

بالخدا



دانشکده علوم ریاضی

رساله دکتری تحقیق در عملیات

رویکرد بهینه سازی پویا برای حل مسائل رگرسیون فازی

نگارنده: دل آرا کرباسی

استادان راهنما

دکتر علیرضا ناظمی
دکتر محمدرضا ربیعی

بهمن ۱۳۹۹

به یاد پدرم که

خفته است اما...

به من، ایستادن آموخت

سپاس‌گزاری...

شکر شایان نثار ایزد منان که توفیق را رفیق راهم ساخت تا این پایان‌نامه را به پایان برسانم. در آغاز وظیفه‌ی خود می‌دانم از زحمات بی دریغ اساتید راهنمای خود، جناب آقایان دکتر علیرضا ناظمی و دکتر محمدرضا ربیعی صمیمانه تشکر و قدردانی کنم که از راهنمایی‌های ارزنده ایشان در راستای پیشبرد پژوهش، حاصل فراوان بردم و همواره شاگرد مکتب علم و انسانیت و منش والای ایشان هستم.

همچنین لازم می‌دانم از اساتید فرهیخته جناب آقایان دکتر روزبه، دکتر غزنوی و دکتر نوری که داوری این پایان‌نامه را به عهده گرفتند با تمام وجود تشکر و قدردانی نمایم.

به‌علاوه، بوسه می‌زنم بر دستان مادر عزیزم، فرشته‌ایی که سایه‌بان عشق و آرامش، تکیه گاه امن و آسایش و برترین آموزگار خوش‌بینی و امید در دوران تحصیل من بود. و تشکر می‌کنم از امین عزیزم به پاس حمایت سرشار و گرمای امیدبخش وجودش، که در سردترین دوران، بهترین پشتیبان من بود.

دل آرا کرباسی

بهمن ۱۳۹۹

تعهد نامه

اینجانب دل آرا کرباسی دانشجوی دکتری رشته ریاضی کاربردی علوم ریاضی دانشگاه شاهرود، نویسنده پایان نامه با عنوان **رویکرد بهینه سازی پویا برای حل مسائل رگرسیون فازی**، تحت راهنمایی **علیرضا ناظمی و محمدرضا ربیعی** متعهد می شوم:

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش های دیگر پژوهش گران، 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 هیچ جا 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 به دست آوردن نتایج اصلی پایان نامه تاثیرگذار بوده اند، در مقالات مستخرج از پایان نامه رعایت می گردد.
- در تمام مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافت های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده شده است)، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

دل آرا کرباسی

بهمن ۱۳۹۹

مالکیت نتایج و حق نشر

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

چکیده

در این پایان نامه یک رویکرد شبکه‌های عصبی بازگشتی برای تخمین ضرایب (پارامترهای) مدل رگرسیون خطی و چندجمله‌ای فازی ارائه می‌شود. در تحلیل رگرسیون در محیط فازی (مانند رگرسیون در نظریه آماری کلاسیک) زمانی که هم خطی چندگانه بین متغیرهای پیش بین وجود دارد از اعتبار مدل کاسته می‌شود. برای حل این مشکل، رگرسیون بریج در محیط فازی برای اولین بار در این پایان نامه معرفی می‌شود و نسبت به حل آن و حالت‌های خاص آن (رگرسیون ریج و لاسو فازی) با استفاده از یک شبکه عصبی کارا اقدام می‌شود. شبکه عصبی پیشنهادی بر اساس برخی مفاهیم بهینه‌سازی محدب و نظریه پایداری ساخته می‌شود که چارچوب آن، رسیدن به یک جواب تقریبی برای مدل‌های رگرسیون فازی را تضمین می‌کند. همچنین، وجود و همگرایی مسیرهای شبکه عصبی مورد مطالعه قرار داده می‌شود و پایداری لیاپانوف برای این شبکه عصبی اثبات می‌شود.

کلمات کلیدی: رگرسیون خطی فازی، رگرسیون چندجمله‌ای فازی، رگرسیون بریج فازی، شبکه عصبی بازگشتی، پایداری، همگرایی

لیست مقالات مستخرج از پایان نامه

1. Delara Karbasi, Alireza Nazemi, Mohammadreza Rabiei. (2020). A parametric recurrent neural network scheme for solving a class of fuzzy regression models with some real-world applications. *Soft Computing*. DOI 10.1007/s00500-020-05008-1.
2. Delara Karbasi, Mohammadreza Rabiei, Alireza Nazemi. (2020). A generalized bridge regression in fuzzy environment and its numerical solution by a capable recurrent neural network. *Journal of Mathematics*. DOI 10.1155/2020/8838040.
3. Delara Karbasi, Alireza Nazemi, Mohammadreza Rabiei, A capable recurrent neural network model for solving a class of ridge fuzzy regression problems, revised.
4. Delara Karbasi, Mohammadreza Rabiei, Alireza Nazemi, Numerical solution of a multivariate fuzzy polynomial regression model with interaction using an efficient recurrent neural network, submitted.

Conference articles

1. Delara Karbasi, Alireza Nazemi, Mohammadreza Rabiei (2017), Efficient neurodynamic model for fuzzy linear regression model, The 1st Seminar on Control and Optimization, Mashhad-Iran.
2. Delara Karbasi, Alireza Nazemi, Mohammadreza Rabiei (2018), Fuzzy linear regression model with crisp coefficients: A dynamic optimization scheme, 11th International Conference of Iranian Operations Research Society, Kermanshah-Iran.
3. Delara Karbasi, Alireza Nazemi, Mohammadreza Rabiei (2018), Efficient neurodynamic model for interval-valued fuzzy linear regression, 14th Iranian Statistics Conference, Shahrood-Iran.
4. Delara Karbasi, Alireza Nazemi, Mohammadreza Rabiei (2018), Artificial neural network approach for solving fuzzy linear regression by fuzzy differential equations, The 2nd national seminar on control and optimization , Shahrood-Iran.

فهرست مطالب

ق فهرست تصاویر

ث فهرست جداول

۱ مفاهیم و تعاریف اولیه

۱ ۱.۱ رگرسیون غیرفازی

۲ ۱.۱.۱ رگرسیون خطی چندگانه

۲ ۲.۱.۱ برآورد کمترین توان‌های دوم

۳ ۳.۱.۱ هم‌خطی در رگرسیون آماری

۵ ۴.۱.۱ روش‌های برخورد با هم‌خطی

۶ ۵.۱.۱ رگرسیون چندجمله‌ای

۷ ۲.۱ رگرسیون فازی

۸ ۱.۲.۱ انواع رگرسیون فازی

۹ ۲.۲.۱ نیکویی برازش

۱۲ ۳.۱ مروری بر شبکه‌های عصبی

۱۲ ۱.۳.۱ شبکه‌های عصبی طبیعی

۱۴ ۲.۳.۱ شبکه‌های عصبی مصنوعی

۳.۳.۱ تاریخچه‌ای از حل مسائل بهینه‌سازی با استفاده از شبکه‌های

عصبی

۱۷ ۴.۳.۱ رویکرد شبکه‌عصبی بازگشتی

۲۹ ۲ مدل شبکه عصبی بازگشتی برای حل مسائل رگرسیون فازی

۳۰ ۱.۲ یک روش شبکه عصبی پارامتریک برای حل مدل‌های رگرسیون خطی فازی

۳۰ ۱.۱.۲ مدل بهینه‌سازی

۳۳ ۲.۱.۲ مدل رگرسیون چندجمله‌ای فازی با یک متغیر و از درجه کامل

۳۳ ۳.۱.۲ مدل رگرسیون چندجمله‌ای فازی با دو متغیر و از درجه دوم

۳۴	۴.۱.۲	رویکرد شبکه عصبی بازگشتی
۳۹	۵.۱.۲	مثال‌های عددی
۶۲	۶.۱.۲	مزایای روش پیشنهادی
۶۶	۲.۲	یک روش شبکه عصبی برای حل مدل رگرسیون چندجمله‌ای فازی
۶۶	۱.۲.۲	مدل بهینه سازی
۶۹	۲.۲.۲	رویکرد شبکه عصبی بازگشتی
۷۱	۳.۲.۲	مثال‌های عددی
۸۱	۴.۲.۲	مزایای روش پیشنهادی
۸۵	۳	مدل شبکه عصبی بازگشتی برای حل مسائل رگرسیون ریج فازی
۸۶	۱.۳	مدل رگرسیون ریج فازی
۸۷	۲.۳	مدل بهینه سازی
۸۹	۳.۳	رویکرد شبکه عصبی بازگشتی
۹۰	۴.۳	مثال‌های عددی
۱۰۲	۵.۳	مزایای روش پیشنهادی
۱۰۳	۴	حل عددی رگرسیون بریج فازی برای مقادیر مختلف γ
۱۰۳	۱.۴	مدل رگرسیون بریج فازی پیشنهادی
۱۰۵	۲.۴	مدل بهینه سازی
۱۰۷	۳.۴	مدل رگرسیون بریج فازی برای $\gamma = 1$ و $\gamma = 2$
۱۰۹	۴.۴	رویکرد شبکه عصبی بازگشتی
۱۱۰	۱.۴.۴	بررسی همگرایی و پایداری
۱۱۱	۵.۴	مثال‌های عددی
۱۲۴	۶.۴	مزایای روش پیشنهادی
۱۲۹		مراجع
۱۳۹	آ	نظریه مجموعه‌های فازی
۱۳۹	۱.آ	مروری بر مفاهیم نظریه مجموعه‌های فازی
۱۴۲	۱.۱.آ	اعداد فازی
۱۴۳	۲.۱.آ	حساب اعداد فازی
۱۴۴	۳.۱.آ	فاصله اعداد فازی
۱۴۷	ب	اثبات قضایا و گزاره‌ها
۱۴۷	ب.۱	فصل ۲

ب.۲ فصل ۴ ۱۵۱

۱۵۵ واژه‌نامه فارسی به انگلیسی

۱۵۷ واژه‌نامه انگلیسی به فارسی

فهرست تصاویر

۱۳	 ساختار نرون	۱.۱
۱۵	 مدل ریاضی نرون عصبی مصنوعی	۲.۱
۱۶	 انواع تابع فعال سازی	۳.۱
۲۲	 نمودار شبکه عصبی (۲۱.۱) و (۲۲.۱)	۴.۱
۴۰	 رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با $h = ۰/۵$ در مثال ۱.۱.۲	۱.۲
	 رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با ۲۰ نقطه شروع متفاوت و	۲.۲
۴۰	 $h = ۰/۵$ در مثال ۱.۱.۲	
۴۷	 تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۱.۱.۲	۳.۲
۴۸	 نمودار حبابی فازی در مثال ۱.۱.۲	۴.۲
۴۹	 رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با $h = ۰/۵$ در مثال ۲.۱.۲	۵.۲
	 رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با ۲۰ نقطه شروع متفاوت و	۶.۲
۴۹	 $h = ۰/۵$ در مثال ۲.۱.۲	
۵۰	 تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۲.۱.۲	۷.۲
۵۱	 نمودار حبابی فازی در مثال ۲.۱.۲	۸.۲
۵۲	 رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با $h = ۱$ در مثال ۳.۱.۲	۹.۲
۵۳	 نمودار حبابی فازی در مثال ۳.۱.۲	۱۰.۲
۵۴	 رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با $h = ۱$ در مثال ۴.۱.۲	۱۱.۲
۵۶	 تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۴.۱.۲	۱۲.۲
۵۷	 نمودار حبابی فازی در مثال ۴.۱.۲	۱۳.۲
۵۹	 رفتار مسیر A شبکه عصبی (۱۹.۲) و (۲۰.۲) با $h = ۰/۳$ در مثال ۵.۱.۲	۱۴.۲
	 رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با ۲۰ نقطه شروع متفاوت و	۱۵.۲
۶۰	 $h = ۰/۳$ در مثال ۵.۱.۲	
۶۱	 تقریب ضرایب مدل در مثال ۵.۱.۲	۱۶.۲
۶۲	 نمودار حبابی فازی در مثال ۵.۱.۲	۱۷.۲
۶۳	 تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۶.۱.۲	۱۸.۲

۶۴	۱۹.۲ رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با $h = 5\%$ در مثال ۶.۱.۲
۶۴	۲۰.۲ نمودار حبابی فازی در مثال ۶.۱.۲
۷۴	۲۱.۲ رفتار مسیر A از شبکه عصبی (۴۲.۲) و (۴۳.۲) در مثال ۱.۲.۲
	۲۲.۲ رفتار مسیر A از شبکه عصبی (۴۲.۲) و (۴۳.۲) با 50° نقطه شروع متفاوت در
۷۵	مثال ۱.۲.۲
۷۵	۲۳.۲ رفتار مسیر A از شبکه عصبی (۴۲.۲) و (۴۳.۲) در مثال ۲.۲.۲
	۲۴.۲ رفتار مسیر A _۱ از شبکه عصبی (۴۲.۲) و (۴۳.۲) با 20° نقطه شروع متفاوت در
۷۶	مثال ۲.۲.۲
۷۷	۲۵.۲ رفتار مسیرهای A _۰ ، A _۱ و A _۲ از شبکه عصبی (۴۲.۲) و (۴۳.۲) در مثال ۳.۲.۲
۷۸	۲۶.۲ رفتار مسیرهای A _۳ ، A _۴ و A _۵ از شبکه عصبی (۴۲.۲) و (۴۳.۲) در مثال ۳.۲.۲
۷۹	۲۷.۲ رفتار مسیر A _۶ از شبکه عصبی (۴۲.۲) و (۴۳.۲) در مثال ۳.۲.۲
	۲۸.۲ رفتار مسیر A _۲ از شبکه عصبی (۴۲.۲) و (۴۳.۲) با 30° نقطه شروع متفاوت در
۷۹	مثال ۴.۲.۲
۸۳	۲۹.۲ رفتار مسیر A از شبکه عصبی (۴۲.۲) و (۴۳.۲) در مثال ۴.۲.۲
۸۴	۳۰.۲ رفتار مسیر A از شبکه عصبی (۴۲.۲) و (۴۳.۲) در مثال ۵.۲.۲
	۱.۳ مقادیر مختلف λ با دو فاصله متفاوت D _۳ و D _۴ با استفاده از روش ارایه شده
۹۱	در [۸۹] در مثال ۱.۴.۳
۹۴	۲.۳ رفتار مسیرهای A _۱ ، A _۲ و A _۳ از شبکه عصبی (۱۰.۳) و (۱۱.۳) در مثال ۱.۴.۳
	۳.۳ رفتار مسیرهای a _{۱u} ، a _{۲u} و a _{۳u} با 20° نقطه شروع متفاوت از شبکه عصبی (۱۰.۳)
۹۵	و (۱۱.۳) در مثال ۱.۴.۳
۹۵	۴.۳ نمودار تیلور برای مراکز در مثال ۱.۴.۳
۹۶	۵.۳ نمودار تیلور در مثال ۱.۱.۲
۹۸	۶.۳ مقادیر مختلف λ با فاصله D _۱ با استفاده از روش ارایه شده در [۸۹] در مثال ۲.۴.۳
۹۹	۷.۳ رفتار مسیر A از شبکه عصبی (۱۰.۳) و (۱۱.۳) در مثال ۲.۴.۳
	۸.۳ رفتار مسیرهای A _۱ با 20° نقطه شروع اولیه از شبکه عصبی (۱۰.۳) و (۱۱.۳) در
۱۰۰	مثال ۲.۴.۳
	۹.۳ نمودار تیلور برای مراکز اعداد فازی در مثال ۲.۴.۳ مدل [۸۶] با مربع سبز و
۱۰۰	مدل [۸۷] با مربع زرد نشان داده شده است.
	۱۰.۳ نمودار تیلور در مثال ۲.۴.۳ مدل [۸۶] با مربع سبز و مدل [۸۷] با مربع زرد
۱۰۱	نشان داده شده است.
	۱.۴ نیکویی برازش برای مقادیر مختلف γ با سه فاصله متفاوت D _۳ ، D _۴ و D _۲ در مثال
۱۱۳	۱.۵.۴

۲.۴	رفتار مسیر A_{24} در مدل رگرسیون بریج با $(\gamma = 3/7)$ برای شبکه عصبی (۴۱.۴)
۱۱۴	و (۴۲.۴) با 20 نقطه شروع متفاوت در مثال ۱.۵.۴
۳.۴	ادامه رفتار مسیرهای شبکه عصبی (۴۱.۴) و (۴۲.۴) در مثال ۱.۵.۴
۴.۴	ادامه رفتار مسیرهای شبکه عصبی (۴۱.۴) و (۴۲.۴) در مثال ۱.۵.۴
۵.۴	ادامه رفتار مسیرهای شبکه عصبی (۴۱.۴) و (۴۲.۴) در مثال ۱.۵.۴
۶.۴	ادامه رفتار مسیرهای شبکه عصبی (۴۱.۴) و (۴۲.۴) در مثال ۱.۵.۴
۷.۴	ادامه رفتار مسیرهای شبکه عصبی (۴۱.۴) و (۴۲.۴) در مثال ۱.۵.۴
۸.۴	نمودار حبابی فازی در مثال ۱.۵.۴
۹.۴	نمودار تیلور برای مراکز اعداد فازی در مثال ۱.۵.۴
۱۰.۴	نمودار تیلور برای مقادیر سمت چپ اعداد فازی ۱.۵.۴
۱۱.۴	نمودار تیلور برای مقادیر سمت راست اعداد فازی ۱.۵.۴
۱۲.۴	نیکویی برازش برای مقادیر مختلف γ با سه فاصله متفاوت D_1, D_2 و D_3 در مثال ۲.۵.۴
۱۲۳	رفتار مسیرهای شبکه عصبی (۴۱.۴) و (۴۲.۴) در مثال ۲.۵.۴
۱۲۵	نمودار حبابی فازی در مثال ۱.۵.۴
۱۲۶	ادامه نمودار حبابی فازی در مثال ۱.۵.۴
۱۲۷	رفتار مسیر A_7 برای شبکه عصبی (۴۱.۴) و (۴۲.۴) با 20 نقطه شروع متفاوت در مثال ۲.۵.۴
۱۲۷	نمودار تیلور برای مراکز اعداد فازی در مثال ۱.۵.۴
۱۲۸	نمودار تیلور برای مقادیر سمت چپ و راست اعداد فازی در مثال ۱.۵.۴

فهرست جداول

۳۹	مقادیر مشاهده شده و خطای به‌دست آمده در مثال ۱.۱.۲	۱.۲
۴۱	تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۱.۱.۲	۲.۲
۴۱	مقادیر مشاهده شده در مثال ۲.۱.۲	۳.۲
۴۲	تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۲.۱.۲	۴.۲
۴۲	ضرایب فازی و خطای به‌دست آمده در مثال ۲.۱.۲	۵.۲
۴۳	ضرایب مدل رگرسیون (۲۸.۲) در مثال ۳.۱.۲	۶.۲
۴۳	مقادیر مشاهده شده در مثال ۳.۱.۲	۷.۲
۵۵	تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۴.۱.۲	۸.۲
۵۵	مقادیر مشاهده شده و خطای به‌دست آمده در مثال ۴.۱.۲	۹.۲
۵۸	مقادیر مشاهده شده در مثال ۵.۱.۲	۱۰.۲
۵۸	تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۵.۱.۲	۱۱.۲
۵۹	ضرایب فازی و خطای به‌دست آمده در مثال ۵.۱.۲	۱۲.۲
۶۰	مقادیر مشاهده شده و خطای به‌دست آمده در مثال ۶.۱.۲	۱۳.۲
۶۴	تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۶.۱.۲	۱۴.۲
۷۱	مقادیر مشاهده شده در مثال ۱.۲.۲	۱۵.۲
۷۲	خطای به‌دست آمده در مثال ۱.۲.۲	۱۶.۲
۷۲	روش MCV در مثال ۱.۲.۲	۱۷.۲
۷۳	مقادیر مشاهده شده و خطای به‌دست آمده در مثال ۲.۱.۲	۱۸.۲
۷۶	تقریب ضرایب مدل در مثال ۳.۲.۲	۱۹.۲
۷۶	خطای به‌دست آمده در مثال ۳.۲.۲	۲۰.۲
۸۰	$MSPE$ در مثال ۴.۲.۲	۲۱.۲
۸۰	روش MCV در مثال ۴.۲.۲	۲۲.۲
۸۰	مقادیر مشاهده شده در مثال ۵.۲.۲	۲۳.۲
۸۱	خطای به‌دست آمده در مثال ۵.۲.۲	۲۴.۲
۸۲	خطای به‌دست آمده در مثال ۶.۲.۲	۲۵.۲

۹۳	مقادیر مشاهده شده در مثال ۱.۴.۳	۱.۳
۹۷	تخمین مقادیر MCV برای داده‌های مثال ۱.۴.۳	۲.۳
۹۷	مقادیر مشاهده شده در مثال ۲.۴.۳	۳.۳
۹۸	تخمین مقادیر MCV برای داده‌های مثال ۱.۴.۳	۴.۳
۹۸	تخمین مقادیر $MSPE$ برای داده‌های مثال ۲.۴.۳	۵.۳
	میانگین خطای ۳۰ مجموعه داده شبیه‌سازی شده برای مقادیر مختلف γ در	۱.۴
۱۱۲	مقایسه با برخی تکنیک‌های متداول برای حل مدل رگرسیون فازی در مثال ۱.۵.۴	۲.۴
	خطای ۱۵ امین مجموعه داده شبیه‌سازی شده برای مقادیر مختلف γ در مثال	
۱۱۳	۱.۵.۴	۳.۴
	ضرایب مدل رگرسیون بریج با $\gamma = 3/7$ ، مطابق با ۱۵ امین مجموعه داده در مثال	
۱۱۴	۱.۵.۴	۴.۴
۱۲۲	مقادیر مشاهده شده و خطای به‌دست آمده در مثال ۲.۵.۴	۵.۴
	مقایسه خطای به‌دست آمده از مدل رگرسیون بریج برای $\gamma = 2$ با برخی تکنیک‌های	
۱۲۴	متداول برای حل مدل رگرسیون فازی در مثال ۲.۵.۴	

فصل ۱

مفاهیم و تعاریف اولیه

۱.۱ رگرسیون غیرفازی

در کتب آماری رگرسیون به عنوان روش تعیین و تحلیل روابط نادقیق بین متغیرهای آماری تعریف می‌شود. بنابراین در رگرسیون، تعیین روابط بین متغیرها و تحلیل روابط به‌دست آمده مورد توجه قرار گرفته است. کلمه رگرسیون در لغت به معنی «برگشت» است. دلیل این نام‌گذاری به مطالعات فرانسیس گالتون^۱ [۵] در مورد رابطه بین قد پدر و قد پسر مربوط می‌شود. گالتون با رسم نمودار قد پسر در مقابل قد پدر (نمودار پراکنش) بر اساس حجم بزرگی از داده‌ها به این نتیجه رسید که اغلب افراد بلندقد، پسرانی کوتاه‌قدتر و اکثر افراد کوتاه‌قد، پسرانی بلندتر از خود دارند و این پدیده را «برگشت به میانگین» نامید و از آن پس نام رگرسیون بر روش‌های بررسی روابط بین متغیرهای آماری باقی ماند.

در یک معادله رگرسیونی، متغیر پاسخ را با Y و متغیرهای پیش‌بین را با $x_1, x_2, \dots, x_p$ نشان می‌دهند. متغیرهای پیش‌بین را «متغیرهای مستقل» نیز می‌نامند، اما به کار بردن این اصطلاح ممکن است تا حدی گمراه‌کننده باشد، زیرا همیشه $x_1, x_2, \dots, x_p$ مستقل از یکدیگر نیستند و معمولاً از Y مستقل نمی‌باشند. متغیر پاسخ نیز گاهی «متغیر وابسته» نامیده می‌شود. برای اطلاعات بیشتر، به [۱] مراجعه شود.

^۱Francis Galton

۱.۱.۱ رگرسیون خطی چندگانه

مدل متداول برای رگرسیون خطی چندگانه (یا به اختصار رگرسیون خطی) به صورت زیر است:

$$Y_i = A_{i0}x_{i0} + A_{i1}x_{i1} + \dots + A_{ip}x_{ip} + \epsilon_i, \quad i = 1, 2, \dots, n \quad (1.1)$$

که در آن Y_i متغیر پاسخ است، ضرایب رگرسیونی مدل هستند و ϵ_i خطای تصادفی است همچنین فرض می‌شود $x_{i0} = 1$ برای $i = 1, 2, \dots, n$. در این مدل فرض می‌کنیم که متغیرها و مشاهدات مربوط به آن‌ها غیرفازی هستند و خطای مدل نیز به خطاهای تصادفی و عدم حضور برخی متغیرها که اثرات جزئی بر متغیر وابسته دارند و در مدل در نظر گرفته نشده‌اند، نسبت داده می‌شود. به علاوه، مفروضاتی نیز در مورد خطاهای مدل در نظر گرفته می‌شود. به عنوان مثال، فرض می‌شود که این خطاهای تصادفی، نرمال، ناهمبسته و دارای میانگین صفر و واریانس ثابت هستند.

برای برآورد ضرایب رگرسیون، معمولاً روش کمترین توان‌های دوم را به کار می‌گیرند. باید توجه داشت که برای پیدا کردن برآورد کمترین توان‌های دوم ضرایب رگرسیون، لازم نیست که فرض شود بردار خطای رگرسیون از توزیع نرمال پیروی می‌کند. مدل رگرسیون خطی چندگانه را می‌توان به صورت ماتریسی زیر بیان نمود:

$$Y = XA + \epsilon \quad (2.1)$$

که در آن $Y = (Y_1, Y_2, \dots, Y_n)^T$ یک بردار n بعدی از پاسخ‌ها، $X = (x_{i0}, x_{i1}, \dots, x_{ip})^T$ یک ماتریس $n \times (p+1)$ بعدی (ماتریس طرح)، $A = (A_0, A_1, \dots, A_p)^T$ بردار ضرایب رگرسیون مجهول (پارامترهای مدل) و $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_n)^T$ بردار خطای تصادفی هستند.

۲.۱.۱ برآورد کمترین توان‌های دوم

برآورد کمترین توان‌های دوم A ، از طریق کمینه کردن مجموع توان‌های دوم خطا به صورت زیر به دست می‌آید:

$$\hat{A} = \arg \min_A [(Y - XA)^T(Y - XA)] \quad (3.1)$$

بنابراین، برآورد مدل رگرسیون (۲.۱) به صورت $\hat{Y} = X\hat{A}$ به دست می‌آید.

برآورد کمترین توان‌های دوم برای مدل رگرسیون (۲.۱) به صورت $\hat{A} = (X^T X)^{-1} X^T Y$ است که در آن $X^T X$ یک ماتریس نامنفرد^۲ (ستون‌های X از هم مستقل‌اند) فرض شده است. واضح است که این روش پارامتری هنگامی مناسب است که فرض خطی بودن رابطه میان متغیر پاسخ و متغیر پیش‌بین دست کم به صورت تقریبی برقرار باشد. اما هنگامی که رابطه میان متغیر پاسخ و متغیرهای پیش‌بین نامعلوم است، استفاده از رگرسیون خطی معمولی ممکن است منجر به نتایج گمراه کننده‌ای شود.

^۲ Non-singular matrix

۳.۱.۱ هم خطی در رگرسیون آماری

تعبیر و تفسیر معادله رگرسیون خطی به طور ضمنی به این فرض وابسته است که متغیرهای پیش‌بین، وابستگی با یکدیگر نداشته و به عبارتی مستقل خطی هستند. زیرا تعبیر یک ضریب رگرسیون به صورت اندازه‌گیری تغییر در متغیر پاسخ وقتی متغیر پیش‌بین متناظر، یک واحد افزایش یافته و سایر متغیرهای پیش‌بین ثابت نگاه داشته شوند امری متداول است؛ اما این تعبیر در صورتی که روابط خطی قوی بین متغیرهای پیش‌بین وجود داشته باشد معتبر نخواهد بود. غالباً تفسیر و استفاده از الگوی رگرسیون خطی به برآورد تک تک ضرایب رگرسیون بستگی دارد. اگر در یک الگوی رگرسیون متغیرهای پیش‌بین به طور کامل یا تقریباً کامل، وابسته خطی باشند، در این حالت ممکن است استنباط برای این الگوی رگرسیون گمراه کننده باشد. در این صورت می‌گوییم بین متغیرها در الگوی رگرسیون، هم خطی^۳ وجود دارد.

اگر یکی از متغیرهای پیش‌بین یک تابع دقیق خطی از یک یا چند متغیر پیش‌بین دیگر باشد گوییم رگرسیون دارای هم خطی کامل^۴ است. هم خطی ناقص^۵ وقتی اتفاق می‌افتد که یکی از متغیرهای پیش‌بین به طور تقریبی یک تابع خطی از یک یا چند متغیر پیش‌بین دیگر باشد. به عبارت دیگر، هم خطی کامل زمانی رخ می‌دهد که دست کم به ازای یک i برای $i = 1, 2, \dots, n$ داشته باشیم $R_i^2 = 1$ و هم خطی ناقص زمانی پیش می‌آید که دست کم به ازای یک i داشته باشیم $R_i^2 \cong 1$ که در آن R_i^2 ضریب تعیین^۶ رگرسیون خطی متغیر x_i بر سایر متغیرهای پیش‌بین است. باید دقت داشته باشیم که ضریب تعیین R^2 نسبت پراکندگی بیان شده به پراکندگی کل است. لذا در مدل رگرسیون (۱.۱) با فرض $Y = (Y_1, Y_2, \dots, Y_n)^T$ و $X = (x_{i1}, x_{i2}, \dots, x_{in})^T$ ضریب تعیین Y روی X برابر است با:

$$R^2 = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (4.1)$$

که در آن $\hat{Y}$ متغیر پاسخ بدست آمده و $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$. برای اطلاع بیشتر در مورد هم خطی چندگانه به [۶] مراجعه شود.

اثرات هم خطی بر برآورد پارامترها

هم خطی ناقص باعث می‌شود که دترمینان ماتریس $X^T X$ بسیار کوچک شود و در نتیجه ماتریس $(X^T X)^{-1}$ ، دارای درایه‌هایی باشند که از لحاظ قدرمطلق بسیار بزرگ هستند و به بینهایت میل می‌کنند. به عبارت دیگر، واریانس برآوردگر کمترین توان‌های دوم A یعنی $\hat{A} = (X^T X)^{-1} X^T Y$ بسیار بزرگ می‌شود. در این صورت، $\hat{A}$ برآورد دقیقی از A نخواهد بود و دقت برآورد ضرایب رگرسیون بسیار کم می‌شود. علاوه بر این، $\hat{A}$ ناپایدار است. یعنی با تغییر

^۳ Collinearity

^۴ Perfect

^۵ Nearly

^۶ Coefficient of determination

کوچکی در مقادیر متغیرهای پیش‌بین، $\hat{A}$ تغییرات زیادی پیدا می‌کنند و در نتیجه محاسبه $\hat{A}$ با دو روش متفاوت ممکن است جواب‌های کاملاً متفاوتی به دست دهد. از طرفی چون $cov(\hat{A}) = (X^T X)^{-1}$ ، هم‌خطی باعث به‌دست آمدن فواصل اطمینان عریض برای پارامتر A می‌شود. علاوه بر این در صورت وجود هم‌خطی کامل، ماتریس $X^T X$ وارون‌پذیر نبوده (زیرا این ماتریس تمام رتبه Y نمی‌شود). و در نتیجه برای برآورد ضرایب، جواب یکتا وجود نخواهد داشت. بنابراین مهم است که بدانیم چه وقت هم‌خطی وجود دارد تا مواظب پیامدهای ممکن آن باشیم. برای مطالعات بیشتر در مورد آثار هم‌خطی به [۷] مراجعه شود.

