control and Optimization

Solving a class of fractional optimal control problems based on a method in artificial intelligence

By: Safiyeh Ghasemi Sahebi

Supervisor:

Dr. Ali Reza Nazemi

2019