عامل تورم واریانس به عنوان یک شاخص بررسی هم‌خطی

شاخص‌های زیادی برای تشخیص هم‌خطی در مدل رگرسیون ارائه شده است. برخی از پرکاربردترین آنها عبارت‌اند از: عامل تورم واریانس، عدد شرطی و محک ماتریس همبستگی. در این پایان‌نامه از شاخص عامل تورم واریانس به منظور بررسی وجود هم‌خطی در مدل استفاده می‌کنیم.

همان‌طور که گفته شد، یکی از شیوه‌های تشخیص وجود هم‌خطی که کاربرد زیادی دارد، عامل تورم واریانس A است که توسط مارکارت 9 [۸] نام‌گذاری شده است. این عامل نشان می‌دهد که واریانس ضرایب برآورد شده تا چه حد نسبت به حالتی که متغیرهای پیش‌بین همبستگی خطی ندارند، متورم شده است. عامل تورم واریانس از رابطه زیر به‌دست می‌آید:

$$VIF_j = \frac{1}{1 - R_j^2}, \quad j = 0, 1, \dots, p$$

که در آن R_j^2 ضریب تعیین مدلی است که متغیر مستقل x_j روی سایر متغیرهای مستقل برازش داده شده است. در حالتی که $R_j^2 = 0$ به عبارت دیگر x_j با سایر متغیرهای مستقل در مدل رگرسیونی رابطه خطی نداشته باشد، $VIF_j = 1$ و هنگامی که $R_j^2 \neq 0$ ، $VIF_j > 1$ خواهد بود که نشان می‌دهد واریانس x_j به دلیل وجود همبستگی بین متغیرهای مستقل چقدر متورم شده است و مقدار آن افزایش پیدا کرده است. هنگامی که یک متغیر مستقل، وابستگی خطی کامل با سایر متغیرهای مستقل مدل داشته باشد، ضریب تعیین مرتبط با آن برابر با یک خواهد بود. در نتیجه عامل تورم واریانس به سمت بینهایت میل خواهد کرد و از آن می‌توان نتیجه گرفت که واریانس x_j نیز بینهایت خواهد شد. مقدار عامل تورم واریانس برای متغیرهای مستقل، اغلب به عنوان شاخصی برای اندازه‌گیری شدت هم‌خطی در مدل استفاده می‌شود. اگر بیشترین مقدار عامل تورم واریانس در مدل بزرگ‌تر از یک باشد، نشان دهنده این است که وجود هم‌خطی در مدل، اثر نامطلوبی بر برآوردهای روش کمترین توان‌های دوم گذاشته است. تجربیات عملی حاکی از این است که اگر هر یک از VIF ها از ۱۰ تجاوز کنند، خطاری

^Y Full rank

^A Variance inflation factor

⁹ Marquardt

مبنی بر این است که ضرایب رگرسیونی ذیربط به علت هم‌خطی به صورت ضعیفی برآورد شده‌اند (برای اطلاعات بیشتر به [۹۰] مراجعه شود).

۴.۱.۱ روش‌های برخورد با هم‌خطی

در بخش‌های قبل گفتیم که وقوع هم‌خطی چندگانه بین متغیرهای پیش‌بین در تحلیل رگرسیون خطی ممکن است باعث ناپایداری شدید در برآوردهای کمترین توان‌های دوم پارامترهای رگرسیونی شود، به این معنی که بزرگی و علامت بردار پارامترها در نمونه‌های مختلف به طور قابل ملاحظه‌ای بی‌ثبات خواهد بود که در نتیجه‌ی آن، برآورد کمترین توان‌های دوم به دست آمده، قابل اعتماد نخواهد بود. روش‌های مختلفی برای غلبه بر مشکل هم‌خطی در مدل‌های خطی ارائه شده است (برای اطلاعات بیشتر به [۹-۱۱] مراجعه شود). در این پایان‌نامه با استفاده از برآوردهای بریج^{۱۰}، ریج^{۱۱} و لاسو^{۱۲} به مقابله با هم‌خطی می‌پردازیم.

برآوردگر بریج

روش‌های متعددی برای مقابله با هم‌خطی در مسائل رگرسیون مطرح شده است. یکی از این روش‌ها استفاده از برآوردگر بریج است که حالت‌های خاص آن برآوردهای ریج و لاسو می‌باشند.

فرانک^{۱۳} و فریدمن^{۱۴} [۱۲] برای مقابله با هم‌خطی، مجموع توان‌های دوم خطای به دست آمده در (۳.۱) را تحت قید $\sum_{j=0}^p |A_j|^\gamma \leq t$ کمینه کرده‌اند و به مسئله بهینه‌سازی زیر رسیده‌اند:

$$\hat{\mathbf{A}}_{Bridge} = \arg \min_{\mathbf{A}} [(\mathbf{Y} - \mathbf{XA})^\top (\mathbf{Y} - \mathbf{XA}) + \lambda \sum_{j=0}^p |A_j|^\gamma]$$

که در آن $\gamma, \lambda \geq 0$ ، پارامترهای کنترل بوده و میزان انقباض تحمیل شده به ضرایب را کنترل می‌کند. لازم به ذکر است که انتخاب بهینه γ منجر به افزایش کارایی برآوردگر خواهد شد. با توجه به ساختار برآوردگر بریج واضح است که به ازای $\gamma \in \{1, 2\}$ ، این برآوردگر بین دو برآوردگر لاسو و ریج رفتار می‌کند و درواقع بین این دو برآوردگر پُل می‌زند که دلیل اصلی نام‌گذاری آن است.

^{۱۰} Bridge

^{۱۱} Ridge

^{۱۲} Lasso

^{۱۳} Frank

^{۱۴} Friedman

برآوردگر لاسو

برآوردگر لاسو که به ازای $\gamma = 1$ در برآوردگر بریج به دست می آید، اولین بار توسط تیبشیرانی^{۱۵} [۱۳] ارائه گردید. در واقع برآوردگر لاسو از حل مسئله کمینه سازی (۳.۱) تحت قید $\sum_{j=0}^p |A_j| \leq t$ به دست می آید. وجود قدرمطلق در این مسئله باعث می شود که برآوردگر لاسو حل نظری نداشته باشد.

برآوردگر ریج

در این بخش ما از یکی از موثرترین روش های موجود برای مقابله با هم خطی به نام برآوردگر ریج که برای اولین بار توسط هورل^{۱۶} و کنار^{۱۷} [۱۴] معرفی شد استفاده می کنیم. مدل رگرسیون خطی (۱.۱) را در نظر بگیرید. برآوردگر ریج به پارامتر ریج λ بستگی دارد که با افزودن این پارامتر به ماتریس $\mathbf{X}^T \mathbf{X}$ به صورت زیر باعث بهبود وضعیت آن می شود:

$$\begin{aligned}\hat{\mathbf{A}}(\lambda) &= (\mathbf{X}^T \mathbf{X} + \lambda I_p)^{-1} (\mathbf{X}^T \mathbf{Y}) \\ &= (\mathbf{X}^T \mathbf{X} + \lambda I_p)^{-1} (\mathbf{X}^T \mathbf{X}) \hat{\mathbf{A}} \\ &= T(\lambda) \hat{\mathbf{A}}\end{aligned}$$

که در آن $T(\lambda) = (\mathbf{X}^T \mathbf{X} + \lambda I_p)^{-1} (\mathbf{X}^T \mathbf{X})$. در واقع برآوردگر ریج یک تبدیل خطی از برآوردگر کمترین توان های دوم است. بنابراین برآوردگر بریج از حل مسئله بهینه سازی (۳.۱) تحت قید $\mathbf{A}^T \mathbf{A} = \sum_{j=0}^p A_j^2 \leq t$ حاصل می شود.

۵.۱.۱ رگرسیون چندجمله ای

روش های رگرسیونی برای مدل سازی داده ها بین متغیرهای مستقل و وابسته به کار می روند. در این بین مدل های رگرسیون خطی از محبوبیت زیادی برخوردار هستند. ولی گاهی استفاده از مدل رگرسیون برای زمانی که ارتباط بین متغیر وابسته و مستقل روی یک خط راست قرار ندارد، نیز قابل استفاده است. رگرسیون چندجمله ای یکی از روش هایی است که با استفاده از یک چندجمله ای^{۱۸} بر حسب متغیر مستقل، ارتباط با متغیر وابسته را مدل بندی می کند. به بیانی دیگر، می توان یک مدل غیرخطی بین متغیرهای توصیفی و پاسخ را به صورت یک چندجمله ای درجه k تعیین کرد. حالت کامل رگرسیون چندجمله ای با اثر متقابل بین متغیرها و برای $i = 1, 2, \dots, n$ ، به صورت زیر نوشته می شود:

$$Y_i = A_0 x_{i0} + \sum_{j=1}^p A_j x_{ij} + \sum_{j=1}^p \sum_{l=1}^p A_{jl} x_{ij} x_{il} + \sum_{j=1}^p \sum_{l=1}^p \sum_{s=1}^p A_{jls} x_{ij} x_{il} x_{is} + \dots \quad (5.1)$$

^{۱۵} Tibshirani

^{۱۶} Horel

^{۱۷} Kenard

^{۱۸} Polynomial

مدل رگرسیون چندجمله‌ای مرتبه اول یک خط مستقیم است (در واقع در این حالت مدل رگرسیون چندجمله‌ای به رگرسیون خطی تبدیل می‌شود)؛ مدل مرتبه دوم یک سهمی است و مدل مرتبه سوم یک منحنی S -شکل است؛ و به همین صورت برای مراتب بالاتر.

۲.۱ رگرسیون فازی

پس از معرفی نظریه مجموعه‌های فازی توسط پروفسور لطفی عسگرزاده، دانشمند ایرانی تبار، در سال ۱۹۶۵، موضوع مدل‌سازی و تحلیل رگرسیونی در محیط فازی مورد توجه پژوهشگران علوم نظری و کاربردی قرار گرفته است (برای اطلاعات بیشتر در مورد نظریه مجموعه‌های فازی به پیوست A مراجعه شود). نخست گفتنی است که منظور از مدل‌سازی و تحلیل رگرسیونی در محیط فازی عبارت است از همه وضعیت‌هایی که در آنها دست کم یکی از سه مؤلفه متغیرهای مورد مطالعه (یا مشاهدات آنها)، پارامترهای مدل یا جمله خطا فازی باشند. در موضوع رگرسیون در محیط فازی دو رویکرد اصلی وجود داشته و دارد: رویکرد مبتنی بر کمترین مجموع فاصله‌ها و رویکرد امکانی (رویکرد کمترین ابهام کل مدل، تحت برخی قیود). رویکرد نخست، یک تعمیم طبیعی از رویکرد کلاسیک کمترین مجموع فاصله‌ها (به ویژه، رویکرد کمترین توان‌های دوم) است و مبتنی بر تعمیم مفهوم فاصله است. در این رویکرد، مفهوم فاصله بین اعداد حقیقی به مفهوم فاصله بین اعداد فازی تعمیم می‌یابد و در محاسبه مجموع توان دوم خطاها (یا مجموع قدر مطلق‌های خطاها) استفاده می‌شود. رویکرد دوم، مبتنی بر کمترین ابهام کل مدل (تحت برخی قیود در مورد اعتبار مدل) است. در این رویکرد، موضوع برآورد پارامترهای مدل، تبدیل به یک مسئله برنامه‌ریزی خطی یا غیرخطی می‌شود که با حل آن، پارامترهای مورد نظر برآورد می‌گردند (برای اطلاعات بیشتر به [۱۵-۱۷] مراجعه شود).

با گذشت زمان، توانایی‌های نظری و کاربردی هر یک از دو رویکرد (و نیز انواع نسخه‌های تعدیل / تصحیح یافته‌ی آنها) آشکار شده است. در کنار این دو رویکرد اصلی، روش‌های ابتکاری متعددی نیز در موضوع رگرسیون فازی پیشنهاد شده‌اند، که البته برخی از این روش‌ها بر پایه ترکیب دو رویکرد بالا هستند (برای مثال به [۱۸-۲۱] مراجعه شود). برخی از روش‌های ابتکاری نیز بر اساس الگوریتم‌های محاسباتی خاص هستند (برای اطلاعات بیشتر در این زمینه به [۲۲، ۲۳] مراجعه شود). برخی دیگر، از سیستم‌های استنتاج فازی استفاده می‌کنند (برای نمونه‌هایی از این رویکرد [۲۴-۲۶] را ببینید). در برخی روش‌ها نیز از شبکه‌های عصبی مصنوعی یا الگوریتم‌های تکاملی استفاده می‌شود (برای مرور و بررسی این روش به [۲۷-۳۲] مراجعه کنید). از دید کلی، وجه قوت روش‌های ابتکاری، توانایی آنها در مسئله‌های خاص است.

۱.۲.۱ انواع رگرسیون فازی

در یک تقسیم‌بندی کلی، انواع رگرسیون فازی را می‌توان به سه حالت زیر تقسیم کرد:

۱. رگرسیون فازی در حالتی که ارتباط بین متغیرها فازی فرض می‌شود. به عبارت دیگر ضرایب معادله رگرسیونی، فازی در نظر گرفته می‌شوند.
 ۲. رگرسیون فازی در حالتی که متغیرها (اعم از پیش‌بین یا پاسخ) نادقیق و فازی هستند.
 ۳. رگرسیون فازی در حالتی که هم متغیرها و هم ضرایب مدل، فازی در نظر گرفته می‌شوند.
- لازم به ذکر است که تنوع در رگرسیون فازی، محدود به حالت‌های فوق نیست. بلکه شیوه‌هایی که برای هر حالت پیشنهاد شده است، تنوع بسیاری را در روش‌های رگرسیون فازی ایجاد کرده است.

رگرسیون کمترین توان‌های دوم خطای فازی

روش کمترین توان‌های دوم خطای فازی^{۱۹} نخستین بار توسط دیاموند^{۲۰} [۳۳] و کلمینس^{۲۱} [۳۴] ارائه شد. این روش تعمیمی از رگرسیون کمترین توان‌های دوم در حالت کلاسیک است، که بر پایه تعریف‌هایی برای فاصله بین اعداد فازی می‌باشد (برای اصلاحات بیشتر در این زمینه به پیوست^۲ مراجعه شود). دیاموند [۳۳] با تعریف فاصله‌ای بر اعداد فازی، مدل‌هایی را برای برازش کمترین توان‌های دوم مدل با متغیر مستقل دقیق و متغیر وابسته فازی و همچنین با متغیر مستقل و وابسته فازی، پیشنهاد نمود. دیاموند و کورنر^{۲۲} [۳۵] روش کمترین توان‌های دوم را برای برآورد متغیرهای تصادفی فازی ارائه کردند. همچنین ژو^{۲۳} و لی^{۲۴} [۳۶] روش کمترین توان‌های دوم را برای برآورد ضرایب فازی مدل با استفاده از متری بر اساس سطوح تراز ارائه کردند و با یک معیار نیکویی برازش^{۲۵} درستی مدل رگرسیونی را اندازه گرفتند. در روش کمترین توان‌های دوم فازی، پارامترها به گونه‌ای برآورد می‌شوند که به ازای هر مقدار، فاصله بین متغیر وابسته مشاهده شده و متغیر وابسته برآورد شده بر اساس فاصله D کمینه گردد. به عبارتی مانند روش کلاسیک به دنبال مینیمم کردن مجموع توان‌های دوم خطا هستیم. برای مثال اگر مشاهدات متغیر وابسته فازی باشند، می‌خواهیم مجموع توان‌های دوم خطا

$$\sum_{i=1}^n D^2(Y_i, \hat{Y}_i) \quad (۶.۱)$$

^{۱۹}Fuzzy least squares

^{۲۰}Diamond

^{۲۱}Celmins

^{۲۲}Korner

^{۲۳}Xu

^{۲۴}Li

^{۲۵}Goodness of fit

را کمینه کنیم که در آن D یک فاصله بین اعداد فازی است. در اینجا می‌خواهیم یک مدل خطی با ضرایب فازی A_j برای $j = 0, 1, \dots, p$ به صورت

$$\hat{Y}_i = A_0 x_{i0} + A_1 x_{i1} + \dots + A_p x_{ip}, \quad i = 1, \dots, n \quad (7.1)$$

را بر پایه‌ی مشاهداتی به صورت

$$(Y_i, x_{i0}, \dots, x_{ip}), \quad i = 1, \dots, n$$

برازش دهیم. بدون کاستن از کلیت مسئله فرض می‌کنیم $x_{ij} > 0$ و $x_{i0} = 1$ (در صورتی که x_{ij} ها اعدادی منفی باشند با انتقال می‌توان آنها را به اعدادی مثبت تبدیل کرد).

۲.۲.۱ نیکویی برازش

در رگرسیون فازی یک موضوع مهم، توان تشریح مدل می‌باشد. در این بخش، تعدادی معیار نیکویی برازش برای اندازه‌گیری توان برازش یک مدل رگرسیونی را معرفی می‌کنیم.

نیکویی برازش از منظر شاخص‌های عددی

۱. **ضریب تعیین:** یکی از شاخص‌های ارزیابی نیکویی برازش مدل رگرسیون آماری، محاسبه ضریب تعیین می‌باشد که در رابطه (۴.۱) نشان داده شده است. با الهام از این تعریف در آمار کلاسیک، برای مدل‌های رگرسیونی فازی، ضریب تعیین تعمیم یافته را می‌توان به صورت زیر تعریف کرد.

$$R^2 = \frac{\sum_{i=1}^n D^2(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n D^2(Y_i - \bar{Y})^2}$$

۲. **ضریب تعیین تعمیم یافته‌ی تعدیل شده:** می‌توان ادعا کرد که شاخص ضریب تعیین تعمیم یافته ممکن است قادر به شناسایی بهترین مدل برای بعضی داده‌های ارائه شده نباشد (برای اطلاعات بیشتر به [۳۷] مراجعه شود). بنابراین برای غلبه بر این مشکل، نسخه تعدیل شده زیر برای R^2 تعریف شده است.

$$\bar{R}^2 = 1 - (1 - R^2) \frac{n-1}{n-p} \quad (8.1)$$

۳. **میانگین توان دوم خطای پیشگویی:** میانگین توان دوم خطای پیشگویی ($MSPE$)^{۲۶} را می‌توان به عنوان معیار دیگری برای ارزیابی نیکویی برازش استفاده کرد. در زیر معیار $MSPE$ را در قالب یک تعریف می‌آوریم.

^{۲۶} Mean squares of prediction error

تعریف ۱.۲.۱. [۳۸] تحت مفروضات مدل رگرسیونی فازی (۷.۱) اگر $\hat{Y}_i$ مقدار پیشگویی Y_i باشد آن گاه میانگین توان دوم خطای پیشگویی برابر است با

$$MSPE = \frac{1}{n} \sum_{i=1}^n D^2(Y_i, \hat{Y}_i)$$

۴. **اعتبار سنجی متقابل:** یکی از شاخص‌های دیگر برای بررسی کارایی مدل، روش اعتبار سنجی متقابل^{۲۷} است. روش اعتبار سنجی متقابل برای سنجش قابلیت پیش‌بینی مدل به کار می‌بریم (برای اطلاعات بیشتر به [۳۹] مراجعه شود). برای این منظور هر دفعه i امین مشاهده از مجموعه داده‌ها خارج می‌شود، در حالی که مابقی مشاهدات به منظور ایجاد مدل رگرسیون فازی استفاده می‌شود. در این روش، مدل به دست آمده بدون در نظر گرفتن i امین مشاهده با $\hat{Y}_{(-i)(x_i)}$ نمایش داده می‌شود و برای پیش‌بینی مقدار پاسخ i امین مشاهده مورد استفاده قرار می‌گیرد. سرانجام برای مقایسه i امین مقدار مشاهدات Y_i و مقدار پیش‌بینی $\hat{Y}_{(-i)(x_i)}$ ، فاصله بین Y_i و $\hat{Y}_{(-i)(x_i)}$ را محاسبه می‌شود. میانگین این فواصل برای $i = 1, 2, \dots, n$ با MCV نشان داده می‌شود و به صورت زیر محاسبه می‌شود:

$$MCV = \frac{1}{n} \sum_{i=1}^n D^2(Y_i, \hat{Y}_{(-i)(x_i)}) \quad (9.1)$$

۵. **اندازه تشابه:** تا کنون چندین معیار تشابه برای سنجش تشابه بین مجموعه‌های فازی معرفی شده است. در این پایان‌نامه از شاخصی استفاده می‌کنیم که دارای مزایای خاصی نسبت به اندازه‌های تشابه دیگر است (برای اطلاعات بیشتر به [۴۰، ۴۱] مراجعه شود).

تعریف ۲.۲.۱. تحت مفروضات مدل (۷.۱)، میانگین اندازه تشابه^{۲۸} (MSM) به صورت زیر تعریف می‌شود:

$$MSM = \frac{1}{n} \sum_{i=1}^n S_{UI}(\hat{Y}_i, Y_i)$$

که در آن

$$S_{UI}(\hat{Y}_i, Y_i) = \frac{Card(\hat{Y}_i \cap Y_i)}{Card(\hat{Y}_i \cup Y_i)}$$

و $Card(A) = \int_x \mu_A(x) dx$ که در آن A یک عدد فازی است. اندازه تشابه بین $\circ$ تا 1 مقدار می‌گیرد و مدل با مقدار MSM بالاتر تناسب بهتری با داده‌ها دارد.

نیکویی برازش از منظر نموداری

برای به تصویر کشیدن عملکرد مدل در این پایان‌نامه از نمودارهای شناخته شده زیر استفاده می‌کنیم.

^{۲۷}Cross- validation

^{۲۸}Mean of similarity measures

۱. **نمودار تیلور:** نمودار تیلور^{۲۹} برای انتقال اطلاعات در مورد شباهت الگو بین مقدار برآورد شده و مشاهده شده به تصویر کشیده شده است. این شباهت از نظر همبستگی بین برآورد و داده‌ها، خطای جذر میانگین توان‌های دوم^{۳۰} (RMS) آنها و انحراف آنها بررسی می‌شود. در واقع این نمودار برای ارزیابی جنبه‌های مختلف مدل‌های پیچیده مناسب است (برای اطلاعات بیشتر به [۴۲] مراجعه شود). علاوه بر این، این نمودار مقایسه عملکرد بین مدل‌های مختلف را امکان پذیر کرده است. در این نمودار مجموعه داده مشاهده شده با یک نقطه در محور x به عنوان «داده‌های واقعی^{۳۱}» نشان داده می‌شود. برآوردها مطابق با سه کمیت آماری قرار داده می‌شوند:

الف. فاصله شعاعی نقطه دارای برچسب «داده‌های واقعی» از مبدا مختصات که نشان دهنده انحراف معیار مجموعه داده‌ها است.

ب. خطای RMS که نشان دهنده فاصله میان داده‌ها و برآوردهای متناسب با آنها است.

ج. همبستگی^{۳۲} ρ میان داده‌ها و برآوردهای آنها که با زاویه ایجاد شده تا محور x نشان داده می‌شود.

فرمول‌ها برای محاسبه خطای RMS مقدار همبستگی ρ ، انحراف معیار داده‌های مشاهده شده (σ_Y) و مقادیر به‌دست آمده ($\sigma_{\hat{Y}}$) به صورت زیر معرفی می‌شوند:

$$\rho = \frac{\frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - \bar{\hat{Y}})(Y_i - \bar{Y})}{\sigma_{\hat{Y}} \sigma_Y}$$

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2}$$

$$\sigma_{\hat{Y}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - \bar{\hat{Y}})^2}$$

$$\sigma_Y = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2}$$

که در آن $\bar{Y}$ و $\bar{\hat{Y}}$ به ترتیب مقدار میانگین محاسبه شده برای $\hat{Y}$ و Y هستند. در نمودار تیلور، وقتی فاصله از نقطه «داده‌های واقعی» نسبتاً کوتاه باشد نشان دهنده این است که داده‌ها به خوبی تقریب زده شده‌اند.

۲. **نمودار حبابی فازی:** نویسندگان در [۳۷] برای ارزیابی تناسب مدل رگرسیون فازی با خروجی فازی و ورودی‌های غیرفازی، از نمودارهای حبابی استفاده کرده‌اند، به طوری

^{۲۹} Taylor's diagram

^{۳۰} Root mean square

^{۳۱} Actual data

^{۳۲} Correlation

که محورهای x و y به ترتیب، مراکز مشاهده شده و تخمین زده شده را نشان می‌دهند. مرکز دایره‌ها به صورت $(y_{im}, \hat{y}_{im})$ و قطر آنها با $\frac{1}{2}(|y_{il} - \hat{y}_{il}| + |y_{ir} - \hat{y}_{ir}|)$ محاسبه می‌شود. در نمودارهای حبابی، نزدیک بودن این دایره‌ها به نیمساز صفحه نشان می‌دهد که مدل به درستی مراکز داده‌های فازی را تقریب می‌زند و هرچه قطر دایره‌های کوچک‌تر باشد، نشان دهنده این است که مدل به درستی پهنای اعداد فازی را تقریب می‌زند.

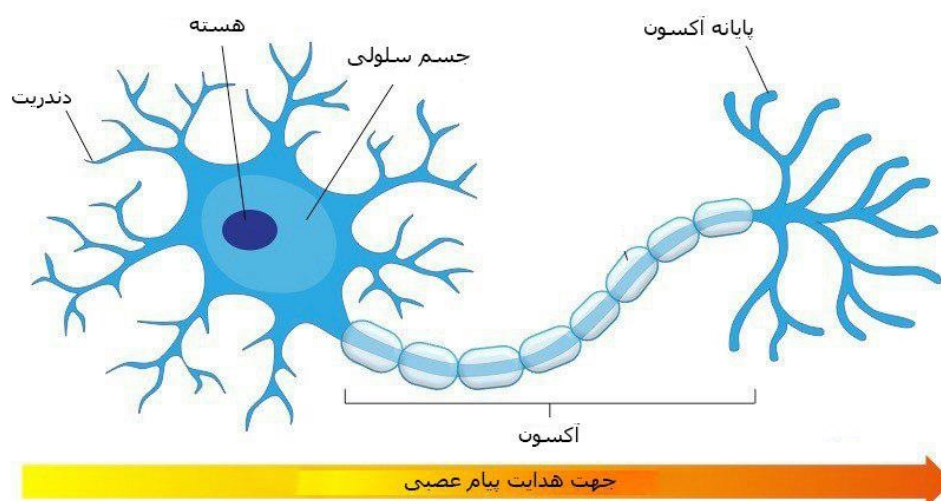
۳.۱ مروری بر شبکه‌های عصبی

در این فصل ابتدا ساختار کلی از شبکه‌های عصبی موجود در مغز انسان را معرفی می‌کنیم در واقع خواهیم دید که عملکرد سیستمی مغز به صورت شماتیک چگونه است. سپس با الگو برداری از رفتار سیستمی مغز، در می‌یابیم که می‌توانیم برخی از محاسبات را برحسب عملکرد مغز انجام دهیم که به آن هوش محاسباتی یا هوش مصنوعی می‌گویند. یکی از روش‌های کارا برای حل مسائل پیچیده، شکستن آن به زیر مساله‌های ساده‌تر است که هر کدام از این زیر مساله‌ها 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 می‌کنند.

در این فصل ابتدا نرون‌های شبکه‌های عصبی طبیعی معرفی شده و طرز کار آنها نشان داده شده است سپس مدل مصنوعی این نرون‌ها و ساختار آنها، مدل ریاضی آنها، شبکه‌های عصبی مصنوعی و نحوه آموزش و بکارگیری این شبکه‌ها نشان داده شده است. در این رساله تمرکز بیشتر بر نوعی از این شبکه‌ها به نام شبکه‌های عصبی مصنوعی بازگشتی می‌باشد. مطالب این فصل غالباً از [۲]. گرفته شده است.

۱.۳.۱ شبکه‌های عصبی طبیعی

مغز انسان در صد کمی از وزن بدن را تشکیل می‌دهد اما از لحاظ مصرف انرژی، بیست درصد انرژی بدن را مصرف می‌کند. ساده ترین واحد ساختاری شبکه عصبی مغز نرون ^{۳۳} ها یا عصب‌ها هستند. مغز از ۱۰۰ تریلیون 10^{11} نرون به هم مرتبط با تعداد 10^{16} پیوند تشکیل شده است و این ساختار، امکان پردازش موازی اطلاعات را موجب می‌شود. در این شبکه‌ها اگر یک سلول آسیب ببیند بقیه سلول‌ها می‌توانند نبود آن را جبران کرده و نیز در بازسازی آن سهیم باشند. این شبکه‌ها قادر به یادگیری‌اند. مثلاً با اعمال سوزش به سلول‌های عصبی لامسه، سلول‌های یاد می‌گیرند که به طرف جسم داغ نروند و با این الگوریتم سیستم می‌آموزد که خطای خود را اصلاح کند. یادگیری در این سیستم‌ها به صورت تطبیقی صورت می‌گیرد، یعنی با استفاده از مثال‌ها وزن سیناپس‌ها به گونه ای تغییر می‌کند که در صورت دادن ورودی‌های جدید سیستم پاسخ درستی تولید کند.



شکل ۱.۱: ساختار نرون

نرون‌ها با وجود تفاوت‌های زیاد از نظر اندازه و شکل ظاهری، مشخصه‌های مشترکی دارند. همان‌طور که در شکل ۱.۱ نشان داده شده است از تنه سلولی نرون تعدادی شاخک کوتاه خارج می‌شود که دندریت^{۳۴} نام دارد. دندریت‌ها و تنه سلولی، سیگنال‌ها را از نرون‌های مجاور دریافت می‌کنند و از طریق یک لوله باریک به نام آکسون^{۳۵} به نرون‌های دیگر منتقل می‌کنند. 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^{۳۶} می‌نامند. همه سیگنال‌های جمع‌بندی شده در تنه نرون ترکیب می‌شوند و اگر وسعت سیگنال‌های ترکیب شده به آستانه نرون برسد، مرحله تحریک شدن^{۳۷} 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^{۳۸} 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. ابتدا تصور می‌شد یک پیام عصبی ماهیت الکتریکی دارد اما بعدها ثابت شد که این پیام‌ها ماهیت الکتروشیمیایی دارند. لذا حفظ اطلاعات و ذخیره سازی آنها در مغز چیزی جز فعالیت‌های شیمیایی در بدنه سلول و پیوندهای آن نیست.

^{۳۴} Dendrite

^{۳۵} Axon

^{۳۶} Synapse

^{۳۷} Firing

^{۳۸} Neuro-transmitter

۲.۳.۱ شبکه‌های عصبی مصنوعی

شبکه عصبی مصنوعی^{۳۹} ایده‌ای است برای پردازش اطلاعات که از سیستم عصبی طبیعی الهام گرفته شده و مانند مغز به پردازش اطلاعات می‌پردازد. شبکه‌های عصبی شاخه‌ای از هوش مصنوعی است و در واقع تلاشی برای پیاده‌سازی نرون‌های عصبی مغز انسان به صورت مصنوعی است. عنصر کلیدی این ایده، ساختار جدید سیستم پردازش اطلاعات است. این سیستم از شمار زیادی عناصر پردازشی به هم پیوسته تشکیل شده که برای حل یک مساله با هم هماهنگ عمل می‌کنند، در این شبکه‌ها به کمک دانش برنامه‌نویسی، ساختار داده‌ای طراحی می‌شود که می‌تواند همانند نرون عمل کند. 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 الگو برداری از قوه یادگیری در انسان و پیاده سازی 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. هر چند شروع تحقیقات در زمینه شبکه‌های عصبی مصنوعی به دهه ۱۹۴۰ میلادی بر می‌گردد اما گسترش آن در دهه ۱۹۸۰ صورت پذیرفت [۳]. 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^{۴۰} بود که در سال ۱۹۸۲ منتشر شد [۴۳]. دومین اتفاق، حل مساله فروشنده دوره گرد^{۴۱} توسط شبکه هاپفیلد بود [۴۴]. هاپفیلد نشان داد که از این شبکه بازگشتی می‌توان برای حل مسائل بهینه سازی استفاده کرد و افق جدیدی را در شبکه‌های عصبی مصنوعی گشود.

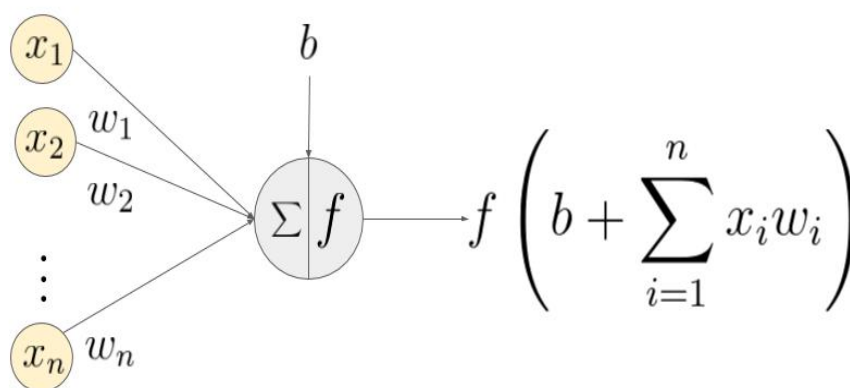
توپولوژی شبکه‌های عصبی مصنوعی

وضعیت نسبی نرون‌ها در شبکه (تعداد و گروه بندی و نوع اتصالات آن‌ها) را توپولوژی شبکه گویند. توپولوژی در واقع سیستم اتصال سخت افزار نرون‌ها به یکدیگر است که توام با نرم‌افزار مربوطه نوع عملکرد شبکه عصبی را تعیین می‌کند. در این توپولوژی یک لایه‌ی ورودی وجود دارد که اطلاعات را دریافت می‌کند، تعدادی لایه‌ی مخفی وجود دارد که اطلاعات را از لایه‌های قبلی می‌گیرند و در نهایت یک لایه‌ی خروجی وجود دارد که نتیجه محاسبات به آنجا می‌رود و جواب‌ها در آن قرار می‌گیرند. در یک نگاه ساده مدل یک نرون باید شامل n ورودی x_i ، $(i = 1, 2, \dots, n)$ باشد. این ورودی‌ها در وزن‌های سیناپسی w_i ، $(i = 1, 2, \dots, n)$ که روی آکسون‌ها واقع هستند و نقش ساختار سیناپسی را ایفا می‌کنند ضرب می‌شوند. حاصل ضرب

^{۳۹} Artificial neural network

^{۴۰} John Hopfield

^{۴۱} Travelling salesman problem



شکل ۲.۱: مدل ریاضی نرون عصبی مصنوعی

ورودی‌ها در وزن‌های متناظرشان با عددی مانند b موسوم به بایاس^{۴۲} جمع می‌شوند. این حاصل جمع را خروجی خطی نرون نامیده و با z نشان می‌دهیم. نهایتاً تابعی به نام تابع فعال‌ساز^{۴۳} به z اثر کرده و بدین ترتیب خروجی نرون حاصل می‌گردد. معادلات ریاضی چنین نرونی به صورت زیر است.

$$f(z) = f\left(\sum_{i=1}^n w_i x_i + b\right) \quad (۱۰.۱)$$

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 ۲.۱ نشان داده شده است. مشخصه‌های اصلی یک شبکه عصبی که در طراحی آن باید مد نظر قرار گیرند عبارت است از:

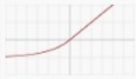
الف. **معماری شبکه عصبی:** نحوه اتصالات بین نرون‌ها، تعداد آنها و تعداد لایه‌های تشکیل دهنده بخش نرون‌های ارتباطی را معماری شبکه عصبی گویند.

ب. **تابع فعال‌ساز:** تابعی که انتخاب می‌شود تا روی نرون قرار گیرد و خروجی نرون را تولید کند تابع فعال‌ساز آن نرون گویند. در مورد ماهیت تابع f باید این نکته مورد توجه قرار بگیرد که ضابطه این تابع بستگی به نوع خروجی‌هایی که انتظار داریم شبکه آنها را حاصل کند مشخص می‌شود. در شکل ۳.۱ انواع توابع فعال‌سازی مشاهده می‌شود.

ج. **یادگیری شبکه عصبی:** یادگیری عبارت است از یافتن وزن‌ها، بایاس‌ها و توابع فعالیت به طوری که در یک دامنه کاری از هر ورودی، خروجی مطلوب حاصل گردد.

^{۴۲} Bias

^{۴۳} Activation function

نام	نمودار	معادله	مشتق
تابع همانی		$f(x) = x$	$f'(x) = 1$
تابع پله‌ای دودویی		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$
تابع لوجستیک (پله‌ای گام نرم)		$f(x) = \frac{1}{1 + e^{-x}}$	$f'(x) = f(x)(1 - f(x))$
تابع تانژانت هذلولوی		$f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
تابع آرک تانژانت		$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2 + 1}$
تابع واحد خطی یکسوسازی شده (ReLU)		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
تابع واحد خطی یکسوسازی شده پارامتریک (PReLU)		$f(x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
تابع واحد خطی نمایی (ELU)		$f(x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
SoftPlus		$f(x) = \log_e(1 + e^x)$	$f'(x) = \frac{1}{1 + e^{-x}}$

شکل ۳.۱: انواع تابع فعال‌سازی

۳.۳.۱ تاریخچه‌ای از حل مسائل بهینه‌سازی با استفاده از شبکه‌های

عصبی

در سال ۱۹۸۶ تانک^{۴۴} و هاپفیلد^[۴۵] اولین شبکه عصبی را برای حل مسائل برنامه‌ریزی خطی معرفی کردند. نشان دادند که وضعیت شبکه در هر تکرار به گونه‌ای تغییر می‌کند که تابع انرژی متناظر با آن به طور یکنواخت کاهش می‌یابد تا جایی که به نقطه کمینه خود می‌رسد که این نقطه کمینه متناظر با نقطه تعادل شبکه عصبی است. آنها این شبکه عصبی را توسط یک مدار الکتریکی پیاده‌سازی کردند. همچنین از این شبکه برای حل مساله فروشنده دوره‌گرد با ۳۰ شهر استفاده نمودند. این شبکه دارای نقص‌هایی بود به‌خصوص این که نقطه تعادل شبکه در شرایط بهینگی صدق نمی‌کرد و لذا جواب مطلوبی از مساله حاصل نمی‌شد. با این وجود کارهای هاپفیلد انگیزه بسیار خوبی را برای محققین به‌وجود آورد تا در این زمینه فعالیت کنند. کندی^{۴۵} و چا^{۴۶} [۴۶] با افزودن یک پارامتر جریمه متناهی به شبکه هاپفیلد آن را توسعه دادند و از آن برای حل مسئله برنامه‌ریزی غیرخطی محدب استفاده کردند. با پارامترهای جریمه این شبکه در یافتن نقطه بهینه دقیق ناتوان است به‌خصوص اگر پارامتر جریمه بزرگ باشد

^{۴۴} Tank^{۴۵} Kennedy^{۴۶} Chua

این شبکه به‌سختی عمل می‌کند.

برای جلوگیری از به‌کار بردن پارامتر جریمه یک شبکه عصبی سوئیچ-خازن^{۴۷} توسط رودریگز-وازکز^{۴۸} و همکاران [۴۷] معرفی شد. ما^{۴۹} و شانبلات^{۵۰} [۴۸] یک شبکه عصبی دو فازی ارائه کردند که فاز اول شبکه مشابه شبکه کندی و چا بود و در فاز دوم مسیر شبکه به جواب دقیق مساله همگرا می‌شد. بنابراین، این روش جواب‌های دقیق‌تری نسبت به شبکه کندی و چا ارائه داد. مشکل این شبکه در این بود که پایداری فاز دوم شبکه بستگی به انتخاب یک مقدار بزرگ از پارامتر جریمه داشت و لذا اگرچه تاثیر پارامتر جریمه در این روش کاهش یافته بود و جواب‌های حاصل دقیق‌تر بودند اما هنوز مستقل از پارامتر نبود.

نویسندگان در [۴۹] بر مبنای روش لاگرانژ یک شبکه عصبی ارائه کردند که کاملاً مستقل از پارامتر جریمه و قادر به حل مسائل غیرخطی بود. نقطه تعادل این شبکه عصبی در شرایط بهینگی مرتبه اول و دوم صدق می‌کرد و هم‌چنین شبکه حاصل همگرا بود. در سال ۱۹۹۳ بوزردوم^{۵۱} و پتیسن^{۵۲} [۵۰] شبکه‌ای را بر مبنای گرادیان و روش تصویر و مستقل از پارامتر جریمه ابداع کردند که تنها قادر به حل مسائل درجه دوم با متغیرهای کران‌دار بود. این روش در عمل روش کارایی بود اما نمی‌توانست حالت کلی مسائل برنامه‌ریزی خطی و درجه دوم را حل کند. چندین مدل شبکه عصبی برای حل مسائل بهینه‌سازی غیرخطی توسط عفتی و همکاران [۵۱-۵۴] ارائه شده است.

۴.۳.۱ رویکرد شبکه‌عصبی بازگشتی

شبکه‌های بازگشتی^{۵۳} دسته‌ای از شبکه‌ها هستند که رفتاری شبیه یک سیستم دینامیکی دارند، لذا به آنها شبکه‌های دینامیکی نیز می‌گویند. بازگشتی بودن آنها به این علت است که خروجی نرون‌ها به ورودی تمام نرون‌های دیگر بازگشت داده می‌شوند. پیش از معرفی شبکه عصبی بازگشتی لازم است به مرور برخی تعاریف و گزاره‌ها در سیستم‌های دینامیکی بپردازیم.

مروری بر سیستم‌های دینامیکی

در این بخش سیستم‌های دینامیکی^{۵۴} را معرفی نموده و قضیه‌ی اساسی وجود و یگانگی وجود جواب مسئله مقدار اولیه و قضایای مربوط به پایداری را ارائه می‌نماییم. مباحث این بخش عمدتاً برگرفته از [۲] است.

^{۴۷} Switched-Capacitor

^{۴۸} Rodriguez-Vazquez

^{۴۹} Maa

^{۵۰} Shanblatt

^{۵۱} Bouzerdoum

^{۵۲} Pattison

^{۵۳} Recurrent networks

^{۵۴} Dynamical systems

دستگاه معادله دیفرانسیل به شکل زیر را در نظر بگیرید:

$$\frac{dx}{dt} = f(x(t), t) \quad (11.1)$$

که در آن $x(t) = [x_i(t)]^T$ ، $t \geq 0$ بردار وضعیت یا حالت^{۵۵} نامیده می‌شود. چنین دستگاهی یک دستگاه وابسته به زمان یا غیرخودگردان^{۵۶} است. به این معنی که متغیر t به طور آشکار در ضابطه f ظاهر می‌شود. اگر f به طور صریح به t وابسته نباشد به عبارتی

$$\frac{dx}{dt} = f(x(t)) \quad (12.1)$$

در این صورت سیستم را خودگردان^{۵۷} یا مستقل از زمان نامند. به چنین دستگاه معادله دیفرانسیلی که بردار وضعیت آن با تغییر زمان تغییر می‌کند، یک سیستم دینامیکی گفته می‌شود. در ادامه از نماد $\dot{x}(t)$ برای نمایش مشتق x نسبت به زمان استفاده می‌کنیم.

قضیه ۱.۳.۱. (قضیه‌ی اساسی وجود و یگانگی جواب مسئله مقدار اولیه) فرض کنید Ω زیرمجموعه‌ی بازی از $\mathbb{R}^n$ و شامل نقطه x_0 باشد و f_i تابعی به طور پیوسته روی $\mathbb{R}^n$ مشتق‌پذیر است یا به عبارتی $f_i \in C^1(\mathbb{R})$ ، در این صورت به ازای هر $a > 0$ که $a \in \mathbb{R}$ ، مسئله مقدار اولیه زیر روی بازه $[-a, a]$ جواب یگانه $x(t)$ دارد:

$$\begin{cases} \dot{x}(t) = f(x(t)) \\ x(t_0) = x_0 \end{cases}$$

تعریف ۱.۳.۱. (نقطه‌ی تعادل^{۵۸}) جواب سیستم دینامیکی (۱۲.۱) را نقطه‌ی تعادل می‌نامیم و با x^e نشان می‌دهیم، هرگاه $f(x^e) = 0$. لذا نقطه تعادل جوابی از سیستم $\dot{x}(t) = f(x(t))$ است که مستقل از زمان می‌باشد.

تعریف ۲.۳.۱. (پایداری^{۵۹}) نقطه تعادل x^e ، نقطه پایدار نامیده می‌شود هرگاه

$$\forall \varepsilon \geq 0 \quad \exists \delta > 0 \quad s.t. \quad \|x(t_0) - x^e\| < \delta \implies \|x(t) - x^e\| < \varepsilon \quad \forall t \geq t_0$$

تعریف ۳.۳.۱. (پایداری مجانبی^{۶۰}) نقطه تعادل x^e به طور مجانبی پایدار است هرگاه پایدار باشد و $\lim_{t \rightarrow \infty} x(t) = x^e$.

یکی از بهترین روش‌ها برای تحلیل پایداری سیستم‌های دینامیکی روش لیاپانف^{۶۱} است. این روش برای سیستم‌های خطی و غیرخطی از هر مرتبه‌ای قابل استفاده است. این روش

^{۵۵}State vector

^{۵۶}Non-autonomous

^{۵۷}Autonomous

^{۵۸}Equilibrium

^{۵۹}Stability

^{۶۰}Asymptotic

^{۶۱}Lyapunov

شرط کافی برای پایداری مجانبی نقطه تعادل یک سیستم دینامیکی غیرخطی را می‌دهد. به عبارتی بدون حل معادلات وضعیت سیستم، می‌توان پایداری نقطه تعادل سیستم غیرخطی را بررسی کرد. در واقع لیاپانف پیشنهاد کرد که ممکن است توابعی وابسته به مسیر وجود داشته باشند که خواص آنها تعیین می‌کند که آیا مسیر به یک نقطه‌ی ساکن همگرا می‌شود یا خیر. چنین تابعی به‌عنوان تابع انرژی برای سیستم دینامیکی در نظر گرفته می‌شود که به تابع لیاپانف نیز معروف است.

قضیه‌ی لیاپانف به‌طور مختصر بیان می‌کند که نقطه‌ی تعادل یک سیستم دینامیکی پایدار است هرگاه یک تابع لیاپانف در همسایگی این نقطه وجود داشته باشد.

تعریف ۴.۳.۱. تابع لیاپانف تابعی مانند $E(x(t))$ است که در شرایط زیر صدق می‌کند:

۱. $E(x(t))$ و مشتقات جزئی مرتبه اول آن همگی پیوسته باشند، به عبارت دیگر $E(x(t)) \in C^1$.

۲. $E(x(t)) \geq 0$ ، به ویژه در یک همسایگی باز نقطه‌ی تعادل x^e ، $E(x(t)) > 0$ و $E(x^e) = 0$.

۳. مشتق تابع $E(x(t))$ نسبت به زمان منفی است، به عبارت دیگر:

$$\frac{dE(x(t))}{dt} = [\nabla_{x(t)} E(x(t))]^T \cdot \dot{x}(t) = [\nabla_{x(t)} E(x(t))]^T \cdot f(x(t)) \leq 0, x \in \Omega - \{x^e\}$$

و در نقطه‌ی تعادل x^e داریم $\frac{dE(x^e)}{dt} = 0$.

قضیه ۲.۳.۱. نقطه‌ی تعادل سیستم $\dot{x}(t) = f(x(t))$ پایدار است اگر یک تابع لیاپانف متناظر با این سیستم وجود داشته باشد.

قضیه ۳.۳.۱. نقطه‌ی تعادل سیستم $\dot{x}(t) = f(x(t))$ به طور مجانبی پایدار است اگر یک تابع لیاپانف متناظر با این سیستم وجود داشته باشد و در یک همسایگی نقطه‌ی تعادل x^e ، $\frac{dE(x(t))}{dt} < 0$.

ملاحظه می‌شود که اگر $\frac{dE(x(t))}{dt} \leq 0$ باشد، آن‌گاه x^e پایدار است ولی الزاماً به‌طور مجانبی پایدار نیست.

مدل شبکه‌ی عصبی بازگشتی انتخابی

مسئله برنامه‌ریزی درجه دوم محدب زیر را در نظر بگیرید:

$$\text{minimize } f(x) \quad (13.1)$$

subject to

$$g(x) \leq 0 \quad (14.1)$$

که در آن $\mathbf{g}(x) = (g_1(x), g_2(x), \dots, g_m(x))^T$ و $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ، $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$ همچنین توابع $f(x), g_1(x), g_2(x), \dots, g_m(x)$ محدب و دوبار مشتق پذیر هستند. در این پایان نامه برای یافتن شبکه عصبی متناسب همواره فرض می شود که مسئله (۱۳.۱) و (۱۴.۱) دارای جواب بهینه متناهی است که در شرط اسلاتر^{۶۲} [۵۵] صدق می کند (یعنی وجود دارد $x^\circ \in \mathbb{R}^n$ به طوری که $(g(x^\circ))^\circ < 0$).

با توجه به مطالب ارائه شده در [۵۶] مدل درجه دوم (۱۳.۱) و (۱۴.۱) به یک سیستم غیرخطی پویا تبدیل می شود. در ابتدا به مرور برخی نکات مورد استفاده در این بخش می پردازیم.

تعریف ۵.۳.۱. تابع $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ را پیوسته لیپشیتس^{۶۳} با ثابت L روی $\mathbb{R}^n$ گوئیم هرگاه برای هر $x, y \in \mathbb{R}^n$ داشته باشیم:

$$\|F(x) - F(y)\| \leq L\|x - y\| \quad (15.1)$$

F را پیوسته لیپشیتس محلی روی $\mathbb{R}^n$ گوئیم هرگاه برای هر نقطه از $\mathbb{R}^n$ یک همسایگی $D_\circ \subset \mathbb{R}^n$ وجود داشته باشد به طوری که نامساوی (۱۵.۱) برای هر $x, y \in \mathbb{R}^n$ برقرار باشد.

تعریف ۶.۳.۱. تابع $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ را یکنوا^{۶۴} گوئیم هرگاه

$$(x - y)^T (F(x) - F(y)) \geq 0, \forall x, y \in \mathbb{R}^n \quad (16.1)$$

و اکیداً یکنوا گوئیم هرگاه نامساوی (۱۶.۱) به طور اکید برای $x \neq y$ برقرار باشد.

لم ۱.۳.۱. [۵۷] اگر نگاشت F روی مجموعه محدب باز D که شامل Ω است، به طور پیوسته مشتق پذیر باشد، آن گاه F یکنوا (اکیداً یکنوا) روی Ω است، اگر و تنها اگر ماتریس ژاکوبین $\nabla F(x)$ برای تمام $x \in \Omega$ نیمه معین مثبت (معین مثبت) باشد.

تعریف ۷.۳.۱. [۵۷] فرض کنید $E: \mathbb{R}^n \rightarrow \mathbb{R}$ به طور پیوسته مشتق پذیر باشد. در این صورت $E(x)$ را به طور شعاعی بیکران گوئیم هرگاه

$$\|x\| \rightarrow \infty \Rightarrow E(x) \rightarrow \infty$$

تعریف ۸.۳.۱. [۵۷] 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 $\tau_1 > \tau$ می باشد گسترش یابد، آن گاه یک جواب ماکسیمال نامیده می شود و $[t_0, \tau)$ بازه ماکسیمال وجود جواب نامیده می شود.

قضیه ۴.۳.۱. [۵۸] اگر $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ یک تابع پیوسته باشد و $x(t)$ برای $t \in [t_0, \tau]$ که $\tau < \infty$ یک جواب ماکسیمال باشد، آن گاه $\lim_{t \rightarrow \tau} \|x\| = \infty$.

^{۶۲} Slater

^{۶۳} Lipschitz

^{۶۴} Monotone

قضیه ۵.۳.۱. [۵۹] (اصل تغییر ناپذیری لسال ^{۶۵}) فرض کنید E یک تابع به طور پیوسته مشتق‌پذیر باشد. همچنین فرض کنید:

۱. $M \subset \mathbb{R}^n$ یک مجموعه فشرد و پایدار نسبت به جواب سیستم (۱۲.۱) باشد.

۲. به ازای $x \in M$ ، $\frac{dE(x(t))}{dt} \leq 0$.

۳. F مجموعه همه نقاط M باشد به طوری که $\frac{dE(x(t))}{dt} = 0$.

۴. N بزرگترین مجموعه پایدار در F باشد.

آن‌گاه به ازای هر $x(t_0) \in M$ که $t_0 \geq 0$ می‌باشد داریم:

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

که در آن $\text{dist}(x, N) = \inf_{y \in N} \|x - y\|$.

از [۶۰] می‌دانیم که $x^* \in \mathbb{R}^n$ یک جواب بهینه از (۱۳.۱) و (۱۴.۱) است، اگر و تنها اگر وجود داشته باشد $u^* \in \mathbb{R}^n$ به طوری که $(x^{*\top}, u^{*\top})^\top$ در شرایط کروش-کان-تاکر ^{۶۶} زیر صدق کند:

$$\begin{cases} u^* \geq 0, & g(x^*) \leq 0, & u^{*\top} g(x^*) = 0 \\ \nabla f(x^*) + \nabla g(x^*)^\top u^* = 0 \end{cases} \quad (۱۷.۱)$$

در این صورت x^* را نقطه کروش-کان-تاکر مدل (۱۳.۱) و (۱۴.۱) و u^* را بردار ضرایب لاگرانژ متناظر با x^* می‌نامیم. علاوه بر این، x^* یک جواب بهینه از (۱۳.۱) و (۱۴.۱) است، اگر و تنها اگر x^* نقطه کروش-کان-تاکر مدل (۱۳.۱) و (۱۴.۱) باشد.

اکنون فرض کنید $x(\cdot)$ و $u(\cdot)$ دو متغیر وابسته به زمان باشند. آن‌گاه شبکه عصبی زیر با نقطه شروع $y_0 = (x_0, u_0)$ معرفی می‌شود:

$$\begin{cases} \frac{dx}{dt} = -(\nabla f(x) + \nabla g(x)^\top (u + g(x)))^+ \\ \frac{du}{dt} = (u + g(x))^+ - u \end{cases} \quad (۱۸.۱)$$

که در آن

$$(u + g(x))^+ = ([(u_1 + g_1(x))]^+, [(u_2 + g_2(x))]^+, \dots, [(u_m + g_m(x))]^+)^\top$$

$$[(u_k + g_k(x))]^+ = \max\{ (u_k + g_k(x)), 0 \}, \quad k = 1, 2, \dots, m$$

لم ۲.۳.۱. $(u + g(x))^+ = u$ اگر و تنها اگر

$$u \geq 0, \quad g(x) \leq 0, \quad u^\top (g(x)) = 0$$

^{۶۵} LaSalle principle of invariance

^{۶۶} Karush-Kuhn-Tucker

برهان. اگر $(u + g(x))^+ = u$ بنابرین $u = 0$ یا $u > 0$. پس حالت‌های زیر را داریم:
 (آ) اگر $u = 0$ ، پس $(u + g(x))^+ = (g(x))^+ = 0$ ؛ و در نتیجه $g(x) \leq 0$ و $u^T(g(x)) = 0$.
 (ب) اگر $u > 0$ ، پس $(u + g(x))^+ = u + g(x) = u$ ؛ و در نتیجه $g(x) = 0$ و $u^T(g(x)) = 0$.
 اثبات حالت برعکس سر راست است. $\square$

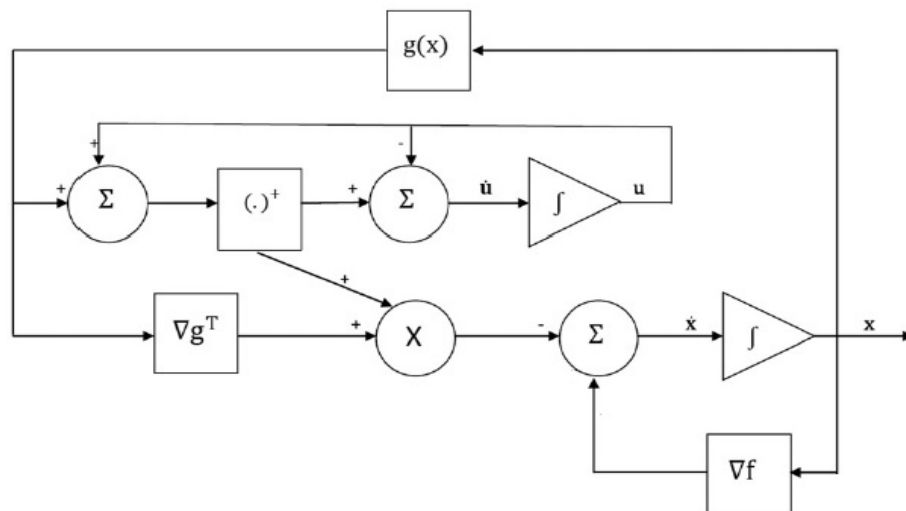
برای سادگی تعریف می‌کنیم:

$$\phi(y) = \begin{pmatrix} -(\nabla f(x) + \nabla g(x)^T(u + g(x))^+) \\ (u + g(x))^+ - u \end{pmatrix} \quad (20.1)$$

که در آن $y = (x^T, u^T)^T$. بنابراین شبکه‌عصبی (۱۸.۱) و (۱۹.۱) به صورت زیر بازنویسی می‌شود:

$$\begin{cases} \frac{dy}{dt} = \tau \phi(y) & (21.1) \\ y(t_0) = y_0 = (x(t_0)^T, u(t_0)^T)^T \in \mathbb{R}^{n+m}, \tau > 0 & (22.1) \end{cases}$$

که در آن τ نشان دهنده سرعت همگرایی شبکه‌عصبی (۲۱.۱) و (۲۲.۱) است (در این پایان‌نامه فرض می‌شود که $\tau = 1$).
 شکل (۴.۱) قابلیت اجرای سخت افزاری مدل شبکه‌عصبی ارائه شده را نشان می‌دهد.



شکل ۴.۱: نمودار شبکه‌عصبی (۲۱.۱) و (۲۲.۱)

بررسی همگرایی و پایداری

در این قسمت به بررسی گزاره‌های مربوط به همگرایی و پایداری مدل (۲۱.۱) و (۲۲.۱) می‌پردازیم.

قضیه ۶.۳.۱. فرض کنید $y^* = (x^{*\top}, u^{*\top})^\top$ نقطه تعادل مدل شبکه عصبی (۲۱.۱) و (۲۲.۱) باشد. در این صورت، x^* یک نقطه کروش-کان-تاکر برای مسئله (۱۳.۱) و (۱۴.۱) است. از سمت دیگر، اگر $x^* \in \mathbb{R}^n$ یک جواب مسئله (۱۳.۱) و (۱۴.۱) باشد، بنابراین $u^* \in \mathbb{R}^n$ وجود دارد به طوری که $y^* = (x^{*\top}, u^{*\top})^\top$ نقطه تعادل مدل شبکه عصبی (۲۱.۱) و (۲۲.۱) است.

برهان. فرض کنید $y^* = (x^{*\top}, u^{*\top})^\top$ ، یک نقطه تعادل از (۲۱.۱) و (۲۲.۱) باشد. بنابراین $\frac{dx^*}{dt} = 0$ و $\frac{du^*}{dt} = 0$ و نتیجه می‌شود که

$$\nabla f(x^*) + \nabla g(x^*)^\top (u^* + g(x^*))^+ = 0 \quad (23.1)$$

$$(u^* + g(x^*))^+ = u^* \quad (24.1)$$

از لم ۲.۳.۱ داریم، $(u^* + g(x^*))^+ = u^*$ اگر و تنها اگر

$$u^* \geq 0, \quad g(x^*) \leq 0, \quad u^{*\top} g(x^*) = 0 \quad (25.1)$$

علاوه بر این، با جایگذاری (۲۴.۱) در (۲۳.۱) داریم:

$$\nabla f(x^*) + \nabla g(x^*)^\top u^* = 0 \quad (26.1)$$

از (۲۵.۱) و (۲۶.۱)، دیده نتیجه می‌شود که $y^* = (x^{*\top}, u^{*\top})^\top$ در شرایط کروش-کان-تاکر (۱۷.۱) صدق می‌کند.

□

اثبات حالت برگشت سر راست است.

در ادامه ثابت می‌کنیم که مدل شبکه عصبی پایدار لیاپانوف است.

لم ۳.۳.۱. ژاکوبین ماتریس ϕ که در (۲۰.۱) تعریف شده است، نیمه معین منفی است.

برهان. بدون کاستن از کلیت فرض کنید وجود دارد $0 < m' < m$ به طوری که،

$$(u + g(x))^+ = \left((u_1 + g_1(x)), (u_2 + g_2(x)), \dots, (u_{m'} + g_{m'}(x)), \underbrace{0, 0, \dots, 0}_{m-m'} \right)^\top$$

بنابراین

$$\nabla \phi(y) = \begin{pmatrix} -\left(\nabla^\top f(x) + \sum_{k=1}^{m'} (u_k + g_k(x)) \nabla^\top g_k^*(x) \right) + \nabla g^*(x)^\top \nabla g^*(x) & -\nabla g^*(x)^\top \\ \nabla g^*(x)^\top & \mathcal{W}_{m \times m} \end{pmatrix}$$

که در آن

$$\nabla \mathbf{g}^*(\mathbf{x}) = \begin{pmatrix} \mathbf{R}_{m' \times n} \\ \mathbf{O}_{(m-m') \times n} \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_{m'-1}} & \frac{\partial g_1}{\partial x_{m'}} & \frac{\partial g_1}{\partial x_{m'+1}} & \dots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \dots & \frac{\partial g_2}{\partial x_{m'-1}} & \frac{\partial g_2}{\partial x_{m'}} & \frac{\partial g_2}{\partial x_{m'+1}} & \dots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_{m'}}{\partial x_1} & \dots & \frac{\partial g_{m'}}{\partial x_{m'-1}} & \frac{\partial g_{m'}}{\partial x_{m'}} & \frac{\partial g_{m'}}{\partial x_{m'+1}} & \dots & \frac{\partial g_{m'}}{\partial x_n} \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

و $\nabla^2 \mathbf{g}_k^*(\mathbf{x})$ برای $k = 1, 2, \dots, m'$ برابر است با

$$\nabla^2 \mathbf{g}_k^*(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 g_k}{\partial x_1^2} & \dots & \frac{\partial^2 g_k}{\partial x_1 \partial x_{m'-1}} & \frac{\partial^2 g_k}{\partial x_1 \partial x_{m'}} & \frac{\partial^2 g_k}{\partial x_1 \partial x_{m'+1}} & \dots & \frac{\partial^2 g_k}{\partial x_1 \partial x_n} \\ \frac{\partial^2 g_k}{\partial x_2 \partial x_1} & \dots & \frac{\partial^2 g_k}{\partial x_2 \partial x_{m'-1}} & \frac{\partial^2 g_k}{\partial x_2 \partial x_{m'}} & \frac{\partial^2 g_k}{\partial x_2 \partial x_{m'+1}} & \dots & \frac{\partial^2 g_k}{\partial x_2 \partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 g_k}{\partial x_{m'} \partial x_1} & \dots & \frac{\partial^2 g_k}{\partial x_{m'} \partial x_{m'-1}} & \frac{\partial^2 g_k}{\partial x_{m'} \partial x_{m'}} & \frac{\partial^2 g_k}{\partial x_{m'} \partial x_{m'+1}} & \dots & \frac{\partial^2 g_k}{\partial x_{m'} \partial x_n} \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

$$\mathcal{W}_{(m \times m)} = \begin{pmatrix} \mathbf{O}_{m' \times m'} & \mathbf{O}_{m' \times (m-m')} \\ \mathbf{O}_{(m-m') \times m'} & -I_{(m-m') \times (m-m')} \end{pmatrix}$$

و $\mathbf{O}$ نشان دهنده ماتریس صفر است.

دیده می‌شود که $\nabla \mathbf{g}^*(\mathbf{x})^\top \nabla \mathbf{g}^*(\mathbf{x})$ یک ماتریس نیمه معین مثبت است. و از آنجایی که توابع $f, g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_m(\mathbf{x})$ فرض شده اند که محدب و دوبار مشتق پذیر باشند بنابراین ماتریس‌های هسین $\nabla^2 f(\mathbf{x})$ و $\nabla^2 g_k(\mathbf{x})$ برای $k = 1, 2, \dots, m'$ نیمه معین مثبت هستند. علاوه بر این واضح است که $\mathcal{W}_{(m \times m)}$ نیمه معین منفی است. از نکات بیان شده نتیجه می‌شود که $\nabla \phi(\mathbf{y})$ یک ماتریس نیمه معین منفی است.

اگر $m' = m$ ، آن گاه $(\mathbf{u} + \mathbf{g}(\mathbf{x}))^+ = ((u_1 + g_1(\mathbf{x})), (u_2 + g_2(\mathbf{x})), \dots, (u_m + g_m(\mathbf{x})))$ بنابراین

$$\nabla \phi(\mathbf{y}) = \begin{pmatrix} -(\nabla^2 f(\mathbf{x}) + \sum_{k=1}^{m'} (u_k + g_k(\mathbf{x})) \nabla^2 g_k(\mathbf{x})) + \nabla \mathbf{g}(\mathbf{x})^\top \nabla \mathbf{g}(\mathbf{x}) & -\nabla \mathbf{g}(\mathbf{x})^\top \\ \nabla \mathbf{g}(\mathbf{x})^\top & \mathcal{W}_{m \times m} \end{pmatrix}$$

مشابه با حالت قبلی، اثبات می‌شود که $\nabla\phi(\mathbf{y})$ نیمه معین منفی است. نهایتاً اگر $m = 0$ ، آن‌گاه

$$(\mathbf{u} + \mathbf{g}(\mathbf{x}))^+ = (\underbrace{0, 0, \dots, 0}_m)^\top$$

بنابراین داریم

$$\nabla\phi(\mathbf{y}) = \begin{pmatrix} -\nabla^\top f(\mathbf{x}) & \mathbf{O}_{(n \times m)} \\ \mathbf{O}_{(m \times n)} & -I_{(m \times m)} \end{pmatrix}$$

در این حالت نیز نتیجه می‌شود که $\nabla\phi(\mathbf{y})$ یک ماتریس نیمه معین منفی است و این اثبات را $\square$ کامل می‌کند.

قضیه ۷.۳.۱. شبکه عصبی (۲۱.۱) و (۲۲.۱) پایدار لیاپانوف است.

برهان. تابع لیاپانوف $E: \mathbb{R}^m \rightarrow \mathbb{R}$ را به صورت زیر در نظر بگیرید:

$$E(\mathbf{y}) = E_1(\mathbf{y}) + E_2(\mathbf{y}) \quad (27.1)$$

که در آن $E_1(\mathbf{y}) = \|\phi(\mathbf{y})\|^2$ و $E_2(\mathbf{y}) = \frac{1}{\gamma} \|\mathbf{y} - \mathbf{y}^*\|^2$ می‌دانیم که $E_1(\mathbf{y})$ یک تابع مشتق‌پذیر است پس با توجه به (۲۰.۱)، دیده می‌شود که

$$\frac{d\phi}{dt} = \frac{\partial\phi}{\partial\mathbf{y}} \frac{d\mathbf{y}}{dt} = \nabla\phi(\mathbf{y})\phi(\mathbf{y})$$

با محاسبه مشتق $E(t)$ در نقطه $\mathbf{y}(t)$ از مدل شبکه عصبی (۲۱.۱) و (۲۲.۱) داریم:

$$\begin{aligned} \frac{dE(\mathbf{y}(t))}{dt} &= \left(\frac{d\phi}{dt}\right)^\top \phi + \phi^\top \left(\frac{d\phi}{dt}\right) + (\mathbf{y} - \mathbf{y}^*)^\top \frac{d\mathbf{y}(t)}{dt} \\ &= \phi^\top (\nabla\phi(\mathbf{y})^\top + \nabla\phi(\mathbf{y}))\phi + (\mathbf{y} - \mathbf{y}^*)^\top \phi(\mathbf{y}) \end{aligned} \quad (28.1)$$

با استفاده از لم ۳.۳.۱، نتیجه می‌شود که

$$\phi^\top (\nabla\phi(\mathbf{y})^\top + \nabla\phi(\mathbf{y}))\phi \leq 0, \quad \forall \mathbf{y} \neq \mathbf{y}^* \quad (29.1)$$

علاوه بر این، با توجه به تعریف (۶.۳.۱) و لم (۱.۳.۱) داریم:

$$(\mathbf{y} - \mathbf{y}^*)^\top (\phi(\mathbf{y}) - \phi(\mathbf{y}^*)) = (\mathbf{y} - \mathbf{y}^*)^\top \phi(\mathbf{y}) \leq 0, \quad \forall \mathbf{y} \neq \mathbf{y}^*$$

بنابراین

$$\frac{dE(\mathbf{y}(t))}{dt} \leq 0 \quad (30.1)$$

و این به این معنی است که شبکه عصبی (۲۱.۱) و (۲۲.۱) پایدار لیاپانوف است. $\square$

لم ۴.۳.۱. ۱. برای هر نقطه ابتدایی $y(t_0) = (x(t_0)^T, u(t_0)^T)^T$ یک جواب پیوسته یکتا $y(t) = (x(t)^T, u(t)^T)^T$ برای سیستم (۲۱.۱) و (۲۲.۱) وجود دارد.

۲. فرض کنید $y(t) = (x(t)^T, u(t)^T)^T$ یک مسیر از شبکه عصبی (۲۱.۱) و (۲۲.۱) با نقطه ابتدایی $y(t_0) = (x(t_0)^T, u(t_0)^T)^T$ باشد. در این صورت اگر، $u(t_0) \geq 0$ ، بنابراین $u(t) \geq 0$.

برهان. ۱. از آنجایی که $\nabla f(x)$ و $\nabla g_k(x)$ برای $k = 1, 2, \dots, m$ به طور پیوسته مشتق پذیر روی مجموعه باز $D \subset \mathbb{R}^{n+m}$ هستند بنابراین

$$\left\{ \nabla f(x^*) + \nabla g(x^*)^T (u^* + g(x^*))^+ \right\}$$

و

$$\left\{ (u^* + g(x^*))^+ - u^* \right\}$$

پیوسته لیپشیتس محلی روی D هستند. با توجه به قضیه ۱.۳.۱ شبکه عصبی (۲۱.۱) و (۲۲.۱) یک جواب پیوسته یکتا $y(t), t \in [t_0, \eta]$ برای بعضی $\eta > t_0$ دارد. از اثبات قضیه ۷.۳.۱ می دانیم که E روی t ، یک تابع غیر صعودی است. بنابراین

$$\frac{1}{2} \|y - y^*\|^2 \leq E(y(t)) \leq E(y(t_0)), \quad \forall t \geq t_0 \quad (31.1)$$

این نشان می دهد که مسیر شبکه عصبی (۲۱.۱) و (۲۲.۱) کراندار است. بنابراین $\eta \rightarrow +\infty$.

۲. برای نقطه ابتدایی داده شده $y(t_0)$ با $u(t_0) \geq 0$ داریم:

$$\begin{aligned} \frac{du}{dt} + u &= (u + g(x))^+ \\ \int_{t_0}^t \left(\frac{du}{dt} + u \right) e^s ds &= \int_{t_0}^t e^s (u + g(x))^+ ds \end{aligned}$$

همچنین

$$u(t) = e^{-(t-t_0)} u(t_0) + e^{-t} \int_{t_0}^t e^s (u + g(x))^+ ds$$

در نتیجه با توجه به این که $(u + g(x))^+ \geq 0$ داریم: $u(t) \geq 0$ برای هر $t \geq t_0$.

□

قضیه ۸.۳.۱. مسیرهای شبکه عصبی (۲۱.۱) و (۲۲.۱) برای هر نقطه ابتدایی

$$y_0 = (x(t_0)^T, u(t_0)^T)^T \in \mathbb{R}^{n+n}$$

به نقطه تعادل شبکه عصبی همگرا می شود. به طور خاص، شبکه عصبی (۲۱.۱) و (۲۲.۱) با هر نقطه ابتدایی $y_0 = (x(t_0)^T, u(t_0)^T)^T \in \mathbb{R}^{n+m}$ پایدار مجانبی سراسری است و D^* یک نقطه تعادل یکتا دارد که در آن D^* نشان دهنده مجموعه نقاط بهینه مسئله اصلی و دوگان است.

برهان. از اثبات لم ۴.۳.۱، داریم که مسیر $((\mathbf{x}(t))^T, \mathbf{u}(t)^T)^T$ از شبکه عصبی (۲۱.۱) و (۲۲.۱) کراندار است. بنابراین، یک دنباله صعودی $\{t_k\}$ با $t_k \rightarrow \infty$ برای $k \rightarrow \infty$ و نقطه حدی $(\bar{\mathbf{x}}^T, \bar{\mathbf{u}}^T)^T$ وجود دارد به طوری که:

$$\lim_{k \rightarrow \infty} (\mathbf{x}(t_k)^T, \mathbf{u}(t_k)^T)^T = (\bar{\mathbf{x}}^T, \bar{\mathbf{u}}^T)^T$$

با استفاده از قضیه ۵.۳.۱ داریم:

$$\{(\mathbf{x}(t)^T, \mathbf{u}(t)^T)^T \rightarrow H\}$$

برای $t \rightarrow \infty$ و H بزرگ‌ترین مجموعه ثابت در

$$K = \{(\mathbf{x}(t)^T, \mathbf{u}(t)^T)^T \mid \frac{dE(\mathbf{y}(t))}{dt} = 0\}$$

است.

از (۲۱.۱) و (۲۲.۱) و (۳۰.۱)، نتیجه می‌شود که $\frac{d\mathbf{x}}{dt} = 0$ و $\frac{d\mathbf{u}}{dt} = 0 \Leftrightarrow \frac{dE(\mathbf{y}(t))}{dt} = 0$. بنابراین $(\bar{\mathbf{x}}^T, \bar{\mathbf{u}}^T)^T \in D^*$ برای $H \subseteq K \subseteq D^*$. از طرف دیگر، اثبات می‌کنیم که $(\mathbf{x}(t)^T, \mathbf{u}(t)^T)^T$ همگرایی سراسری به نقطه تعادل $(\bar{\mathbf{x}}(t)^T, \bar{\mathbf{u}}(t)^T)^T$ است. اکنون تابع لیپانوف دیگری تعریف می‌کنیم:

$$\bar{E}(\mathbf{y}) = \|\phi(\mathbf{y})\|^2 + \frac{1}{\gamma} \|\mathbf{y} - \bar{\mathbf{y}}\|^2 \quad (۳۲.۱)$$

با جایگذاری $\mathbf{x}^* = \bar{\mathbf{x}}$ و $\mathbf{u}^* = \bar{\mathbf{u}}$ در (۲۷.۱). نتیجه می‌شود که $\bar{E}(\mathbf{y})$ به طور پیوسته مشتق‌پذیر است و $\bar{E}(\bar{\mathbf{y}}) = 0$. توجه شود که $\lim_{k \rightarrow \infty} (\mathbf{x}(t_k)^T, \mathbf{u}(t_k)^T)^T = (\bar{\mathbf{x}}^T, \bar{\mathbf{u}}^T)^T$. بنابراین داریم

$$\lim_{k \rightarrow \infty} \bar{E}(\mathbf{x}(t_k)^T, \mathbf{u}(t_k)^T)^T = \bar{E}(\bar{\mathbf{x}}^T, \bar{\mathbf{u}}^T)^T$$

پس برای هر $\epsilon > 0$ وجود دارد $q > 0$ به طوری که برای هر $t \geq t_q$ داریم $\bar{E}(\mathbf{y}(t)) < \epsilon$. به طریق مشابه $\frac{d\bar{E}(\mathbf{y}(t))}{dt} \leq 0$ در نتیجه برای $t \geq t_q$

$$\frac{1}{\gamma} \|\mathbf{y}(t) - \bar{\mathbf{y}}\|^2 \leq \bar{E}(\mathbf{y}(t)) \leq \epsilon$$

همچنین $\lim_{t \rightarrow \infty} \|\mathbf{y}(t) - \bar{\mathbf{y}}\| = 0$ و $\lim_{t \rightarrow \infty} \mathbf{y}(t) = \bar{\mathbf{y}}$. بنابراین شبکه (۲۱.۱) و (۲۲.۱) همگرایی سراسری به نقطه تعادل $(\bar{\mathbf{x}}^T, \bar{\mathbf{u}}^T)^T$ است. که در آن $\bar{\mathbf{x}}$ جواب بهینه مسئله (۱۳.۱) و (۱۴.۱) است.

در حالت خاص، اگر $D^* = \{(\mathbf{x}^{*T}, \mathbf{u}^{*T})^T\}$ ، بنابراین مسیر $(\mathbf{x}(t)^T, \mathbf{u}(t)^T)^T$ با هر نقطه ابتدایی $(\mathbf{x}(t_0)^T, \mathbf{u}(t_0)^T)^T$ پایدار مجانبی سراسری به نقطه تعادل $\mathbf{y}^* = (\mathbf{x}^{*T}, \mathbf{u}^{*T})^T$ است.

□

قضیه ۹.۳.۱. سرعت همگرایی شبکه عصبی (۲۱.۱) و (۲۲.۱) با افزایش τ افزایش می‌یابد.

برهان. از (۲۹.۱) و (۳۰.۱) دیده می‌شود که برای هر $\tau > 0$ در (۲۱.۱) و (۲۲.۱) داریم:

$$\frac{dE(\mathbf{y})}{dt} \leq \tau \phi(\mathbf{y})^\top \left(\nabla \phi(\mathbf{y})^\top + \nabla \phi(\mathbf{y}) \right) \phi(\mathbf{y}) \leq 0$$

در نتیجه

$$E(\mathbf{y}(t)) \leq E(\mathbf{y}(t_0)) + \tau \int_{t_0}^t \phi(\mathbf{y}(s))^\top \left(\nabla \phi(\mathbf{y}(s))^\top + \nabla \phi(\mathbf{y}(s)) \right) \phi(\mathbf{y}(s)) ds.$$

از طرفی با توجه به این که $E(\mathbf{y}) \geq \frac{1}{2} \|\mathbf{y} - \mathbf{y}^*\|^2$ و $\mathbf{y}^*$ یک نقطه کروش-کان-تاکر از (۱۳.۱) و (۱۴.۱) است، داریم:

$$\|\mathbf{y}(t) - \mathbf{y}^*\|^2 \leq 2E(\mathbf{y}(t_0)) + 2\tau \int_{t_0}^t \phi(\mathbf{y}(s))^\top \left(\nabla \phi(\mathbf{y}(s))^\top + \nabla \phi(\mathbf{y}(s)) \right) \phi(\mathbf{y}(s)) ds$$

□

پس سرعت همگرایی مسیر $\mathbf{y}(t)$ با افزایش τ افزایش می‌یابد.

با توجه به مطالب ارائه شده در این بخش نتیجه زیر را داریم:

نتیجه ۱.۳.۱. اگر $D^* = \{(\mathbf{x}^{*\top}, \mathbf{u}^{*\top})^\top\}$ بنابراین شبکه عصبی (۲۱.۱) و (۲۲.۱) برای حل (۱۳.۱) و (۱۴.۱) پایدار مجانبی سراسری به نقطه تعادل یکتا $\mathbf{y}^* = (\mathbf{x}^{*\top}, \mathbf{u}^{*\top})^\top$ است.

فصل ۲

مدل شبکه عصبی بازگشتی برای حل مسائل رگرسیون فازی

استفاده از شبکه‌های عصبی برای حل مسائل رگرسیون فازی، اولین بار توسط ایشیبوچی^۱ و تاناکا^۲ برای حل مسائل پیچیده‌تر از مسائل قابل حل با روش کمترین توان‌های دوم ارائه گردید [۶۴]. پس از آن محققان زیادی با استفاده از این روش، به حل مسائل مختلف رگرسیون فازی پرداخته‌اند (برای مثال به [۲۷، ۲۹، ۶۵-۷۰] مراجعه شود). با این حال تا کنون هیچ مدل شبکه عصبی بازگشتی با تحلیل پایداری و قضایای مربوط به همگرایی سراسری برای حل این دست از مسائل ارائه نشده است.

در بیشتر منابع ذکر شده، نویسندگان تلاش کرده‌اند تا با استفاده از یک الگوریتم یادگیری برای بدست آوردن وزن‌های فازی از یک شبکه عصبی پیش‌رونده اقدام کنند. مهمترین ایراد این روش که به طور کلی به حل مسائل رگرسیون خطی، غیرخطی و چندجمله‌ای فازی پرداخته است؛ اشاره نکردن به قضایا و گزاره‌های مربوط به همگرایی و پایداری شبکه عصبی پیش‌رونده بیان شده است. از طرف دیگر، در این روش‌ها برای بدست آوردن وزن‌ها از یک الگوریتم تکراری استفاده شده است که نیازمند محاسبه ماتریس وارون در هر تکرار است. واضح است که هر چه بعد مسئله بالاتر برود پیچیدگی محاسباتی بیشتر می‌شود. علاوه بر این،

^۱Ishibuchi

^۲Tanaka

در روش‌های پیشنهادی همگرایی شبکه عصبی به انتخاب یک نقطه اولیه مناسب بستگی دارد. بنابراین پیشنهاد یک شبکه عصبی مناسب برای حل مسائل رگرسیون فازی بدون نیاز به محاسبه پارامتر جریمه و با تجزیه و تحلیل دقیق پایداری و همگرایی، کاملاً ضروری و معنی دار است.

برای این منظور با توجه به شرایط بهینه سازی کروش-کان-تاکر، یک مدل شبکه عصبی برای حل مسئله برنامه‌ریزی درجه دو معادل با مسئله رگرسیون فازی ارائه می‌شود. ثابت می‌شود که نقطه تعادل شبکه عصبی پیشنهادی معادل با نقطه کروش-کان-تاکر از مسئله برنامه‌ریزی درجه دوم است. وجود و منحصر به فرد بودن نقطه تعادل شبکه نیز مورد تجزیه و تحلیل قرار گرفته است. با ساخت یک تابع مناسب لیپانوف، شرط کافی برای وجود و پایداری مجانبی سراسری نقطه تعادل منحصر به فرد شبکه عصبی به دست می‌آید. همچنین در پایان اثبات می‌شود که شبکه عصبی پیشنهادی همگرایی سراسری است.

۱.۲ یک روش شبکه عصبی پارامتریک برای حل مدل‌های رگرسیون خطی فازی

با توجه به مطالب ارائه شده در ۲.۱ مدل رگرسیون خطی فازی

$$Y_i = A_0 x_{i0} + A_1 x_{i1} + \dots + A_p x_{ip} + \epsilon_i, \quad i = 1, \dots, n \quad (1.2)$$

را در نظر می‌گیریم. در این بخش ما علاقه مند به یافتن ضرایب رگرسیون $A_0, A_1, \dots, A_p$ هستیم به طوری که $\hat{Y}_i$ برای $i = 1, 2, \dots, n$ تا حد امکان نزدیک به Y_i تقریب زده شود. به عبارت دیگر:

$$\text{minimize } \|\hat{Y}_i^\dagger(h) - Y_i^\dagger(h)\|, \quad h \in [0, 1], \quad i = 1, 2, \dots, n \quad (2.2)$$

که در آن $\dagger$ بدان معناست که این گزاره برای حد بالا و پایین با هم و به طور مستقل برقرار است. بنابراین این مسئله به یک مسئله بهینه سازی تبدیل می‌شود.

چندین روش بهینه سازی برای حل مسئله برنامه‌ریزی (۲.۲) وجود دارد که هر یک از این روش‌ها نقاط ضعف و قوت خاص خود را دارند. در اینجا ما به روش شبکه عصبی برای حل مسائل رگرسیون فازی علاقه‌مند هستیم که تفاوت اساسی با روش‌هایی که در گذشته اشاره شده است دارد (برای اطلاعات بیشتر به [۵۶] مراجعه شود).

۱.۱.۲ مدل بهینه سازی

فرض کنید h -برش‌های متغیر پاسخ Y_i برای $i = 1, \dots, n$ به صورت زیر نشان داده شود:

$$Y_i[h] = [\underline{Y}_i(h), \bar{Y}_i(h)], \quad h \in [0, 1]$$

که در آن $\underline{Y}_i(h)$ و $\overline{Y}_i(h)$ به ترتیب نشان دهنده نقاط سمت چپ و نقاط سمت راست $-h$ -برش‌های متغیر پاسخ هستند. تابع خطا که باید برای هر $-h$ -برش کمینه شود به صورت زیر است:

$$C(A_0, A_1, \dots, A_p)(h) = \underline{C}(A_0, A_1, \dots, A_p)(h) + \overline{C}(A_0, A_1, \dots, A_p)(h)$$

که در آن $\underline{C}(A_0, A_1, \dots, A_p)(h)$ نشان دهنده خطا میان نقاط سمت چپ $-h$ -برش‌های متغیر پاسخ و مقدار برآورد شده است و $\overline{C}(A_0, A_1, \dots, A_p)(h)$ نشان دهنده خطا میان نقاط سمت راست $-h$ -برش‌های متغیر پاسخ و مقدار برآورد شده است (برای اطلاعات بیشتر به [۷۱] مراجعه شود).

با توجه به (۲.۲)، $C^\dagger$ به صورت زیر تعریف می‌شود:

$$C^\dagger(A_0, A_1, \dots, A_p)(h) = \|\hat{Y}_i^\dagger(h) - Y_i^\dagger(h)\|, \quad h \in [0, 1], \quad i = 1, 2, \dots, n$$

که در آن

$$\|\hat{Y}_i^\dagger(h) - Y_i^\dagger(h)\| = \frac{1}{\sqrt{p}} \sum_{i=1}^n (\hat{Y}_i^\dagger(h) - Y_i^\dagger(h))^2 = \frac{1}{\sqrt{p}} \sum_{i=1}^n \left(\sum_{j=0}^p x_{ij} A_j^\dagger(h) - Y_i^\dagger(h) \right)^2 \quad (3.2)$$

اکنون با فرض $0 \leq x_{ij}$ برای $i = 1, \dots, n$ و $j = 0, \dots, p$ به یک تابع درجه دوم به شکل زیر تبدیل می‌شود:

$$\begin{aligned} C^\dagger(A_0, A_1, \dots, A_p)(h) &= \frac{1}{\sqrt{p}} \sum_{i=1}^n \left(\sum_{j=0}^p x_{ij} A_j^\dagger(h) - Y_i^\dagger(h) \right)^2 \\ &= \frac{1}{\sqrt{p}} \sum_{i=1}^n \left[\left(\sum_{j=0}^p x_{ij} A_j^\dagger(h) \right)^2 - 2 Y_i^\dagger(h) \sum_{j=0}^p x_{ij} A_j^\dagger(h) + (Y_i^\dagger(h))^2 \right] \\ &= \frac{1}{\sqrt{p}} (\mathbf{A}^\dagger(h))^T \mathbf{Q} \mathbf{A}^\dagger(h) + (\mathbf{B}^\dagger(h))^T \mathbf{A}^\dagger(h) + M^\dagger(h) \end{aligned}$$

که در آن

$$\mathbf{Q} = \begin{bmatrix} n & \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i2} & \dots & \sum_{i=1}^n x_{ip} \\ \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1}^2 & \sum_{i=1}^n x_{i1} x_{i2} & \dots & \sum_{i=1}^n x_{i1} x_{ip} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^n x_{ip} & \sum_{i=1}^n x_{ip} x_{i1} & \sum_{i=1}^n x_{ip} x_{i2} & \dots & \sum_{i=1}^n x_{ip}^2 \end{bmatrix} \quad (4.2)$$

$$\mathbf{A}^\dagger(h) = (A_0^\dagger(h), A_1^\dagger(h), \dots, A_p^\dagger(h))^T$$

$$\mathbf{B}^\dagger(h) = (B_0^\dagger(h), B_1^\dagger(h), \dots, B_p^\dagger(h))^T$$

$$M^\dagger(h) = \frac{1}{\sqrt{p}} \sum_{i=1}^n (Y_i^\dagger(h))^2$$

$$B_j^\dagger(h) = -\sum_{i=1}^p x_{ij} Y_i^\dagger(h) \text{ و}$$

با توجه به مطالب ارائه شده در پیوست آ، برای اعداد فازی A_j که $j = 0, 1, \dots, p$ داریم:

$$\underline{A}_j(h) \leq \overline{A}_j(h), \quad h \in [0, 1]$$

بنابراین مدل بهینه سازی درجه دوم محدب زیر به دست می آید:

$$\text{minimize } C(\mathbf{A})(h) = \frac{1}{2}(\mathbf{A}(h))^T \mathbf{N} \mathbf{A}(h) + (\mathbf{B}(h))^T \mathbf{A}(h) + M(h) \quad (5.2)$$

subject to

$$\mathbf{G} \mathbf{A}(h) \leq 0, \quad h \in [0, 1] \quad (6.2)$$

که در آن

$$\mathbf{A}(h) = (\underline{A}_0(h), \underline{A}_1(h), \dots, \underline{A}_p(h), \overline{A}_0(h), \overline{A}_1(h), \dots, \overline{A}_p(h))^T \quad (7.2)$$

$$\mathbf{B}(h) = (\underline{B}_0(h), \underline{B}_1(h), \dots, \underline{B}_p(h), \overline{B}_0(h), \overline{B}_1(h), \dots, \overline{B}_p(h))^T \quad (8.2)$$

$$M(h) = (\underline{M}(h) + \overline{M}(h)) \quad (9.2)$$

$$\mathbf{G} = \begin{bmatrix} -1 & \dots & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & -1 & 0 & \dots & 1 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix}_{(2p+2) \times (2p+2)} \quad (10.2)$$

$$\mathbf{N} = \begin{bmatrix} \mathbf{Q} & \mathbf{O} \\ \mathbf{O} & \mathbf{Q} \end{bmatrix}_{(2p+2) \times (2p+2)}$$

$$\underline{B}_j(h) = -\sum_{i=1}^n x_{ij} \underline{Y}_i(h)$$

$$\overline{B}_j(h) = -\sum_{i=1}^n x_{ij} \overline{Y}_i(h)$$

در اینجا $\mathbf{O}$ ماتریس $(p+1) \times (p+1)$ صفر می باشد و $\mathbf{Q}$ در (۴.۲) تعریف شده است.

دیده می شود که هسین ماتریس $C(\mathbf{A})(h)$ برابر است با $\mathbf{N}$ $\nabla^2 C(\mathbf{A})(h) = \nabla(\nabla C(\mathbf{A})(h)) = \mathbf{N}$ زیرا $\mathbf{Q}$ یک ماتریس معین مثبت است.

۲.۱.۲ مدل رگرسیون چندجمله‌ای فازی با یک متغیر و از درجه کامل

در مدل رگرسیون چندجمله‌ای

$$Y_i = A_0 x_{i0} + \sum_{j=1}^p A_j x_{ij} + \sum_{j=1}^p \sum_{l=1}^p A_{jl} x_{ij} x_{il} + \sum_{j=1}^p \sum_{l=1}^p \sum_{s=1}^p A_{jls} x_{ij} x_{il} x_{is} + \dots \quad (11.2)$$

که در ۵.۱.۱ به معرفی آن پرداختیم، فرض کنید ضرایب فازی هستند و تنها یک متغیر داریم. در این صورت مدل رگرسیون چندجمله‌ای فازی زیر به دست می‌آید:

$$\hat{Y}_i = A_0 x_{i0} + A_1 x_{i1} + A_2 x_{i2} + \dots + A_p x_{ip} \quad (12.2)$$

با روشی مشابه آنچه در ۱.۱.۲ به کار برده شد، و با فرض $x_i^j \leq 1$ و $x_i^0 = 1$ برای $i = 1, \dots, n$ و $j = 0, \dots, p$ مسئله درجه دوم مقید (۵.۲) و (۶.۲) به دست می‌آید که در آن

$$N = \begin{bmatrix} Q & O \\ O & Q \end{bmatrix}_{(2p+2) \times (2p+2)}$$

$$Q = \begin{bmatrix} n & \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \dots & \sum_{i=1}^n x_i^p \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 & \dots & \sum_{i=1}^n x_i^p x_i \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^n x_i^p & \sum_{i=1}^n x_i^p x_i & \sum_{i=1}^n x_i^p x_i^2 & \dots & \sum_{i=1}^n x_i^p x_i^p \end{bmatrix}$$

$$B(h) = (\underline{B}_0(h), \underline{B}_1(h), \dots, \underline{B}_p(h), \overline{B}_0(h), \overline{B}_1(h), \dots, \overline{B}_p(h))^T$$

$$\underline{B}_j(h) = -\sum_{i=1}^n x_i^j \underline{Y}_i(h)$$

$$\overline{B}_j(h) = -\sum_{i=1}^n x_i^j \overline{Y}_i(h)$$

و $A(h)$ ، $M(h)$ و G به ترتیب در (۷.۲)، (۹.۲) و (۱۰.۲) معرفی شده‌اند. هسین ماتریس $C(A)(h)$ برابر است با N بنابراین یک ماتریس معین مثبت است.

۳.۱.۲ مدل رگرسیون چندجمله‌ای فازی با دو متغیر و از درجه دوم

اگر فرض شود که در مدل

$$Y_i = A_0 + \sum_{j=1}^p A_j x_{ij} + \sum_{j=1}^p \sum_{l=1}^p A_{jl} x_{ij} x_{il} + \sum_{j=1}^p \sum_{l=1}^p \sum_{s=1}^p A_{jls} x_{ij} x_{il} x_{is} + \dots \quad (13.2)$$

ارائه شده در ۵.۱.۱ ضرایب فازی باشند، تنها دو متغیر داشته باشیم و مدل از مرتبه دوم باشد، بنابراین داریم:

$$\hat{Y}_i = A_0 x_{i0} + A_1 x_{i1} + A_2 x_{i2} + A_{11} x_{i1}^2 + A_{22} x_{i2}^2 + A_{12} x_{i1} x_{i2} \quad (14.2)$$

توجه داشته باشید که مدل رگرسیون (۱۴.۲) شامل اجزای خطی و درجه دوم جداگانه برای هر یک از دو متغیر پیش بینی کننده و یک ضرب برداری است. مانند قبل هدف پیدا کردن ضرایب $A_0, A_1, A_2, A_{11}, A_{22}, A_{12}$ از (۱۴.۲) است به طوری که $\hat{Y}_i$ بهترین تقریب برای Y_i باشد. مشابه قبل با فرض $0 \leq x_{i1}, x_{i2}$ و $x_{i0} = 1$ برای $i = 1, \dots, n$

$$\mathbf{A}(h) = (\underline{A}_0(h), \underline{A}_1(h), \underline{A}_2(h), \underline{A}_{11}(h), \underline{A}_{22}(h), \underline{A}_{12}(h), \overline{A}_0(h), \overline{A}_1(h), \overline{A}_2(h), \overline{A}_{11}(h), \overline{A}_{22}(h), \overline{A}_{12}(h))^T$$

به مسئله بهینه سازی درجه دوم محدب (۵.۲) و (۶.۲) می‌رسیم که در آن

$$\mathbf{N} = \begin{bmatrix} \mathbf{Q} & \mathbf{O} \\ \mathbf{O} & \mathbf{Q} \end{bmatrix}_{12 \times 12}$$

$$\mathbf{Q} = \begin{bmatrix} n & \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i2} & \sum_{i=1}^n x_{i1}^2 & \sum_{i=1}^n x_{i2}^2 & \sum_{i=1}^n x_{i1} x_{i2} \\ \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1}^2 & \sum_{i=1}^n x_{i1} x_{i2} & \sum_{i=1}^n x_{i1}^3 & \sum_{i=1}^n x_{i1}^2 x_{i2} & \sum_{i=1}^n x_{i1}^2 x_{i2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^n x_{i1} x_{i2} & \sum_{i=1}^n x_{i1}^2 x_{i2} & \sum_{i=1}^n x_{i1} x_{i2}^2 & \sum_{i=1}^n x_{i1}^3 x_{i2} & \sum_{i=1}^n x_{i1}^2 x_{i2}^2 & \sum_{i=1}^n x_{i1}^2 x_{i2}^2 \end{bmatrix}$$

$$\mathbf{B}(h) = - \left(\sum_{i=1}^n x_{i0} Y_i(h), \sum_{i=1}^n x_{i1} Y_i(h), \sum_{i=1}^n x_{i2} Y_i(h), \sum_{i=1}^n x_{i1}^2 Y_i(h), \sum_{i=1}^n x_{i2}^2 Y_i(h), \sum_{i=1}^n x_{i1} x_{i2} Y_i(h), \sum_{i=1}^n x_{i0} \bar{Y}_i(h), \sum_{i=1}^n x_{i1} \bar{Y}_i(h), \sum_{i=1}^n x_{i2} \bar{Y}_i(h), \sum_{i=1}^n x_{i1}^2 \bar{Y}_i(h), \sum_{i=1}^n x_{i2}^2 \bar{Y}_i(h), \sum_{i=1}^n x_{i1} x_{i2} \bar{Y}_i(h) \right)^T$$

و $M(h)$ و G مشابه تعاریف ارائه شده در (۹.۲) و (۱۰.۲) می‌باشند.

مشابه قبل واضح است که هسین ماتریس $C(\mathbf{A})(h)$ معین مثبت است.

۴.۱.۲ رویکرد شبکه عصبی بازگشتی

مدل (۵.۲) و (۶.۲)، یک مسئله بهینه سازی درجه دوم محدب است. بنابراین مطابق با مطالب ارائه شده در ۴.۳.۱ در این فصل نیز تلاش می‌کنیم که با استفاده از روش‌های بهینه سازی مدل درجه دوم (۵.۲) و (۶.۲) را به یک سیستم غیرخطی پویا تبدیل کنیم. فرض کنید $\mathbf{A}(h)(.)$ و $\mathbf{U}(.)$ متغیرهای وابسته به زمان باشند. هدف ما طراحی یک مدل پویا است که به نقطه تعادل برسد. بنابراین برای هر h -برش مدل شبکه عصبی بازگشتی زیر را ارائه می‌دهیم. به طوری که مسئله بهینه سازی (۵.۲) و (۶.۲) را حل کند.

$$\begin{cases} \frac{d\mathbf{A}(h)}{dt} = -\tau [\mathbf{N}\mathbf{A}(h) + \mathbf{B}(h) + \mathbf{G}^T (\mathbf{U} + \mathbf{G}\mathbf{A}(h))^+] & h \in [0, 1] & (15.2) \\ \frac{d\mathbf{U}}{dt} = -\tau [(\mathbf{U} + \mathbf{G}\mathbf{A}(h))^+ - \mathbf{U}] & & (16.2) \\ \mathbf{A}(0) = \mathbf{A}_0, & & (17.2) \end{cases}$$

همانطور که در ۴.۳.۱ اشاره شد، $\tau > 0$ یک عامل تنظیم کننده سرعت همگرایی مدل (۱۵.۲)–(۱۷.۲) است.

از ۴.۳.۱ می‌دانیم که $A^*(h) \in \mathbb{R}^{2p+2}$ برای $h \in [0, 1]$ جواب بهینه (۵.۲) و (۶.۲) است اگر و تنها اگر وجود داشته باشد $U^* \in \mathbb{R}^{2p+2}$ به طوری که $(A^*(h)^T, U^{*T})^T$ در شرایط کروش–کان–تاکر زیر صدق کند:

$$\begin{cases} U^* \geq 0, \quad GA^*(h) \leq 0, \quad U^{*T}GA^*(h) = 0 \\ NA^*(h) + B(h) + G^T U^* = 0, \quad h \in [0, 1] \end{cases} \quad (18.2)$$

که در آن $A^*(h)$ نقطه کروش–کان–تاکر مدل (۵.۲) و (۶.۲) و U^* بردار ضریب لاگراژ متناظر با $A^*(h)$ است.

اکنون به ازای هر $h \in [0, 1]$ مدل شبکه عصبی پیشنهادی زیر را در نظر می‌گیریم:

$$\begin{cases} \frac{dy}{dt} = \tau \phi(y) \\ y(t_0) = y_0 = ((A(t_0)(h))^T, U(t_0)^T)^T \in \mathbb{R}^{4p+4}, \quad \tau > 0, \end{cases} \quad (19.2) \quad (20.2)$$

که در آن

$$\phi(y) = \begin{pmatrix} - (NA(h) + B(h) + G^T(U + GA(h)))^+ \\ (U + GA(h))^+ - U \end{pmatrix} \quad (21.2)$$

برای راحتی فرض شده $\tau = 1$.

بررسی همگرایی و پایداری

در این قسمت به بررسی گزاره‌های مربوط به همگرایی و پایداری مدل (۱۹.۲) و (۲۰.۲) می‌پردازیم.

قضیه ۱.۱.۲. فرض کنید $y^* = (A^*(h)^T, U^{*T})^T$ ، برای $h \in [0, 1]$ نقطه تعادل مدل شبکه عصبی (۱۹.۲) و (۲۰.۲) باشد. بنابراین $A^*(h)$ یک نقطه کروش–کان–تاکر برای مسئله (۵.۲) و (۶.۲) است. از سمت دیگر، اگر $A^*(h) \in \mathbb{R}^{2p+2}$ جواب بهینه مسئله (۵.۲) و (۶.۲) باشد؛ بنابراین وجود دارد $U^* \in \mathbb{R}^{2p+2}$ به طوری که $y^* = (A^*(h)^T, U^{*T})^T$ نقطه تعادل مدل شبکه عصبی (۱۹.۲) و (۲۰.۲) است.

برهان. فرض کنید $y^* = (A^*(h)^T, U^{*T})^T$ برای $h \in [0, 1]$ نقطه تعادل مدل (۱۹.۲) و (۲۰.۲) باشد. بنابراین $\frac{dA^*(h)}{dt} = 0$ و $\frac{dU^*}{dt} = 0$. به راحتی نتیجه می‌شود که:

$$NA^*(h) + B(h) + G^T(U^* + GA^*(h))^+ = 0 \quad (22.2)$$

$$(U^* + GA^*(h))^+ - U^* = 0 \quad (23.2)$$

با استفاده از لم ۲.۳.۱ و با جایگذاری (۲۳.۲) در (۲۲.۲) دیده می‌شود که $\mathbf{y}^* = (\mathbf{A}^*(h)^\top, \mathbf{U}^{*\top})^\top$ در شرایط کروش-کان-تاکر (۱۸.۲) صدق می‌کند.

□

اثبات حالت برگشت سر راست است.

لم ۱.۱.۲. ژاکوبین ماتریس ϕ که در (۲۱.۲) تعریف شده است، یک ماتریس نیمه معین منفی است.

برهان. بدون کاستن از کلیت، فرض کنید وجود دارد $0 < m' < (2p+2)$ به طوری که:

$$(\mathbf{U} + \mathbf{GA}(h))^+ = \left((\mathbf{U} + \mathbf{GA}(h))_1, (\mathbf{U} + \mathbf{GA}(h))_2, \dots, (\mathbf{U} + \mathbf{GA}(h))_{m'}, \underbrace{0, 0, \dots, 0}_{(2p+2)-m'} \right)^\top$$

با روشی مشابه آنچه در لم ۳.۳.۱ انجام شد، داریم:

$$\nabla \phi(\mathbf{y}) = \begin{pmatrix} -(\mathbf{N} + (\mathbf{G}^*)^\top \mathbf{G}^*) & -(\mathbf{G}^*)^\top \\ \mathbf{G}^* & \mathcal{W}_{(2p+2) \times (2p+2)} \end{pmatrix}$$

که در آن

$$\mathbf{G}^* = \begin{pmatrix} \mathbf{R}_{m' \times (2p+2)} \\ \mathbf{O}_{(2p+2-m') \times (2p+2)} \end{pmatrix} = \begin{pmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \\ \dots \\ \dots \\ \mathbf{G}_{m'} \\ \mathbf{O}_{1 \times (2p+2)} \\ \mathbf{O}_{1 \times (2p+2)} \\ \dots \\ \dots \\ \mathbf{O}_{1 \times (2p+2)} \end{pmatrix}$$

$$\mathcal{W}_{(2p+2) \times (2p+2)} = \begin{pmatrix} \mathbf{O}_{m' \times m'} & \mathbf{O}_{m' \times (2p+2-m')} \\ \mathbf{O}_{(2p+2-m') \times m'} & -I_{(2p+2-m') \times (2p+2-m')} \end{pmatrix}$$

و $\mathbf{O}$ ماتریس صفر است. از نکات بیان شده نتیجه می‌شود که $\nabla \phi(\mathbf{y})$ یک ماتریس نیمه معین منفی است.

اگر $m' = 2p+2$ آن‌گاه،

$$(\mathbf{U} + \mathbf{GA}(h))^+ = \left((\mathbf{U} + \mathbf{GA}(h))_1, (\mathbf{U} + \mathbf{GA}(h))_2, \dots, (\mathbf{U} + \mathbf{GA}(h))_{2p+2} \right)^\top$$

بنابراین:

$$\nabla\phi(\mathbf{y}) = \begin{pmatrix} -(\mathbf{N} + \mathbf{G}^\top \mathbf{G}) & -\mathbf{G}^\top \\ \mathbf{G} & \mathbf{O}_{(2p+2) \times (2p+2)} \end{pmatrix}$$

مشابه با حالت قبلی، اثبات می‌شود که $\nabla\phi(\mathbf{y})$ یک ماتریس نیمه معین منفی است. نهایتاً اگر $m' = 0$ ، آن گاه $(\mathbf{U} + \mathbf{GA}(h))^+ = (\underbrace{0, 0, \dots, 0}_{2p+2})^\top$ بنابرین داریم:

$$\nabla\phi(\mathbf{y}) = \begin{pmatrix} -\mathbf{N} & \mathbf{O}_{(2p+2) \times (2p+2)} \\ \mathbf{O}_{(2p+2) \times (2p+2)} & -I_{(2p+2) \times (2p+2)} \end{pmatrix}.$$

در این حالت نیز نتیجه می‌شود که $\nabla\phi(\mathbf{y})$ یک ماتریس نیمه معین منفی است و این اثبات را $\square$ کامل می‌کند.

قضیه ۲.۱.۲. تحت مفروضات لم ۱.۱.۲ شبکه عصبی (۱۹.۲) و (۲۰.۲) پایدار لیپانوف است.

برهان. تابع لیپانوف $E: \mathbb{R}^{p+4} \rightarrow \mathbb{R}$ را به صورت زیر در نظر بگیرید:

$$E(\mathbf{y}) = E_1(\mathbf{y}) + E_2(\mathbf{y}) \quad (24.2)$$

که در آن $E_1(\mathbf{y}) = \|\phi(\mathbf{y})\|^2$ و $E_2(\mathbf{y}) = \frac{1}{\epsilon} \|\mathbf{y} - \mathbf{y}^*\|^2$. با روشی مشابه با روش ارائه شده در لم ۷.۳.۱ نتیجه می‌شود شبکه عصبی (۱۹.۲) و (۲۰.۲) پایدار لیپانوف است. $\square$

لم ۲.۱.۲. (i) برای هر نقطه ابتدایی $\mathbf{y}(t_0) = ((\mathbf{A}(t_0)(h))^\top, \mathbf{U}(t_0)^\top)^\top$ که $h \in [0, 1]$ ، یک جواب پیوسته یکتا $\mathbf{y}(t) = ((\mathbf{A}(t)(h))^\top, \mathbf{U}(t)^\top)^\top$ برای سیستم (۱۹.۲) و (۲۰.۲) وجود دارد. (ii) فرض کنید $\mathbf{y}(t) = ((\mathbf{A}(t)(h))^\top, \mathbf{U}(t)^\top)^\top$ برای $h \in [0, 1]$ ، یک مسیر از شبکه عصبی (۱۹.۲) و (۲۰.۲) با نقطه ابتدایی $\mathbf{y}(t_0) = ((\mathbf{A}(t_0)(h))^\top, \mathbf{U}(t_0)^\top)^\top$ باشد. در این صورت، اگر $\mathbf{U}(t_0) \geq 0$ آنگاه $\mathbf{U}(t) \geq 0$ ، برای هر $t \geq t_0$.

برهان. (i) با توجه به برهان لم ۴.۳.۱ واضح است که $\{\mathbf{NA}(h) + \mathbf{B}(h) + \mathbf{G}^\top (\mathbf{U} + \mathbf{GA}(h))^+\}$ و $\{(\mathbf{U} + \mathbf{GA}(h))^+ - \mathbf{U}\}$ پیوسته لیپشیتس^۳ محلی روی مجموعه محدب باز $D \subseteq \mathbb{R}^{p+4}$ می‌باشند. با رویکردی مشابه آنچه در اثبات لم ۴.۳.۱ دیدیم، نتیجه می‌گیریم که نشان می‌دهد مسیر شبکه عصبی (۱۹.۲) و (۲۰.۲) کراندار است. بنابراین $\eta \rightarrow +\infty$. $\square$ (ii) با توجه به برهان لم ۴.۳.۱ نتیجه واضح است.

قضیه ۳.۱.۲. مسیرهای شبکه عصبی (۱۹.۲) و (۲۰.۲) برای هر نقطه ابتدایی

$$\mathbf{y}_0 = ((\mathbf{A}(t_0)(h))^\top, \mathbf{U}(t_0)^\top)^\top \in \mathbb{R}^{p+4} \quad h \in [0, 1]$$

^۳Lipschitz

به نقطه تعادل شبکه عصبی همگرا می‌شوند. به طور به خاص، شبکه عصبی (۱۹.۲) و (۲۰.۲) با هر نقطه ابتدایی $y_0 = (A(t_0)(h))^T, U(t_0)^T)^T \in \mathbb{R}^{p+q}$ پایدار مجانبی سراسری است و D^* یک نقطه تعادل یکتا دارد. که در آن D^* نشان دهنده مجموعه نقاط بهینه مسئله اصلی و دوگان آن است.

برهان. از اثبات لم ۲.۱.۲، داریم که مسیر $((A(t)(h))^T, U(t)^T)^T$ از شبکه عصبی (۱۹.۲) و (۲۰.۲) کراندار است. بنابراین یک دنباله صعودی $\{t_k\}$ با $t_k \rightarrow \infty$ که $k \rightarrow \infty$ و نقطه حدی $(\bar{A}(h)^T, \bar{U}^T)^T$ وجود دارد به طوری که:

$$\lim_{k \rightarrow \infty} ((A(t_k)(h))^T, U(t_k)^T)^T = (\bar{A}(h)^T, \bar{U}^T)^T$$

با استفاده از قضیه ۵.۳.۱ داریم: $\{((A(t)(h))^T, U(t)^T)^T \rightarrow M\}$ که $t \rightarrow \infty$ و M بزرگ‌ترین مجموعه ثابت در $K = \{((A(t)(h))^T, U(t)^T)^T | \frac{dE(y(t))}{dt} = 0\}$ است. از (۱۹.۲) و (۲۰.۲) نتیجه می‌شود که $\frac{dE(y(t))}{dt} = 0 \Leftrightarrow \frac{dU}{dt} = 0$ و $\frac{dA(h)}{dt} = 0$. بنابراین $(\bar{A}(h)^T, \bar{U}^T)^T \in D^*$. $M \subseteq K \subseteq D^*$

از طرف دیگر، اثبات کردیم که مسیر $((A(t)(h))^T, U(t)^T)^T$ همگرای سراسری به نقطه تعادل $(\bar{A}(h)^T, \bar{U}^T)^T$ است. اکنون تابع لیاپانوف دیگری تعریف می‌کنیم:

$$\bar{E}(y) = \|\phi(y)\|^2 + \frac{1}{\gamma} \|y - \bar{y}\|^2 \quad (25.2)$$

با جایگذاری $A^*(h) = \bar{A}(h)$ و $U^* = \bar{U}$ در (۲۴.۲) نتیجه می‌شود که $\bar{E}(y)$ به طور پیوسته مشتق پذیر است و $\bar{E}(\bar{y}) = 0$. توجه شود که $\lim_{k \rightarrow \infty} (A(t_k)(h))^T = \bar{A}(h)^T$ و $\lim_{k \rightarrow \infty} (U(t_k)^T)^T = \bar{U}^T$. بنابراین داریم $\lim_{k \rightarrow \infty} \bar{E}((A(t_k)(h))^T, U(t_k)^T)^T = \bar{E}(\bar{A}(h)^T, \bar{U}^T)^T$. پس برای هر $\epsilon > 0$ وجود دارد $q > 0$ به طوری که برای هر $t \geq t_q$ داریم $\bar{E}(y(t)) < \epsilon$. به طریق مشابه $\frac{d\bar{E}(y(t))}{dt} \leq 0$. در نتیجه برای $t \geq t_q$

$$\frac{1}{\gamma} \|y(t) - \bar{y}\|^2 \leq \bar{E}(y(t)) \leq \epsilon$$

همچنین $\lim_{t \rightarrow \infty} y(t) = \bar{y}$ و $\lim_{t \rightarrow \infty} \|y(t) - \bar{y}\| = 0$. بنابراین شبکه (۱۹.۲) و (۲۰.۲) همگرای سراسری به نقطه تعادل $(\bar{A}(h)^T, \bar{U}^T)^T$ است. که در آن $\bar{A}(h)$ جواب بهینه مسئله (۵.۲) و (۶.۲) است. در حالت خاص، اگر $D^* = \{(A^*(h)^T, U^{*T})^T\}$ ، بنابراین مسیر $((A(t)(h))^T, U(t)^T)^T$ با هر نقطه ابتدایی $((A(t_0)(h))^T, U(t_0)^T)^T$ پایدار مجانبی سراسری به نقطه تعادل $y^* = (A^*(h)^T, U^{*T})^T$ است. $\square$

قضیه ۴.۱.۲. سرعت همگرایی شبکه عصبی (۱۹.۲) و (۲۰.۲) با افزایش τ افزایش می‌یابد.

$\square$

برهان. به برهان قضیه ۹.۳.۱ مراجعه شود.

۵.۱.۲ مثال‌های عددی

در این قسمت برای نشان دادن کارایی روش پیشنهادی به حل چند مثال معروف می‌پردازیم. در بعضی از مثال‌ها، درستی روش پیشنهادی را با چندین نقطه شروع مختلف آزمایش می‌کنیم. همچنین برای تاکید بیشتر بر مدل پیشنهادی، مقایسه‌هایی با مدل‌های موجود در [۲۹، ۶۵، ۷۴-۷۸] با استفاده از فاصله‌های فازی D_1 ، D_2 و D_3 انجام شده است.

ملاحظه ۱.۱.۲. در این بخش اعداد فازی به فرم $A = (a_l, a_m, a_r)_T$ نشان داده شده‌اند. که در آن a_m نشان دهنده مرکز و a_l و a_r به ترتیب نشان دهنده نقاط انتهایی سمت چپ و راست عدد فازی می‌باشند.

مثال ۱.۱.۲. داده‌های فازی ارائه شده در جدول ۱.۲ را برای یک متغیر وابسته فازی Y و دو متغیر غیر فازی x_1 و x_2 در نظر بگیرید.

جدول ۱.۲: مقادیر مشاهده شده و خطای به‌دست آمده در مثال ۱.۱.۲

خطای به‌دست آمده								مقادیر مشاهده شده		
مدل (۱۹.۲) و (۲۰.۲)				مدل ارائه شده در [۲۹]						
S_{UI}	D_3	D_2	D_1	S_{UI}	D_3	D_2	D_1	Y	x_2	x_1
۰/۹۷۶	۰/۰۰۱	۰/۱۵۱	۰/۰۴۲	۰/۹۷۸	۰/۰۰۱	۰/۱۵۰	۰/۰۳۵	$(-2, 2, 7)_T$	۲	۱
۰/۹۱۷۵	۰/۰۹۶	۱/۱۶۳	۱/۸۱۲	۰/۹۱۸	۰/۰۹۱	۱/۱۶۴	۱/۸۰۶	$(-3, 3, 11)_T$	۳	۲
۰/۷۷۹	۱/۲۷۹	۳/۸۵۱	۲۳/۲۶۷	۰/۷۷۷	۱/۳۰۶	۳/۸۸۱	۲۳/۵۴۹	$(-3, 5, 12)_T$	۴	۴
۰/۸۵۵	۰/۹۹۷	۳/۳۲۰	۱۷/۴۲۵	۰/۸۵۷	۰/۹۶۷	۳/۳۲۳	۱۷/۳۷۲	$(-4, 7, 21)_T$	۵	۵
۰/۹۸۳	۰/۰۱۹۸	۰/۵۰۹	۰/۴۱۵	۰/۹۸۰	۰/۰۲۵۷	۰/۵۶۰	۰/۴۷۱	$(-7, 6, 23)_T$	۸	۶
۰/۴۷۹				۰/۴۷۸				MSPE		
۱/۷۹۹				۱/۸۱۶				MSM		
۰/۹۰۲				۰/۹۰۲						

با استفاده از مدل (۱۹.۲) و (۲۰.۲) به حل تقریبی مدل رگرسیونی زیر می‌پردازیم.

$$Y = A_0 + A_1 x_1 + A_2 x_2 \quad (26.2)$$

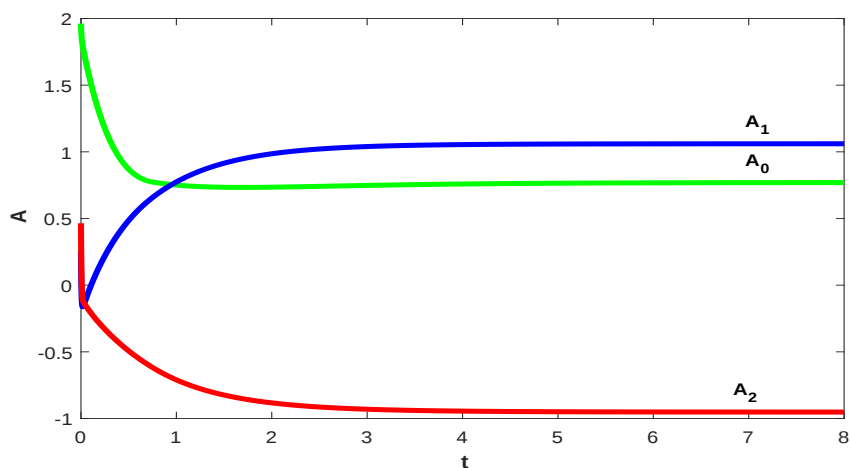
شکل ۱.۲ نشان می‌دهد که مسیرهای شبکه عصبی (۱۹.۲) و (۲۰.۲) به جواب بهینه مسئله (۵.۲) و (۶.۲)، همگرا است. شکل ۲.۲ مسیرهای شبکه عصبی (۱۹.۲) و (۲۰.۲) برای تقریب ضریب A از مدل رگرسیون با 20 نقطه شروع متفاوت را نشان می‌دهد. تقریب ضرایب A_0 ، A_1 و A_2 از مدل رگرسیون خطی فازی (۲۶.۲) با ۱۱ مقدار مختلف از h در شکل ۳.۲ و نتایج عددی در جدول ۲.۲ نشان داده شده است.

بنابراین مدل رگرسیون خطی زیر را داریم:

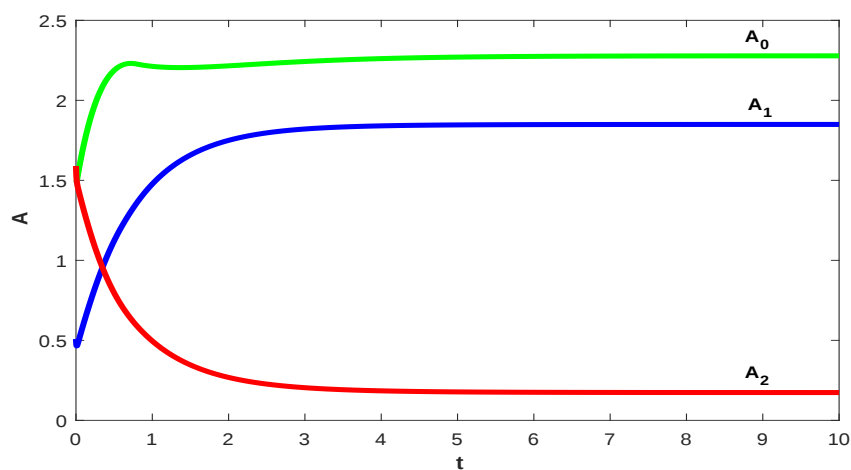
$$\begin{aligned} \hat{Y} = & (-0/1466, 1/6847, 2/8696)_T + (0/4477, 1/6714, 2/0247)_{Tx_1} \\ & + (-1/1965, -0/7106, 1/0544)_{Tx_2} \end{aligned}$$

نتایج نشان می‌دهند که مسیرهای شبکه عصبی (۱۹.۲) و (۲۰.۲) با هر نقطه شروع، همواره به جواب بهینه مسئله (۵.۲) و (۶.۲)، متناظر با این مثال همگرا می‌باشند.

۴۰ مدل شبکه عصبی بازگشتی برای حل مسائل رگرسیون فازی

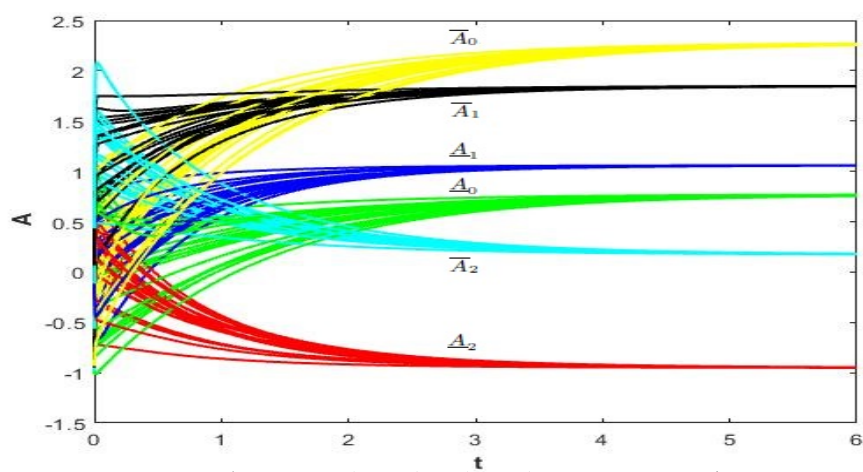


(آ) مسیرهای $\underline{A}_0$ ، $\underline{A}_1$ و $\underline{A}_2$



(ب) مسیرهای $\bar{A}_0$ ، $\bar{A}_1$ و $\bar{A}_2$

شکل ۱.۲: رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با $h = 5\%$ در مثال ۱.۱.۲.



شکل ۲.۲: رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با 20° نقطه شروع متفاوت و $h = 5\%$ در مثال ۱.۱.۲.

جدول ۲.۲: تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۱.۱.۲

$A_2[h]$	$A_1[h]$	$A_0[h]$	h
$[-1/1965, 1/544]$	$[0/4477, 2/247]$	$[-0/1466, 2/8696]$	۰
$[-1/1476, 0/8784]$	$[0/5703, 1/9897]$	$[0/0366, 2/7513]$	۰/۱
$[-1/0986, 0/7022]$	$[0/6929, 1/9546]$	$[0/2198, 2/6329]$	۰/۲
$[-1/0497, 0/5261]$	$[0/8154, 1/9195]$	$[0/4030, 2/5145]$	۰/۳
$[-1/0007, 0/3500]$	$[0/9380, 1/8844]$	$[0/5863, 2/3961]$	۰/۴
$[-0/9517, 0/1738]$	$[1/0606, 1/8492]$	$[0/7695, 2/2777]$	۰/۵
$[-0/9028, -0/0023]$	$[1/1832, 1/8141]$	$[0/9527, 2/1593]$	۰/۶
$[-0/8538, -0/1785]$	$[1/3058, 1/7789]$	$[1/1359, 2/0409]$	۰/۷
$[-0/8048, -0/3546]$	$[1/4284, 1/7438]$	$[1/3192, 1/9225]$	۰/۸
$[-0/7558, -0/5308]$	$[1/5510, 1/7086]$	$[1/5025, 1/8039]$	۰/۹
$[-0/7106, -0/7132]$	$[1/6714, 1/6756]$	$[1/6847, 1/6866]$	۱

برای بررسی کارایی مدل پیشنهادی، این مثال با شبکه عصبی ارائه شده در [۲۹] (مثال ۶.۱) و با ۱۵ تکرار از الگوریتم یادگیری حل شده است.

خطای محاسبه شده برای این روش در مقایسه با خطای به دست آمده از روش پیشنهادی در جدول ۱.۲ نشان داده شده است. واضح است که خطای به دست آمده از روش پیشنهادی، کمتر از خطای به دست آمده برای روش ارائه شده در است. علاوه بر این، شکل نشان می‌دهد که شبکه عصبی پیشنهادی سازگاری بیشتری با مقادیر مشاهده شده دارد.

مثال ۲.۱.۲. نویسندگان در [۲۹]، [۷۶] و [۷۷] از داده‌های ارائه شده در جدول ۳.۲ برای نشان دادن روش پیشنهادیشان استفاده کرده‌اند.

جدول ۳.۲: مقادیر مشاهده شده در مثال ۲.۱.۲

x	۱	۲	۳	۴	۵
Y	$(6/2, 8/0, 9/8)_T$	$(4/2, 6/4, 8/6)_T$	$(6/9, 9/5, 12/1)_T$	$(10/9, 13/5, 16/1)_T$	$(10/6, 13/0, 15/4)_T$

مدل رگرسیون زیر را در نظر بگیرید:

$$Y = A_0 + A_1 x \quad (27.2)$$

با استفاده از مدل (۱۹.۲) و (۲۰.۲)، ضرایب مدل رگرسیون (۲۷.۲) به صورت زیر تخمین زده می‌شوند:

$$\hat{Y} = (-2/6491, -0/8095, 1/0302)_T + (2/9897, 3/1499, 3/3098)_T x$$

شکل ۵.۲ نشان می‌دهد که مسیرهای پیشنهادی (۱۹.۲) و (۲۰.۲)، به جواب بهینه مسئله همگرا می‌شوند. شکل ۶.۲ نشان می‌دهد که مسیرهای شبکه پیشنهادی (۱۹.۲) و (۲۰.۲)، با هر نقطه شروع، همواره به جواب بهینه مدل همگرا هستند. تقریب ضرایب A_0 و A_1 از مدل رگرسیون (۲۷.۲) برای ۱۱ مقدار مختلف h در شکل ۷.۲ نشان داده شده است. جدول ۴.۲ نتایج عددی متناظر را نشان می‌دهد.

جدول ۴.۲: تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۲.۱.۲

$A_1[h]$	$A_0[h]$	h
[۲/۹۸۹۷, ۳/۳۰۹۸]	[-۲/۶۴۹۱, ۱/۰۳۰۲]	۰
[۳/۰۰۳۸, ۳/۲۹۲۸]	[-۲/۴۶۵۶, ۰/۸۴۶۰]	۰/۱
[۳/۰۲۱۷, ۳/۲۷۷۸]	[-۲/۲۸۱۲, ۰/۶۶۲۳]	۰/۲
[۳/۰۳۷۷, ۳/۲۶۱۸]	[-۲/۰۹۷۲, ۰/۴۷۸۳]	۰/۳
[۳/۰۵۳۸, ۳/۲۴۵۹]	[-۱/۹۱۳۲, ۰/۲۹۴۴]	۰/۴
[۳/۰۶۹۸, ۳/۲۲۹۹]	[-۱/۷۲۹۳, ۰/۱۱۰۴]	۰/۵
[۳/۰۸۵۸, ۳/۲۱۳۹]	[-۱/۵۴۵۳, -۰/۰۷۳۵]	۰/۶
[۳/۱۰۱۷, ۳/۱۹۷۷]	[-۱/۳۶۱۴, -۰/۲۵۷۵]	۰/۷
[۳/۱۱۷۵, ۳/۱۸۱۵]	[-۱/۱۷۷۵, -۰/۴۴۱۵]	۰/۸
[۳/۱۳۳۷, ۳/۱۶۵۷]	[-۰/۹۹۳۵, -۰/۶۲۵۴]	۰/۹
[۳/۱۴۹۹, ۳/۱۴۹۸]	[-۰/۸۰۹۵, -۰/۸۰۹۳]	۱

برای اثبات کارایی روش پیشنهادی، خطای به دست آمده از این روش را با روش های معروف ارائه شده در [۲۹]، [۷۶] و [۷۷] مقایسه می کنیم (جدول ۵.۲ مشاهده شود). واضح است که روش پیشنهادی، عملکرد بهتری روی مشاهدات داشته است. به منظور

جدول ۵.۲: ضرایب فازی و خطای به دست آمده در مثال ۲.۱.۲

MSPE			ضرایب فازی	روش
D_3	D_2	D_1		
۴/۱۳۵	۵/۰۷۰۰	۳۷/۳۶۴۵	$A_0 = (0, 3/850, 7/700)_T$ $A_1 = (2/100, 2/100, 2/100)_T$	مدل ارائه شده در [۷۷]
۵/۴۴۲	۴/۲۹۴۴	۱۶/۳۳۱۳	$A_0 = (-2/202, 0/118, 2/438)_T$ $A_1 = (1/718, 1/718, 1/718)_T$	مدل ارائه شده در [۷۶]
۵/۶۵۴	۲/۴۷۵۹	۱۵/۸۰۲۳	$A_0 = (3/110, 4/949, 6/789)_T$ $A_1 = (1/550, 1/710, 1/870)_T$	مدل ارائه شده در [۲۹]
۱/۵۰۷	۱/۷۴۴۵	۱۵/۵۲۷۲	$A_0 = (-2/649, -0/809, 1/030)_T$ $A_1 = (2/989, 3/149, 3/309)_T$	مدل (۱۹.۲) و (۲۰.۲)

تجسم بهتر روش پیشنهادی، نمودارهای حبابی فازی برای نشان دادن چگونگی تناسب این چهار مدل با داده های مشاهده شده ترسیم شده اند (شکل ۸.۲ مشاهده شود). نمودار حبابی فازی به وضوح نشان می دهد که شبکه عصبی پیشنهادی عملکرد بهتری نسبت به دو مدل قبلی در [۷۶] و [۷۷] داشته است (به دلیل شعاع کمتر دایره ها). علاوه بر این، نزدیکی دایره ها به نیمساز نشان می دهد که مدل در مقایسه با مدل ارائه شده در [۲۹]، تخمین بهتری برای مراکز بوده است.

مثال ۳.۱.۲. (سطح مقرون به صرفه قیمت خانه ها) نویسندگان در [۷۵] با استفاده از داده های ارائه شده در جدول ۷.۲ یک مدل رگرسیون خطی برای نشان دادن رابطه میان سطح مقرون به صرفه قیمت خانه از دیدگاه مصرف کننده و برخی از عوامل کلیدی پیشنهاد کرده اند. در مدل رگرسیون پیشنهادی، عوامل تاثیرگذار بر قیمت خانه شامل سه عامل اقتصادی (نرخ بهره وام مسکن، مالیات بر املاک و مستغلات و نسبت پیش پرداخت) و سه عامل غیر اقتصادی (اندازه مسکن، درآمد سالانه خانوار و جمعیت خانواده) بوده است.

جدول ۶.۲: ضرایب مدل رگرسیون (۲۸.۲) در مثال ۳.۱.۲

MSM	MSPE	ضرایب فازی	روش
۰/۴۰۶	۱۵/۱۸۶۷	$A_0 = (7/72, 13/01, 23/89)_T$ $A_1 = (468/16, 709/72, 805/62)_T$ $A_2 = (26/15, 20/12, 14/97)_T$ $A_3 = (45/65, 34/11, 27/52)_T$ $A_4 = (28/97, 25/7, 24/9)_T$ $A_5 = (664/92, 879/19, 1034/69)_T$ $A_6 = (9/52, 7/65, 4/59)_T$	مدل ارائه شده در [۷۵]
۰/۷۷۱	۸/۵۱۱۶	$A_0 = (-744/68, -744/68, -700/09)_T$ $A_1 = (6/4594, 7/2617, 7/2617)_T$ $A_2 = (-17/305, -17/30511, -16/783)_T$ $A_3 = (-53/130, -50/32998, -50/329)_T$ $A_4 = (-2/5652, -2/5652, -2/3139)_T$ $A_5 = (9/8045, 13/117, 15/428)_T$ $A_6 = (91/6282, 103/54, 116/50)_T$	مدل (۱۹.۲) و (۲۰.۲)

جدول ۷.۲: مقادیر مشاهده شده در مثال ۳.۱.۲

i	x_1	x_2	x_3	x_4	x_5	x_6	Y
۱	۸۰	۵/۹	۰/۶	۲۰	۲۷	۴	$(180, 300, 380)_T$
۲	۹۰	۴/۹	۰	۳۵	۳۰	۴	$(330, 470, 620)_T$
۳	۷۰	۶/۵۵	۰/۴	۳۰	۳۵	۳	$(230, 330, 430)_T$
۴	۱۲۰	۶/۸	۰/۶	۳۰	۴۰	۳	$(530, 750, 930)_T$
۵	۱۰۰	۴/۹	۰	۳۵	۳۳	۴	$(400, 600, 800)_T$
۶	۱۰۰	۶/۵۵	۰	۳۰	۴۰	۵	$(430, 670, 820)_T$
۷	۸۰	۵/۹	۰/۴	۲۰	۱۸	۵	$(120, 220, 270)_T$
۸	۹۰	۶/۸	۰/۶	۳۰	۴۵	۵	$(450, 630, 800)_T$
۹	۹۰	۶/۱۵	۰/۶	۳۰	۱۲	۴	$(100, 200, 250)_T$
۱۰	۱۰۰	۶/۱۵	۰	۳۰	۳۰	۳	$(380, 550, 630)_T$
۱۱	۸۰	۴/۹	۰	۳۵	۲۵	۳	$(220, 350, 400)_T$
۱۲	۹۰	۵/۹	۰/۶	۲۰	۲۲	۳	$(200, 330, 420)_T$
۱۳	۱۰۰	۶/۸	۰/۶	۳۰	۶۵	۳	$(700, 1000, 1270)_T$
۱۴	۷۰	۶/۱۵	۰/۶	۳۰	۲۳	۴	$(100, 180, 250)_T$
۱۵	۸۰	۶/۵۵	۰/۶	۳۰	۳۰	۴	$(200, 370, 470)_T$
۱۶	۷۰	۴/۹	۰/۴	۳۰	۲۳	۵	$(120, 200, 270)_T$
۱۷	۱۰۰	۶/۱۵	۰	۳۰	۳۴	۵	$(450, 620, 730)_T$
۱۸	۸۰	۵/۹	۰/۶	۲۰	۲۰	۴	$(120, 220, 250)_T$
۱۹	۸۰	۵/۹	۰/۶	۲۰	۱۸	۲	$(100, 200, 270)_T$
۲۰	۱۰۰	۶/۵۵	۰	۳۰	۴۲	۴	$(500, 680, 870)_T$
۲۱	۸۰	۴/۹	۰/۶	۳۵	۳۰	۳	$(230, 420, 480)_T$
۲۲	۱۰۰	۶/۱۵	۰	۳۰	۲۰	۳	$(250, 470, 500)_T$
۲۳	۸۰	۶/۱۵	۰/۶	۲۰	۲۳	۴	$(150, 280, 330)_T$
۲۴	۱۱۰	۶/۸	۰	۳۰	۳۵	۵	$(500, 680, 820)_T$
۲۵	۸۰	۳/۲۵	۰/۶	۳۰	۳۳	۳	$(280, 430, 470)_T$
۲۶	۱۰۰	۶/۱۵	۰	۳۰	۴۷	۴	$(530, 720, 900)_T$
۲۷	۱۲۰	۴/۷	۰/۴	۳۰	۵۷	۵	$(730, 1130, 1220)_T$

۲۸	۹۰	۵/۹	۰	۳۰	۲۳	۳	$(۲۷۰, ۳۸۰, ۴۸۰)_T$
۲۹	۱۰۰	۵/۹	۰	۳۰	۳۵	۴	$(۴۳۰, ۶۳۰, ۷۰۰)_T$
۳۰	۹۰	۴/۷	۰	۳۰	۲۵	۳	$(۲۷۰, ۳۵۰, ۵۰۰)_T$
۳۱	۱۰۰	۶/۸	۰	۳۰	۳۷	۴	$(۴۰۰, ۶۰۰, ۷۸۰)_T$
۳۲	۸۰	۶/۱۵	۰/۴	۲۰	۲۸	۳	$(۲۲۰, ۳۷۰, ۴۲۰)_T$
۳۳	۹۰	۶/۵۵	۰	۳۰	۲۰	۳	$(۲۰۰, ۳۲۰, ۳۸۰)_T$
۳۴	۱۰۰	۶/۵۵	۰/۴	۳۰	۳۲	۴	$(۳۵۰, ۵۰۰, ۶۵۰)_T$
۳۵	۹۰	۶/۵۵	۰	۳۰	۳۵	۴	$(۳۵۰, ۵۵۰, ۷۰۰)_T$
۳۶	۹۰	۵/۹	۰/۶	۲۰	۱۷	۴	$(۱۳۰, ۲۷۰, ۳۵۰)_T$
۳۷	۱۰۰	۶/۱۵	۰	۳۰	۳۵	۴	$(۴۰۰, ۶۳۰, ۷۵۰)_T$
۳۸	۸۰	۵/۹	۰	۳۰	۲۳	۲	$(۲۰۰, ۳۲۰, ۳۸۰)_T$
۳۹	۸۰	۶/۸	۰/۶	۲۰	۱۵	۴	$(۱۰۰, ۱۵۰, ۱۸۰)_T$
۴۰	۹۰	۶/۱۵	۰	۳۰	۲۰	۳	$(۲۳۰, ۳۵۰, ۳۸۰)_T$
۴۱	۸۰	۴/۹	۰/۶	۳۵	۲۷	۳	$(۲۰۰, ۳۲۰, ۴۲۰)_T$
۴۲	۸۰	۶/۱۵	۰	۳۰	۲۲	۳	$(۲۰۰, ۳۰۰, ۳۳۰)_T$
۴۳	۱۰۰	۶/۸	۰	۳۰	۴۳	۴	$(۴۸۰, ۶۷۰, ۸۲۰)_T$
۴۴	۱۰۰	۴/۹	۰	۳۵	۳۵	۵	$(۴۳۰, ۶۳۰, ۷۵۰)_T$
۴۵	۶۰	۶/۵۵	۰/۶	۲۰	۲۷	۴	$(۱۲۰, ۱۷۰, ۲۳۰)_T$
۴۶	۱۲۰	۴/۹	۰	۳۵	۵۰	۳	$(۷۳۰, ۹۷۰, ۱۱۷۰)_T$
۴۷	۷۰	۶/۱۵	۰/۶	۳۰	۲۵	۳	$(۱۲۰, ۱۸۰, ۲۵۰)_T$
۴۸	۱۰۰	۵/۹	۰/۴	۳۰	۲۸	۳	$(۳۰۰, ۵۳۰, ۶۰۰)_T$
۴۹	۸۰	۶/۸	۰/۶	۲۰	۲۰	۳	$(۱۲۰, ۲۲۰, ۳۰۰)_T$
۵۰	۱۰۰	۶/۱۵	۰/۴	۳۰	۳۷	۴	$(۳۷۰, ۶۲۰, ۷۲۰)_T$
۵۱	۸۰	۶/۱۵	۰	۲۰	۳۰	۴	$(۲۸۰, ۴۳۰, ۵۲۰)_T$
۵۲	۹۰	۶/۵۵	۰/۴	۳۰	۳۰	۳	$(۳۰۰, ۴۲۰, ۵۲۰)_T$
۵۳	۸۰	۵/۹	۰/۶	۲۰	۲۰	۳	$(۱۳۰, ۲۵۰, ۳۰۰)_T$
۵۴	۷۰	۴/۹	۰/۶	۳۵	۲۵	۲	$(۱۲۰, ۱۸۰, ۳۰۰)_T$
۵۵	۸۰	۵/۹	۰	۲۰	۳۰	۴	$(۲۸۰, ۴۰۰, ۴۸۰)_T$
۵۶	۹۰	۶/۵۵	۰/۶	۳۰	۳۵	۳	$(۳۲۰, ۵۰۰, ۶۸۰)_T$
۵۷	۸۰	۶/۸	۰	۲۰	۳۰	۴	$(۲۷۰, ۴۰۰, ۵۰۰)_T$
۵۸	۷۰	۵/۹	۰/۶	۲۰	۲۳	۳	$(۱۰۰, ۱۸۰, ۲۳۰)_T$
۵۹	۱۰۰	۶/۱۵	۰	۳۰	۳۰	۴	$(۳۷۰, ۵۵۰, ۷۰۰)_T$
۶۰	۸۰	۶/۵۵	۰/۶	۲۰	۳۳	۴	$(۲۲۰, ۴۵۰, ۵۰۰)_T$
۶۱	۸۰	۴/۹	۰	۳۵	۲۷	۴	$(۲۳۰, ۳۲۰, ۴۰۰)_T$
۶۲	۹۰	۳/۲۵	۰	۳۰	۲۵	۳	$(۲۸۰, ۳۸۰, ۴۷۰)_T$
۶۳	۶۰	۶/۵۵	۰/۴	۳۰	۳۰	۲	$(۱۵۰, ۲۵۰, ۳۲۰)_T$
۶۴	۸۰	۴/۹	۰	۳۵	۲۷	۴	$(۲۳۰, ۳۵۰, ۴۷۰)_T$
۶۵	۱۵۰	۶/۸	۰/۶	۳۰	۷۵	۵	$(۱۰۵۰, ۱۵۲۰, ۱۸۰۰)_T$
۶۶	۸۰	۵/۹	۰	۲۰	۲۰	۴	$(۱۷۰, ۲۷۰, ۳۰۰)_T$
۶۷	۷۰	۶/۱۵	۰/۶	۳۰	۲۳	۴	$(۱۰۰, ۱۷۰, ۲۳۰)_T$
۶۸	۸۰	۵/۹	۰/۴	۳۰	۲۵	۳	$(۲۰۰, ۳۰۰, ۳۷۰)_T$
۶۹	۸۰	۵/۹	۰/۶	۲۰	۳۳	۲	$(۲۵۰, ۴۵۰, ۵۷۰)_T$
۷۰	۸۰	۴/۹	۰	۳۵	۲۸	۳	$(۲۳۰, ۳۷۰, ۴۷۰)_T$
۷۱	۷۰	۶/۵۵	۰/۴	۳۰	۲۳	۲	$(۱۳۰, ۲۰۰, ۲۵۰)_T$
۷۲	۸۰	۶/۸	۰	۲۰	۳۷	۴	$(۳۳۰, ۴۵۰, ۶۷۰)_T$
۷۳	۷۰	۶/۱۵	۰/۶	۳۰	۳۰	۴	$(۱۵۰, ۲۷۰, ۳۲۰)_T$
۷۴	۸۰	۶/۵۵	۰/۴	۳۰	۲۲	۳	$(۱۵۰, ۲۷۰, ۳۲۰)_T$

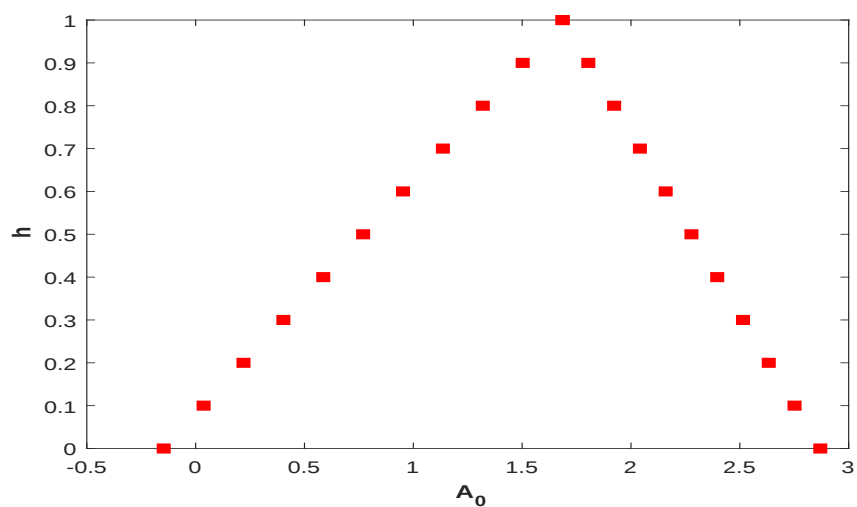
۷۵	۱۰۰	۴/۹	۰/۶	۳۵	۳۶	۳	(۳۷۰،۶۵۰،۶۸۰) _T
۷۶	۸۰	۵/۹	۰	۲۰	۲۰	۴	(۱۷۰،۲۵۰،۳۳۰) _T
۷۷	۱۰۰	۶/۱۵	۰	۳۰	۳۵	۵	(۴۵۰،۵۸۰،۷۵۰) _T
۷۸	۸۰	۶/۸	۰/۶	۳۰	۳۷	۳	(۳۰۰،۴۷۰،۶۰۰) _T
۷۹	۱۰۰	۵/۹	۰/۴	۲۰	۲۵	۲	(۳۲۰،۵۰۰،۶۲۰) _T
۸۰	۹۰	۶/۵۵	۰	۳۰	۲۰	۳	(۲۵۰،۳۵۰،۳۸۰) _T
۸۱	۱۰۰	۶/۸	۰	۳۰	۳۷	۲	(۴۲۰،۶۷۰،۷۳۰) _T
۸۲	۸۰	۶/۸	۰/۶	۲۰	۳۵	۳	(۳۰۰،۴۷۰،۵۸۰) _T
۸۳	۹۰	۶/۱۵	۰/۶	۲۰	۴۰	۲	(۳۸۰،۵۳۰،۶۵۰) _T
۸۴	۸۰	۶/۵۵	۰	۳۰	۳۰	۳	(۲۸۰،۴۳۰،۵۰۰) _T
۸۵	۶۰	۶/۱۵	۰/۴	۳۰	۲۸	۳	(۱۰۰،۱۷۰،۲۳۰) _T
۸۶	۷۰	۴/۹	۰/۶	۳۵	۲۲	۳	(۱۲۰،۱۸۰،۲۵۰) _T
۸۷	۱۰۰	۴/۷	۰/۶	۳۰	۴۳	۳	(۴۸۰،۶۵۰،۹۲۰) _T
۸۸	۹۰	۴/۹	۰/۶	۳۵	۳۰	۲	(۳۰۰،۴۷۰،۵۸۰) _T
۸۹	۷۰	۴/۹	۰/۶	۳۵	۲۷	۴	(۱۵۰،۲۷۰،۳۲۰) _T
۹۰	۱۱۰	۵/۹	۰/۶	۲۰	۴۷	۳	(۵۸۰،۸۳۰،۱۰۵۰) _T
۹۱	۸۰	۴/۷	۰	۳۰	۲۰	۲	(۱۷۰،۲۵۰،۳۲۰) _T
۹۲	۱۰۰	۶/۸	۰/۴	۳۰	۳۵	۴	(۴۰۰،۶۲۰،۷۰۰) _T
۹۳	۸۰	۴/۹	۰/۴	۴۰	۳۲	۴	(۲۳۰،۴۳۰،۵۵۰) _T
۹۴	۱۰۰	۶/۱۵	۰	۳۰	۴۰	۳	(۴۸۰،۷۸۰،۸۰۰) _T
۹۵	۸۰	۶/۱۵	۰	۳۰	۳۵	۳	(۳۳۰،۴۸۰،۵۲۰) _T
۹۶	۹۰	۵/۹	۰/۶	۳۰	۲۷	۴	(۲۳۰،۴۲۰،۴۳۰) _T
۹۷	۷۰	۵/۹	۰/۶	۳۰	۲۳	۴	(۱۰۰،۱۷۰،۲۵۰) _T
۹۸	۸۰	۵/۹	۰	۳۰	۲۵	۲	(۲۲۰،۳۲۰،۳۸۰) _T
۹۹	۸۰	۳/۲۵	۰/۴	۳۰	۱۸	۳	(۱۳۰،۲۲۰،۲۷۰) _T
۱۰۰	۷۰	۶/۱۵	۰	۳۰	۳۰	۴	(۲۰۰،۳۵۰،۳۷۰) _T
۱۰۱	۱۲۰	۶/۸	۰	۳۰	۵۰	۳	(۶۸۰،۱۰۰۰،۱۱۵۰) _T
۱۰۲	۹۰	۶/۱۵	۰	۳۰	۳۳	۴	(۳۵۰،۵۳۰،۵۸۰) _T
۱۰۳	۷۰	۶/۸	۰	۳۰	۲۳	۴	(۱۲۰،۲۲۰،۲۸۰) _T
۱۰۴	۱۰۰	۶/۸	۰/۶	۳۰	۳۵	۳	(۳۸۰،۵۸۰،۷۳۰) _T
۱۰۵	۹۰	۴/۹	۰	۳۵	۳۵	۳	(۴۰۰،۵۷۰،۶۵۰) _T
۱۰۶	۷۰	۶/۱۵	۰/۶	۳۰	۲۵	۲	(۱۲۰،۲۰۰،۲۸۰) _T
۱۰۷	۸۰	۶/۱۵	۰/۴	۳۰	۲۵	۳	(۲۰۰،۳۰۰،۳۳۰) _T
۱۰۸	۱۰۰	۶/۵۵	۰	۲۰	۳۲	۳	(۳۸۰،۵۷۰،۷۳۰) _T
۱۰۹	۹۰	۵/۹	۰	۲۰	۳۰	۳	(۳۳۰،۵۲۰،۵۷۰) _T
۱۱۰	۱۲۰	۶/۱۵	۰	۲۰	۴۷	۲	(۷۳۰،۱۰۰۰،۱۰۵۰) _T
۱۱۱	۱۰۰	۶/۵۵	۰/۴	۳۰	۵۰	۴	(۵۷۰،۸۰۰،۹۲۰) _T
۱۱۲	۹۰	۶/۵۵	۰/۴	۳۰	۲۰	۳	(۱۸۰،۳۰۰،۳۸۰) _T
۱۱۳	۶۰	۶/۱۵	۰	۳۰	۳۰	۲	(۱۵۰،۲۵۰،۳۲۰) _T
۱۱۴	۷۰	۵/۹	۰	۶۰	۲۷	۳	(۱۷۰،۲۵۰،۳۳۰) _T
۱۱۵	۷۰	۶/۵۵	۰	۳۰	۳۲	۳	(۲۵۰،۳۲۰،۳۷۰) _T
۱۱۶	۸۰	۴/۹	۰	۴۰	۳۵	۳	(۳۲۰،۴۵۰،۵۳۰) _T
۱۱۷	۸۰	۵/۹	۰	۲۰	۳۵	۵	(۳۵۰،۵۰۰،۶۲۰) _T
۱۱۸	۹۰	۶/۱۵	۰	۳۰	۲۳	۲	(۲۵۰،۳۸۰،۴۷۰) _T
۱۱۹	۱۰۰	۶/۵۵	۰/۶	۲۰	۳۵	۴	(۳۷۰،۶۲۰،۸۲۰) _T
۱۲۰	۸۰	۵/۹	۰	۳۰	۲۸	۳	(۲۵۰،۳۷۰،۴۳۰) _T
۱۲۱	۱۰۰	۴/۹	۰/۶	۳۰	۴۸	۳	(۵۲۰،۷۸۰،۹۵۰) _T

۱۲۲	۸۰	۶/۵۵	۰/۶	۳۰	۳۷	۴	$(280, 530, 580)_T$
۱۲۳	۶۰	۵/۹	۰/۶	۳۰	۲۷	۴	$(120, 180, 220)_T$
۱۲۴	۸۰	۶/۱۵	۰/۶	۲۰	۴۷	۳	$(430, 600, 680)_T$
۱۲۵	۸۰	۶/۱۵	۰	۳۰	۳۰	۴	$(250, 400, 470)_T$
۱۲۶	۷۰	۵/۹	۰	۳۰	۲۳	۳	$(150, 230, 270)_T$
۱۲۷	۷۰	۶/۱۵	۰/۶	۳۰	۲۳	۳	$(100, 170, 230)_T$
۱۲۸	۱۰۰	۵/۹	۰	۶۰	۴۰	۴	$(480, 720, 820)_T$
۱۲۹	۷۰	۶/۱۵	۰	۲۰	۳۳	۲	$(270, 380, 480)_T$
۱۳۰	۹۰	۶/۸	۰/۶	۳۰	۳۴	۳	$(280, 480, 570)_T$
۱۳۱	۸۰	۶/۵۵	۰	۲۰	۳۰	۴	$(270, 380, 470)_T$
۱۳۲	۹۰	۵/۹	۰	۳۰	۲۷	۴	$(280, 400, 500)_T$
۱۳۳	۸۰	۶/۱۵	۰	۳۰	۳۲	۴	$(280, 380, 450)_T$
۱۳۴	۱۰۰	۳/۲۵	۰/۴	۳۰	۳۷	۳	$(450, 630, 730)_T$
۱۳۵	۸۰	۶/۵۵	۰/۴	۳۰	۳۵	۳	$(280, 420, 520)_T$
۱۳۶	۱۰۰	۶/۵۵	۰/۴	۳۰	۳۰	۳	$(350, 520, 630)_T$
۱۳۷	۱۲۰	۶/۱۵	۰	۳۰	۴۷	۳	$(650, 1020, 1130)_T$
۱۳۸	۷۰	۵/۹	۰/۶	۳۰	۲۵	۲	$(120, 200, 280)_T$
۱۳۹	۷۰	۵/۹	۰	۳۰	۲۵	۳	$(300, 400, 420)_T$
۱۴۰	۹۰	۶/۱۵	۰	۳۰	۲۰	۳	$(200, 320, 430)_T$
۱۴۱	۱۰۰	۶/۸	۰	۳۰	۲۸	۲	$(350, 520, 680)_T$
۱۴۲	۸۰	۶/۸	۰/۶	۳۰	۱۷	۳	$(100, 180, 230)_T$
۱۴۳	۸۰	۶/۱۵	۰/۴	۳۰	۲۳	۳	$(180, 250, 320)_T$
۱۴۴	۹۰	۴/۹	۰	۳۵	۲۳	۳	$(270, 370, 470)_T$
۱۴۵	۹۰	۶/۱۵	۰	۳۰	۳۰	۴	$(300, 480, 580)_T$

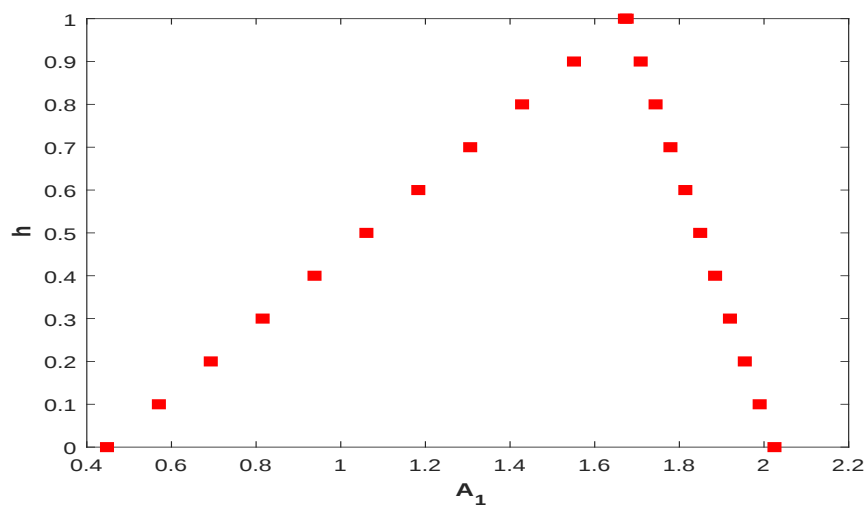
با استفاده از این داده‌ها و با توجه به مدل پیشنهادی (۱۹.۲) و (۲۰.۲) به دنبال تقریب ضرایب مدل رگرسیون زیر هستیم:

$$Y = A_0 + A_1x_1 + A_2x_2 + A_3x_3 + A_4x_4 + A_5x_5 + A_6x_6 \quad (28.2)$$

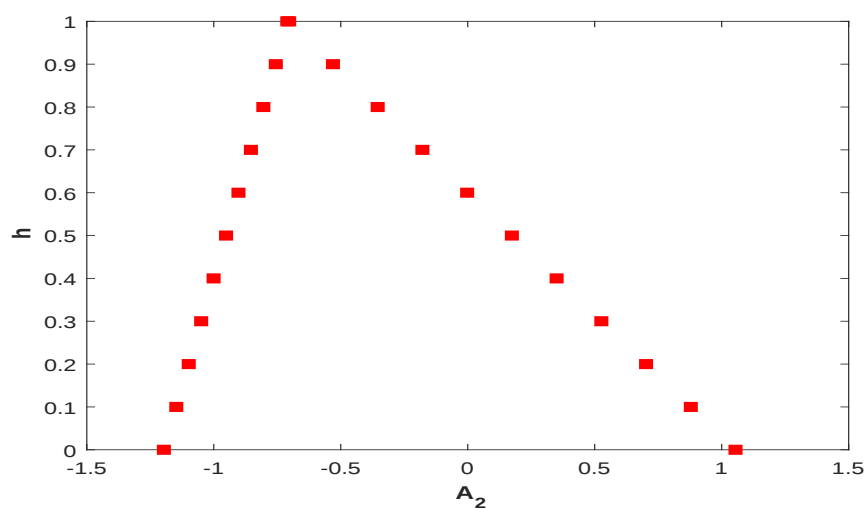
نتایج به دست آمده در جدول ۶.۲ نشان داده شده است. شکل ۹.۲ مسیرهای شبکه عصبی (۱۹.۲) و (۲۰.۲) را نشان می‌دهد. برای مقایسه کارایی، مدل ارائه شده در [۷۵] را در نظر می‌گیریم. نتایج عددی برای ۱۴۵ مشاهده در جدول ۶.۲ گزارش شده است. واضح است که مدل پیشنهادی عملکرد بهتری نسبت به مدل ارائه شده در [۷۵] داشته است. علاوه بر این، برای تجسم چگونگی تناسب این مدل با داده‌های فازی مشاهده شده؛ در شکل ۱۰.۲ نمودار حبابی فازی ترسیم شده است. دیده می‌شود که شبکه عصبی (۱۹.۲) و (۲۰.۲) مراکز و پهنای داده‌های فازی را بهتر از مدل ارائه شده در [۷۵] تخمین زده است (به دلیل نزدیک تر بودن دایره‌ها به نیمساز).



(ا) تقریب ضریب A_0

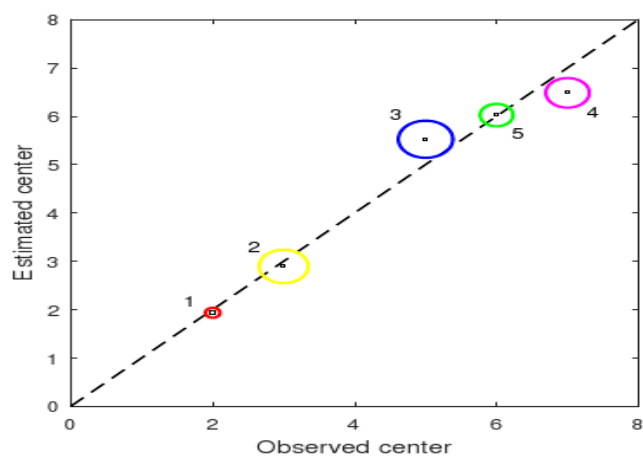


(ب) تقریب ضریب A_1

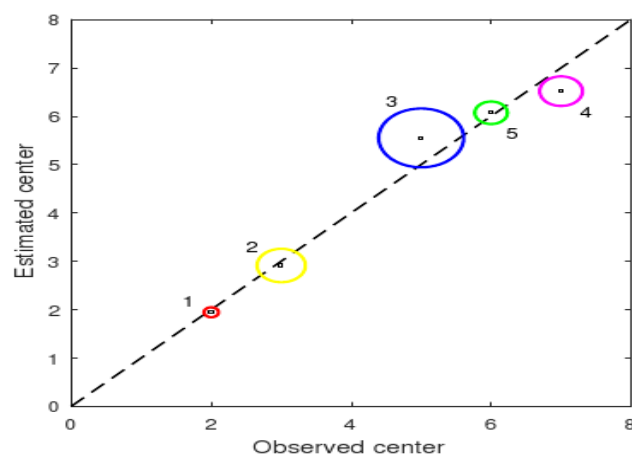


(ج) تقریب ضریب A_2

شکل ۳.۲: تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۱.۱.۲

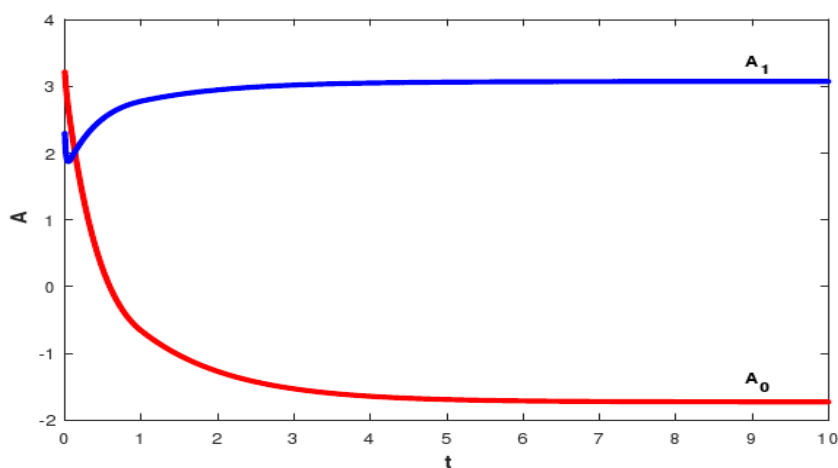


(آ) مدل (۱۹.۲) و (۲۰.۲)

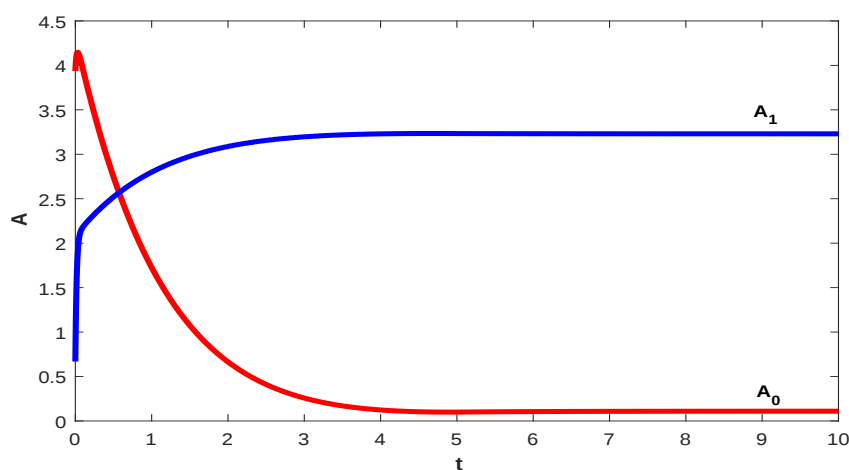


(ب) مدل ارائه شده در [۲۹]

شکل ۴.۲: نمودار حبابی فازی در مثال ۱.۱.۲.

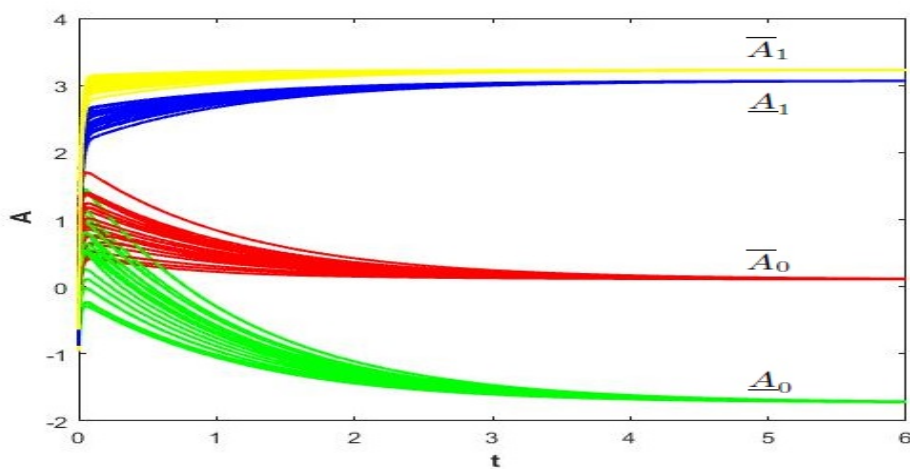


(آ) مسیرهای $\underline{A}_1$ و $\underline{A}_0$

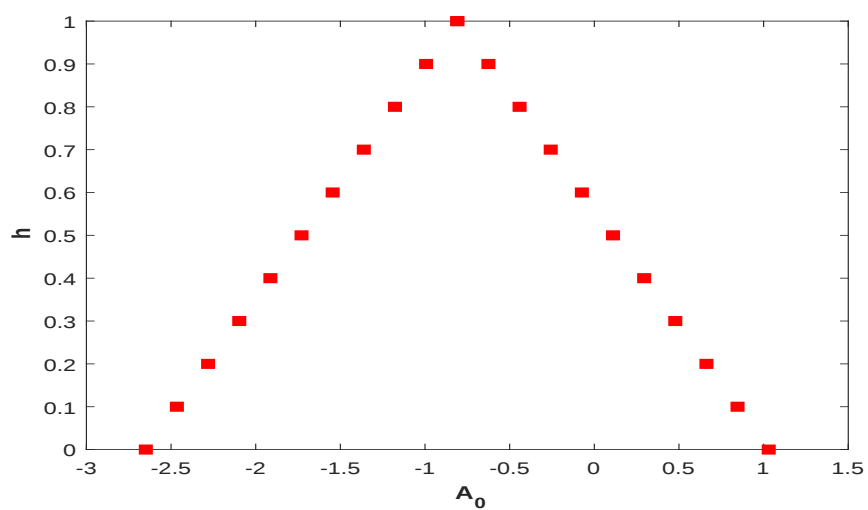


(ب) مسیرهای $\bar{A}_1$ و $\bar{A}_0$

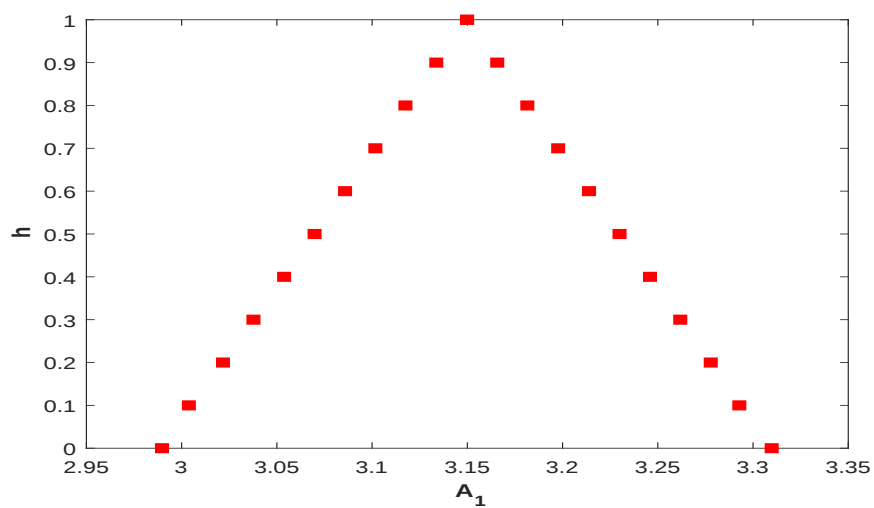
شکل ۵.۲: رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با $h = 0.5$ در مثال ۲.۱.۲.



شکل ۶.۲: رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با ۲۰ نقطه شروع متفاوت و $h = 0.5$ در مثال ۲.۱.۲.

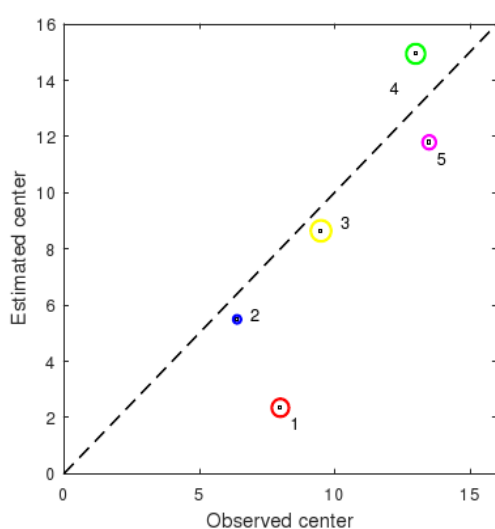


(آ) تقریب ضریب A_0

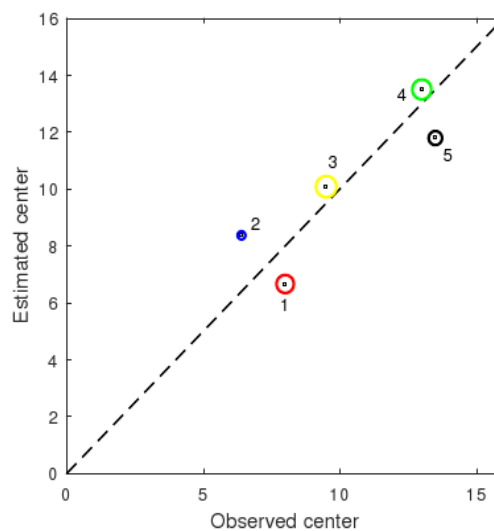


(ب) تقریب ضریب A_1

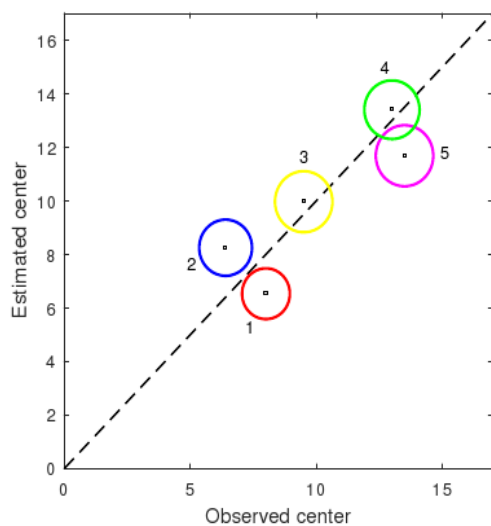
شکل ۷.۲: تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۲.۱.۲



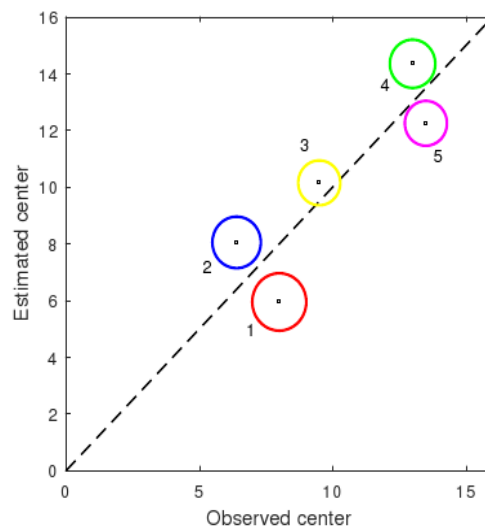
(ب) مدل ارائه شده در [۲۹]



(آ) مدل (۱۹.۲) و (۲۰.۲)

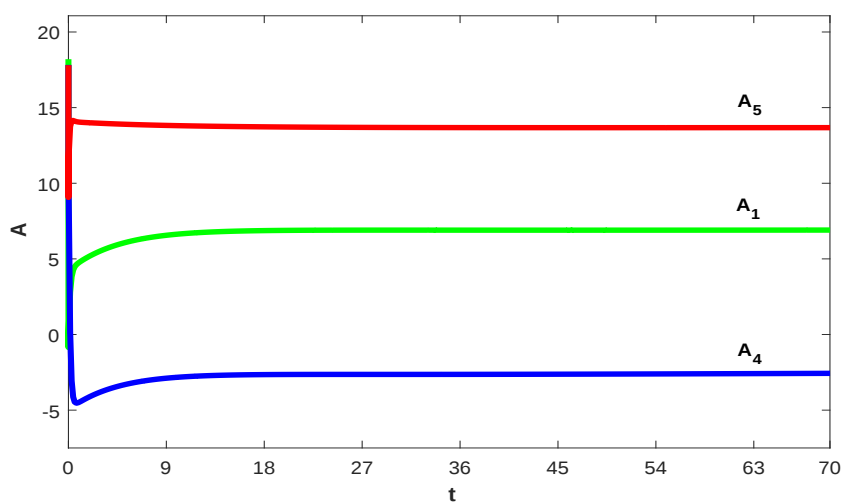


(د) مدل ارائه شده در [۷۶]

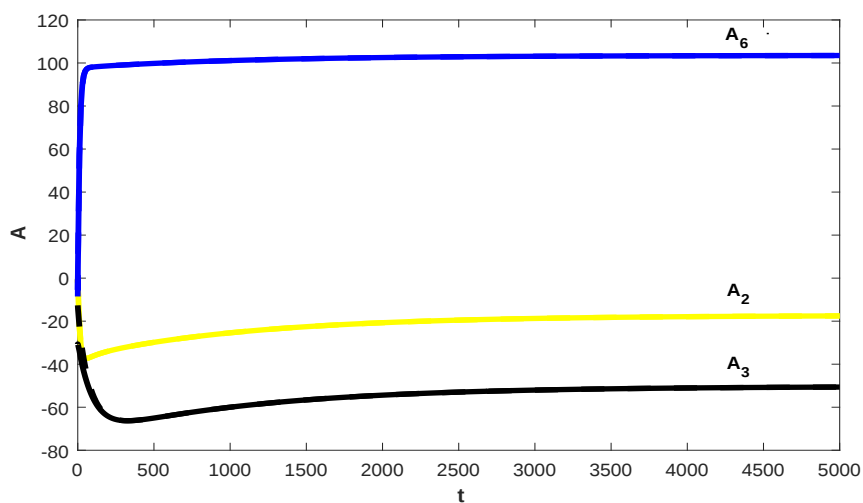


(ج) مدل ارائه شده در [۷۷]

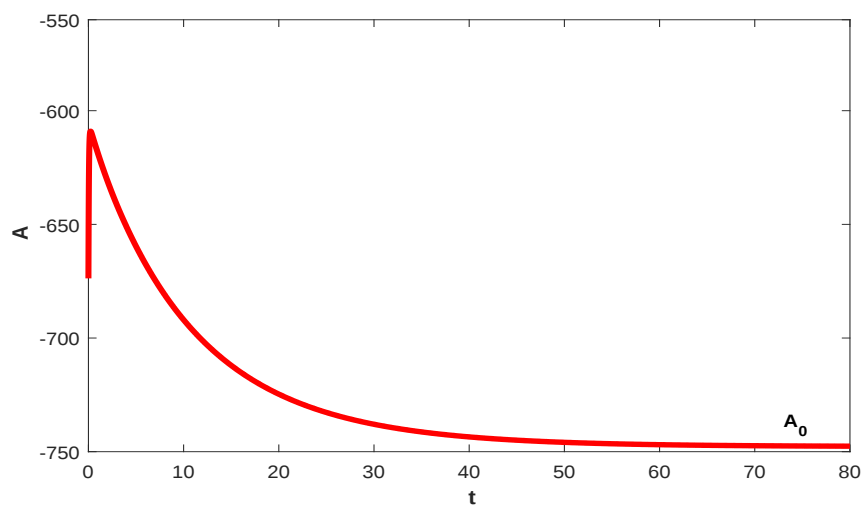
شکل ۸.۲: نمودار حبابی فازی در مثال ۲.۱.۲.



(آ) مسیرهای A_5 ، A_4 و A_1

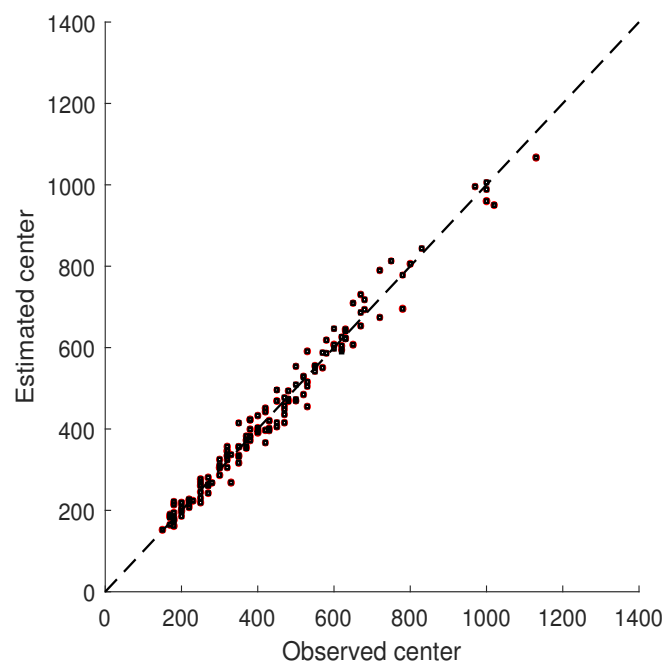


(ب) مسیرهای A_6 ، A_3 و A_2

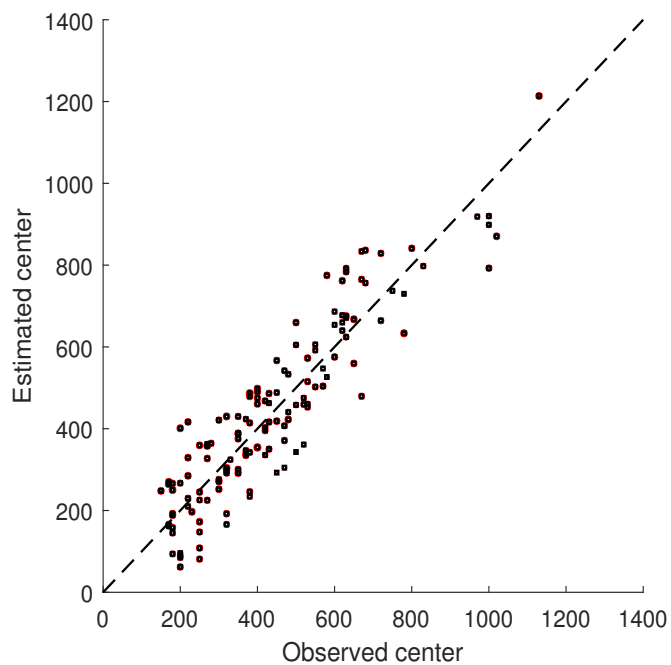


(ج) مسیر A_0

شکل ۹.۲: رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با $h = 1$ در مثال ۳.۱.۲.



(آ) مدل (۱۹.۲) و (۲۰.۲)



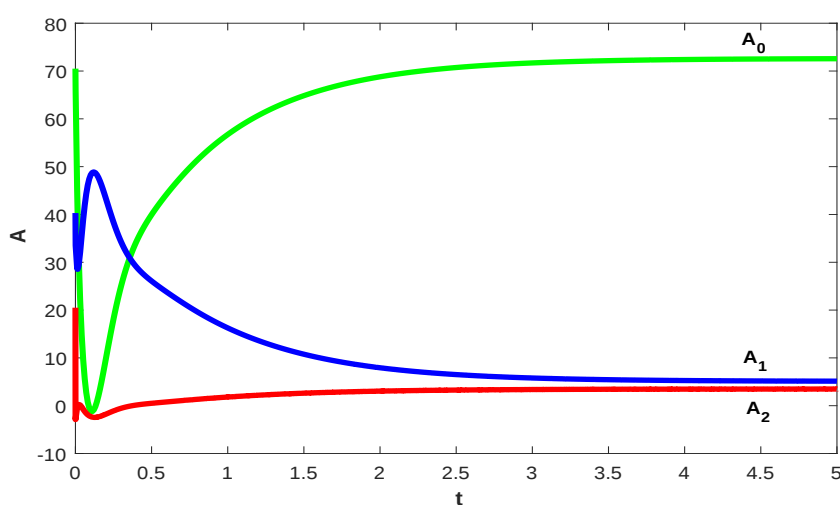
(ب) مدل ارائه شده در [۷۵]

شکل ۱۰.۲: نمودار حبابی فازی در مثال ۳.۱.۲.

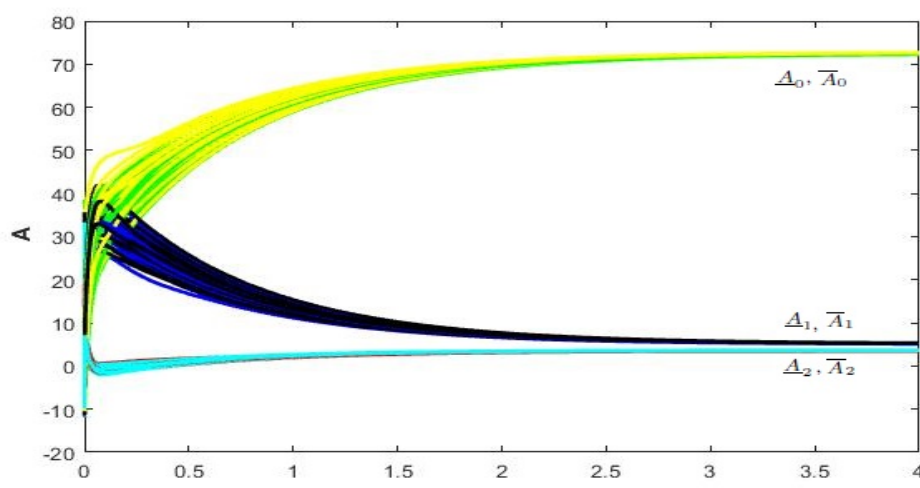
مثال ۴.۱.۲. در یک نظرسنجی در مورد هزینه سالانه تعمیرات برای ۵ مدل خودرو داده ها عبارتند از: هزینه تعمیر سالانه (دلار) و سن خودرو (سال) (جدول ۹.۲ مشاهده شود). معادله رگرسیون فازی را به صورت زیر در نظر بگیرید:

$$Y = A_0 + A_1x + A_2x^2 \quad (29.2)$$

با استفاده از مدل (۱۹.۲) و (۲۰.۲) به تقریب ضرایب مدل (۲۹.۲) می پردازیم. شکل ۱۱.۲ به ترتیب مسیرهای شبکه عصبی پیشنهادی را با یک و پنجاه نقطه شروع متفاوت نشان می دهد.



(آ) مسیرهای A_0 ، A_1 و A_2



(ب) مسیرهای A_0 ، A_1 و A_2 با ۵۰ نقطه شروع متفاوت

شکل ۱۱.۲: رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با $h = 1$ در مثال ۴.۱.۲.

بنابراین ضرایب مدل رگرسیون به صورت زیر تخمین زده می‌شود:

$$\hat{Y} = (71/7634, 72/5568, 74/1616)_T + (3/7704, 5/1923, 5/3716)_{Tx} \\ + (3/1389, 3/5612, 4/1381)_{Tx^2}$$

برای مقادیر مختلف h ، ضرایب رگرسیون A_0 ، A_1 و A_2 در شکل ۱۲.۲ و مقادیر عددی متناظرش در جدول ۸.۲ نشان داده شده است.

جدول ۸.۲: تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۴.۱.۲

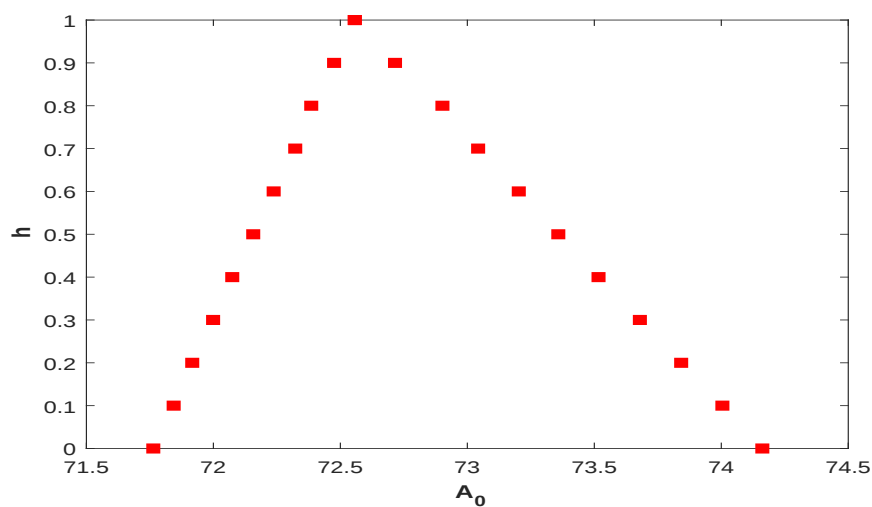
h	$A_0[h]$	$A_1[h]$	$A_2[h]$
۰	[71/7634, 74/1616]	[3/7704, 5/3716]	[3/1389, 4/1381]
۰/۱	[71/8436, 74/0043]	[3/9130, 5/3524]	[3/1814, 4/0814]
۰/۲	[71/9167, 73/8420]	[4/0612, 5/3369]	[3/2240, 4/0236]
۰/۳	[71/9989, 73/6794]	[4/2020, 5/3216]	[3/2659, 3/9656]
۰/۴	[72/0743, 73/5166]	[4/3485, 5/3069]	[3/3088, 3/9094]
۰/۵	[72/1569, 73/3582]	[4/4894, 5/2885]	[3/3519, 3/8521]
۰/۶	[72/2374, 73/2026]	[4/6321, 5/1964]	[3/3955, 3/7952]
۰/۷	[72/3223, 73/0425]	[4/7711, 5/2224]	[3/4385, 3/7384]
۰/۸	[72/3883, 72/9019]	[4/9268, 5/2393]	[3/4793, 3/6837]
۰/۹	[72/4753, 72/7152]	[5/0619, 5/2510]	[3/5226, 3/6239]
۱	[72/5568, 72/5568]	[5/1923, 5/2680]	[3/5612, 3/5612]

برای مقایسه، با استفاده از شبکه عصبی ارائه شده در مثال ۶.۳ از [۶۵] به حل این مسئله می‌پردازیم. خطای محاسبه شده از این روش با ۵۰ تکرار از الگوریتم یادگیری در جدول ۹.۲ نشان داده شده است. همچنین نمودارهای حبابی فازی برای هر دو مدل ترسیم شده است

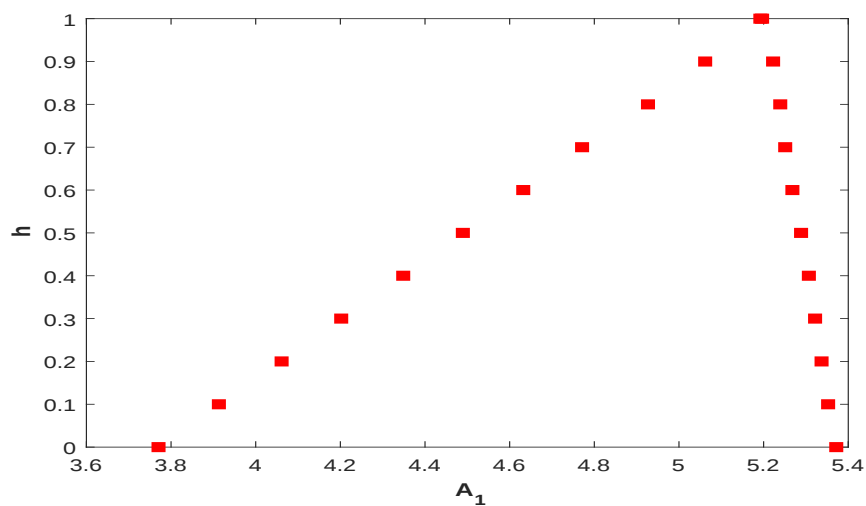
جدول ۹.۲: مقادیر مشاهده شده و خطای به‌دست آمده در مثال ۴.۱.۲

مقادیر مشاهده شده								خطای به‌دست آمده	
مدل ارائه شده در [۶۵]				مدل (۱۹.۲) و (۲۰.۲)					
$Repair$	Age	D_1	D_2	D_3	S_{UI}	D_1	D_2	D_3	S_{UI}
$(78, 81, 83)_T$	۱	3/044	1/027	0/124	0/665	2/816	1/029	0/137	0/676
$(93, 98, 103)_T$	۲	12/790	1/915	0/685	0/654	12/149	1/877	0/655	0/650
$(112, 120, 127)_T$	۳	1/598	1/458	0/119	0/918	1/495	1/391	0/146	0/921
$(135, 150, 161)_T$	۴	12/355	2/485	0/404	0/862	11/593	2/603	0/420	0/869
$(170, 188, 205)_T$	۵	2/959	1/171	0/157	0/946	3/878	1/008	0/106	0/936
MSPE		6/549	1/611	0/298					0/293
MSM						0/809			
									0/810

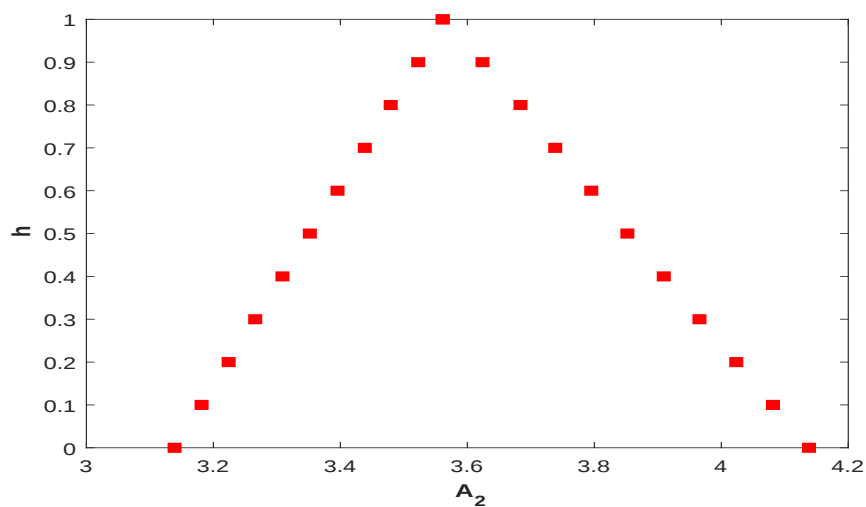
(به شکل ۱۳.۲ مراجعه شود). نتایج این مثال نشان می‌دهد که ما با موفقیت یک روش برآورد مناسب را تعریف کرده ایم که دارای اثر برازش مناسب و کارایی در عمل می‌باشد.



(آ) تقریب ضریب A_0

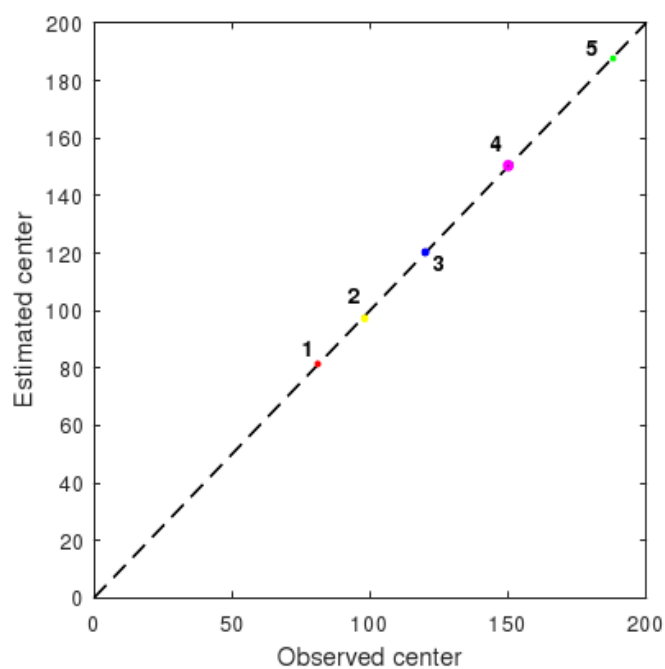


(ب) تقریب ضریب A_1

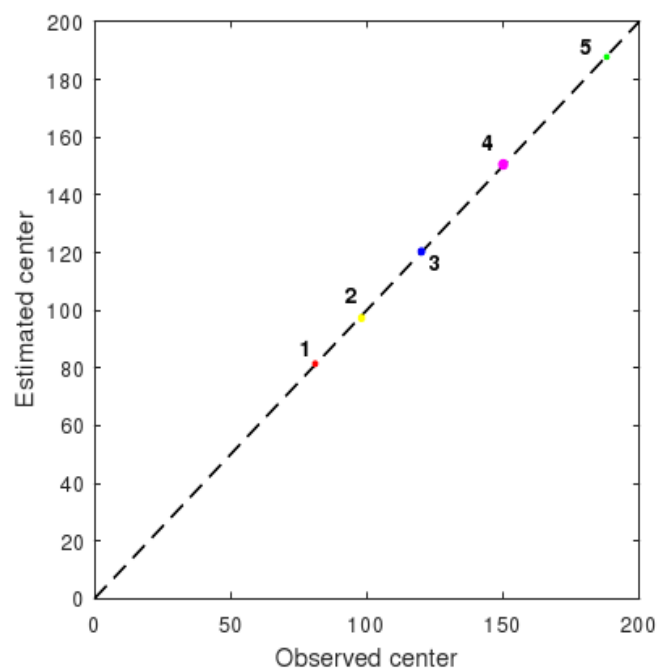


(ج) تقریب ضریب A_2

شکل ۱۲.۲: تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۴.۱.۲



(آ) مدل (۱۹.۲) و (۲۰.۲)



(ب) مدل ارائه شده در [۶۵]

شکل ۱۳.۲: نمودار حبابی فازی در مثال ۴.۱.۲.

مثال ۵.۱.۲. در این مثال برای مقایسه کارایی مدل با دیگر روش‌های موجود در [۷۴] و [۷۸] از داده‌های مشاهده شده در جدول ۱۰.۲ زمانی که در میان داده‌ها داده پرت وجود دارد، استفاده می‌کنیم.

جدول ۱۰.۲: مقادیر مشاهده شده در مثال ۵.۱.۲

x	۱	۲	۳	۴	۵
Y	$(۶/۴, ۲/۲, ۲/۲)_T$	$(۸, ۱/۸, ۱/۸)_T$	$(۱۶/۵, ۲/۶, ۲/۶)_T$	$(۱۱/۵, ۲/۶, ۲/۶)_T$	$(۱۳, ۲/۴, ۲/۴)_T$

مدل رگرسیون خطی زیر را در نظر بگیرید:

$$Y = A_0 + A_1 x \quad (۳۰.۲)$$

با استفاده از مدل (۱۹.۲) و (۲۰.۲) ضرایب به صورت زیر تقریب زده می‌شوند:

$$\hat{Y} = (۴/۱۱۰۰, ۶/۰۷۰۰, ۸/۰۳۰۰)_T + (۱/۵۵۰۰, ۱/۶۵۰۰, ۱/۷۹۰۰)_T x$$

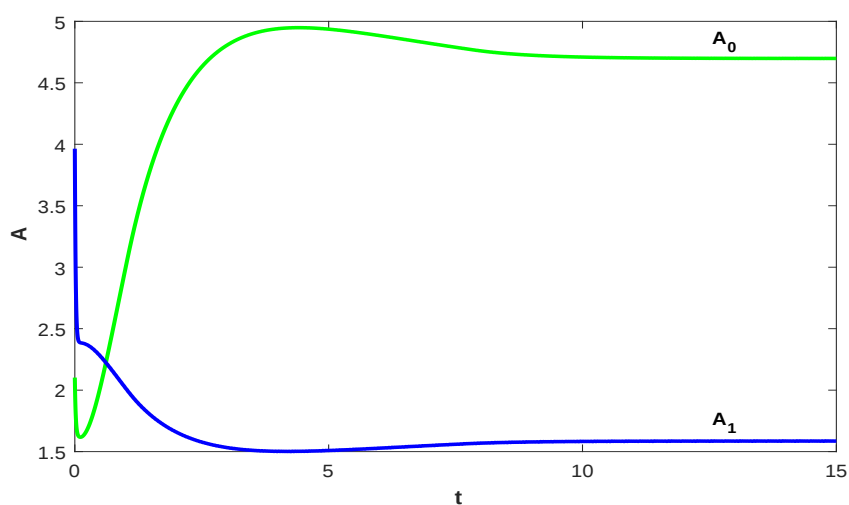
شکل ۱۴.۲، نشان می‌دهد که مسیرهای (۱۹.۲) و (۲۰.۲) به جواب بهینه مسئله همگرا می‌باشند. شکل ۱۵.۲ رفتار مسیرهای ضرایب رگرسیون A_0 و A_1 را با ۲۰ نقطه شروع مختلف نشان می‌دهد. تقریب ضرایب رگرسیون A_0 و A_1 از مدل (۳۰.۲) با ۱۱ مقدار مختلف h در شکل ۱۶.۲ و نتایج عددی در جدول ۱۱.۲ نشان داده شده است.

جدول ۱۱.۲: تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۵.۱.۲

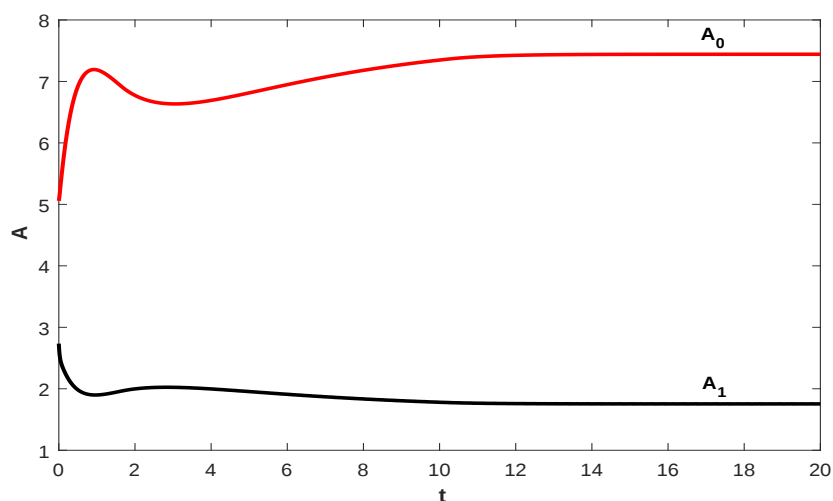
$A_1[h]$	$A_0[h]$	h
$[۲/۹۸۹۷, ۱/۷۹۰۰]$	$[۴/۱۱۰۰, ۱/۵۵۰۰]$	۰
$[۱/۵۶۲۰, ۱/۷۷۸۰]$	$[۴/۳۰۶۰, ۷/۸۳۴۰]$	۰/۱
$[۱/۵۷۴۰, ۱/۷۶۶۰]$	$[۴/۵۰۲۰, ۷/۶۳۸۰]$	۰/۲
$[۱/۵۸۶۰, ۱/۷۵۴۰]$	$[۴/۶۹۸۰, ۷/۴۴۲۰]$	۰/۳
$[۱/۵۹۸۰, ۱/۷۴۲۰]$	$[۴/۸۹۴۰, ۷/۲۴۶۰]$	۰/۴
$[۱/۶۱۰۰, ۱/۷۳۰۰]$	$[۵/۰۹۰۰, ۷/۰۵۰۰]$	۰/۵
$[۱/۶۲۲۰, ۱/۷۱۸۰]$	$[۵/۲۸۶۰, ۶/۸۵۴۰]$	۰/۶
$[۱/۶۳۴۰, ۱/۷۰۶۰]$	$[۵/۴۸۲۰, ۶/۶۵۸۰]$	۰/۷
$[۱/۶۴۶۰, ۱/۶۹۴۰]$	$[۵/۶۷۸۰, ۶/۴۶۲۰]$	۰/۸
$[۱/۶۵۸۰, ۱/۶۸۲۰]$	$[۵/۸۷۴۰, ۶/۲۶۶۰]$	۰/۹
$[۱/۶۷۰۰, ۱/۶۷۰۰]$	$[۶/۰۷۰۰, ۶/۰۷۰۰]$	۱

ضرایب فازی به‌دست آمده با استفاده از روش پیشنهادی و روش‌های موجود در [۷۴] و [۷۸] در جدول ۱۲.۲ مقایسه شده‌اند. مشاهده می‌شود که روش پیشنهادی، عملکرد بهتری نسبت به دیگر روش‌های موجود داشته است.

نمودار حبابی فازی نیز برای مقایسه کارایی در شکل ۱۷.۲ به تصویر کشیده شده است. به وضوح نشان دهنده عملکرد بهتر شبکه عصبی پیشنهادی (۱۹.۲) و (۲۰.۲) نسبت به دو مدل قبلی در [۲۹] و [۷۷] می‌باشد.



(آ) مسیرهای A_0 و A_1



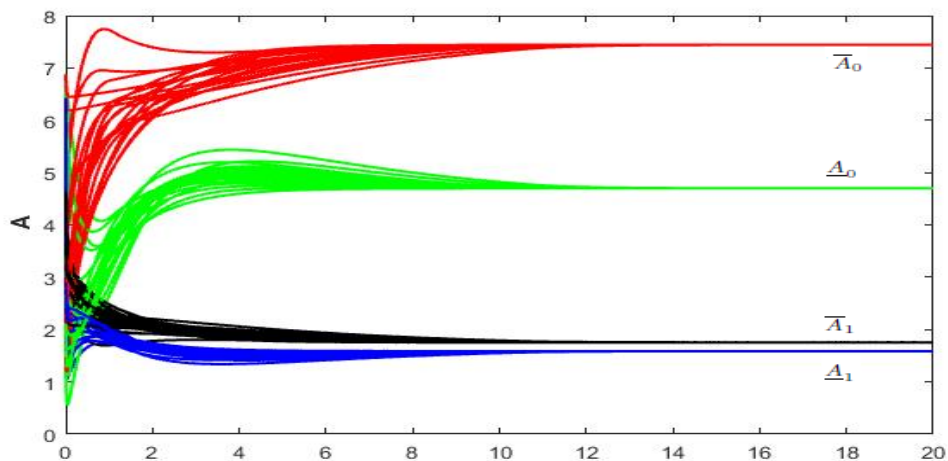
(ب) مسیرهای $\bar{A}_0$ و $\bar{A}_1$

شکل ۵.۱.۲: رفتار مسیر A شبکه عصبی (۱۹.۲) و (۲۰.۲) با $h = 0.3$ در مثال ۵.۱.۲.

جدول ۱۲.۲: ضرایب فازی و خطای به‌دست آمده در مثال ۵.۱.۲

MSM	MSPE		ضرایب فازی	روش
	D_3	D_2		
۰/۲۲۱۹	۲۰/۱۹۵	۱۶/۴۲۴۵	$A_0 = (5/5067, 5/5067, 5/5067)_T$ $A_1 = (1/3341, 2/4656, 3/5971)_T$	مدل ارائه شده در [۷۸]
۰/۲۶۴۱	۵/۰۴۰	۲/۶۱۰۰	$A_0 = (4/100, 6/060, 8/0200)_T$ $A_1 = (1/550, 1/67, 1/790)_T$	مدل ارائه شده در [۷۴]
۰/۲۶۷۸	۵/۰۳۰	۲/۵۸۰۰	$A_0 = (4/1100, 6/0700, 8/0300)_T$ $A_1 = (1/5500, 1/6700, 1/7900)_T$	مدل (۱۹.۲) و (۲۰.۲)

مثال ۶.۱.۲. (تأمین نیروی کار) هنگامی که برخی از مشکلات خرد و کلان در زمینه اقتصاد مورد مطالعه قرار می‌گیرند، اغلب معادلاتی را ایجاد می‌کنند که برخی از پارامترها در آن دقیق

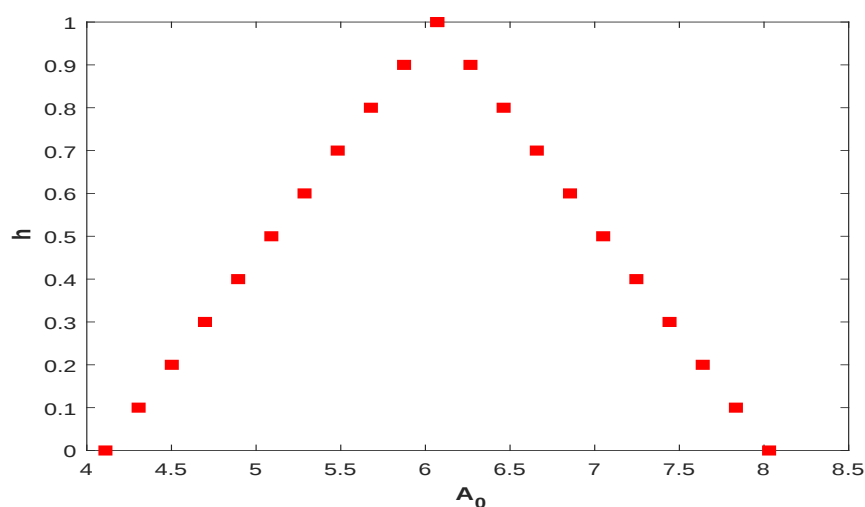


شکل ۱۵.۲: رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با ۲۰ نقطه شروع متفاوت و $h = ۰/۳$ در مثال ۵.۱.۲.

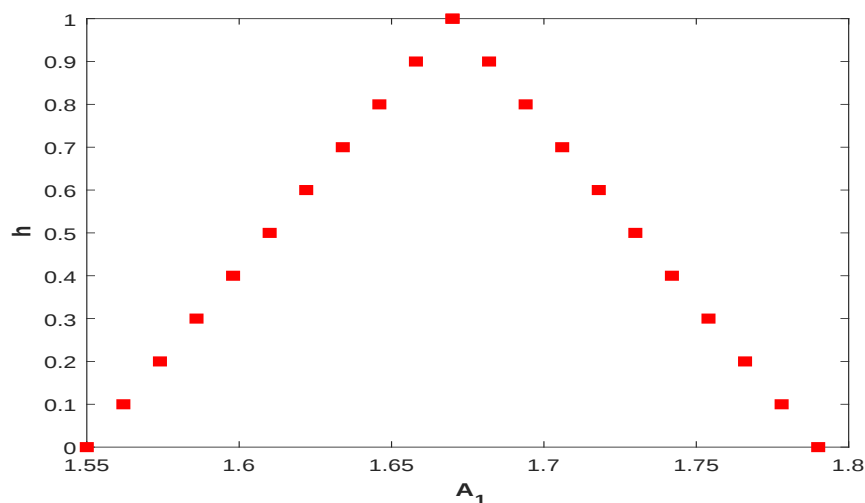
نیستند و مبهم به نظر می‌رسد. در این مثال، یک کاربرد اقتصادی از مدل رگرسیون فازی ارائه می‌شود. داده‌های عرضه بازار برای یک متغیر فازی وابسته Y که نرخ مشارکت در نیروی کار است و دو متغیر مستقل x_1 و x_2 که به ترتیب تولید ناخالص داخلی و دستمزد را نشان می‌دهند، در جدول ۱۳.۲ داده شده است.

جدول ۱۳.۲: مقادیر مشاهده شده و خطای به‌دست آمده در مثال ۶.۱.۲

خطای به‌دست آمده		مقادیر مشاهده شده			
مدل	مدا ارائه شده در [۶۵]	Y	x_2	x_1	Year
۳/۹۴۲۵	۳/۹۴۲۵	$(۳۷/۶۹۹, ۳۸/۶۹۹, ۳۹/۱۹۹)_T$	۱۹۰۵۰	۱۸۰۸۲۲	۱۹۸۸
۳/۱۳۸۳	۳/۱۳۸۳	$(۳۷/۲۱۶, ۳۸/۱۱۶, ۳۸/۹۱۶)_T$	۱۹۲۶۴	۱۸۹۵۹۷	۱۹۸۹
۲/۷۵۱۹	۲/۷۵۱۹	$(۳۶/۷۹۹, ۳۸/۲۹۹, ۳۹/۲۹۹)_T$	۲۰۶۱۱	۲۱۸۵۳۸	۱۹۹۰
۳/۴۶۸۶	۳/۴۶۸۶	$(۳۶/۹۹۹, ۳۸/۰۹۹, ۳۹/۲۹۹)_T$	۲۲۷۷۸	۲۴۵۰۳۶	۱۹۹۱
۲/۸۹۳۴	۲/۸۹۳۴	$(۳۶/۴۹۹, ۳۷/۶۹۹, ۳۸/۸۹۹)_T$	۲۲۸۵۹	۲۵۴۰۵۹	۱۹۹۲
۴/۴۱۴۴	۴/۴۱۴۴	$(۳۶/۲۹۹, ۳۷/۱۹۹, ۳۷/۶۹۹)_T$	۲۴۸۶۶	۲۵۸۶۰۱	۱۹۹۳
۴/۱۲۶۳	۴/۱۲۶۳	$(۳۵/۹۹۹, ۳۶/۷۹۹, ۳۷/۵۹۹)_T$	۲۵۴۴۵	۲۵۹۸۷۶	۱۹۹۴
۳/۸۷۷۴	۳/۸۷۷۴	$(۳۵/۹۰۰, ۳۶/۴۰۰, ۳۷/۴۰۰)_T$	۲۴۶۲۴	۲۶۷۵۳۴	۱۹۹۵
۵/۴۱۴۲	۵/۴۱۴۲	$(۳۴/۸۰, ۳۵/۳۰۰, ۳۶/۹۰۰)_T$	۲۴۷۴۲	۲۸۳۸۰۶	۱۹۹۶
۲/۱۲۶۷	۵/۴۲۹۷	$(۳۴/۸۰۰, ۳۵/۷۰۰, ۳۶/۳۵)_T$	۲۳۹۴۸	۲۹۱۷۶۸	۱۹۹۷
۱/۵۷۹۳	۴/۵۱۰۷	$(۳۵/۲۰۰, ۳۶/۲۰۰, ۳۷/۲۰۰)_T$	۲۴۲۹۴	۳۰۰۱۳۹	۱۹۹۸
۲/۲۹۰۷	۵/۶۳۰۷	$(۳۶/۱۹۹, ۳۶/۶۹۹, ۳۷/۴۹۹)_T$	۲۲۵۷۶	۳۰۴۹۴۱	۱۹۹۹
۱/۰۶۹۶	۴/۵۱۳۹	$(۳۶/۲۰۰, ۳۷/۱۰۰, ۳۷/۶۰۰)_T$	۲۴۶۳۸	۳۲۰۰۶۹	۲۰۰۰
۰/۸۳۸۰	۴/۵۹۲۲	$(۳۷/۰۹۹, ۳۷/۵۹۹, ۳۸/۰۹۹)_T$	۲۴۷۸۰	۳۳۰۵۶۵	۲۰۰۱
۰/۸۲۶۱	۴/۸۲۸۶	$(۳۷/۴۹۹, ۳۸/۰۹۹, ۳۸/۵۹۹)_T$	۲۵۶۹۱	۳۵۵۵۵۴	۲۰۰۲
۰/۸۳۴۶	۵/۰۰۱۸	$(۳۷/۸۹۹, ۳۸/۴۹۹, ۳۹/۰۹۹)_T$	۲۶۰۱۳	۳۷۹۸۳۷	۲۰۰۳
۱/۲۹۳۱	۵/۶۴۳۴	$(۳۸/۳۹۹, ۳۸/۹۹۹, ۳۹/۵۹۹)_T$	۲۶۶۵۳	۳۹۸۲۳۴	۲۰۰۴
۱/۰۰۱۱	۵/۱۶۴۷	$(۳۸/۴۰۰, ۳۹/۲۰۰, ۳۹/۸۰۰)_T$	۲۶۳۱۷	۴۱۹۷۰۵	۲۰۰۵
۰/۶۲۷۴	۵/۱۳۶۳	$(۳۸/۶۹۹, ۳۹/۳۹۹, ۴۰/۰۹۹)_T$	۲۶۵۰۱	۴۴۵۷۹۰	۲۰۰۶
۱/۲۰۷۹	۴/۴۴۶۳				MSPE



(ا) تقریب ضریب A_0



(ب) تقریب ضریب A_1

شکل ۱۶.۲: تقریب ضرایب مدل در مثال ۵.۱.۲

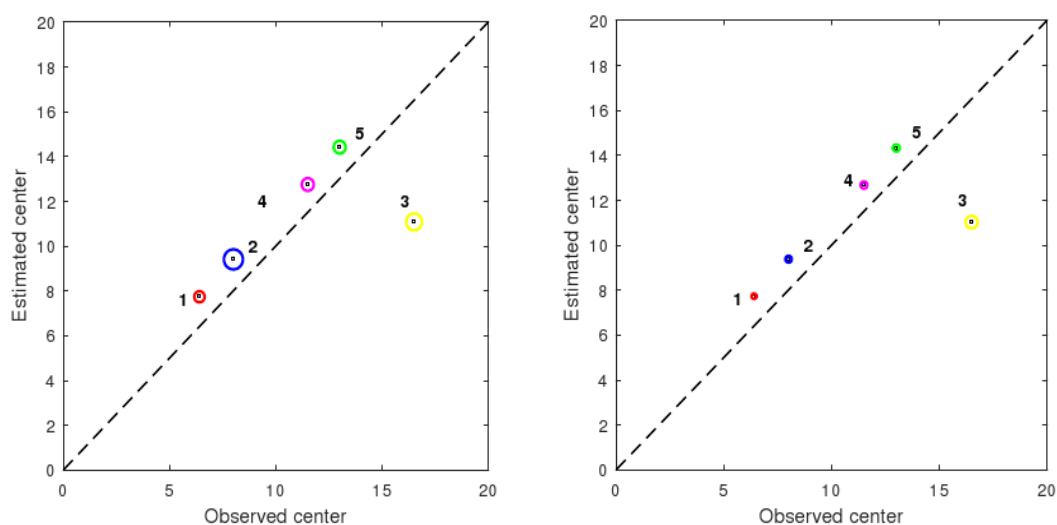
هدف این است که با استفاده از این داده‌ها، یک معادله رگرسیون فازی به صورت زیر ساخته شود:

$$Y = A_0 + A_1 x_1 + A_2 x_1^2 \quad (31.2)$$

شکل ۱۹.۲ مسیرهای شبکه عصبی (۱۹.۲) و (۲۰.۲) را نشان می‌دهد. شکل ۱۸.۲ و جدول ۱۴.۲ به ترتیب ضرایب رگرسیون A_0 ، A_1 و A_2 برای مقادیر مختلف h و نتایج عددی متناظر با آن را نشان می‌دهند.

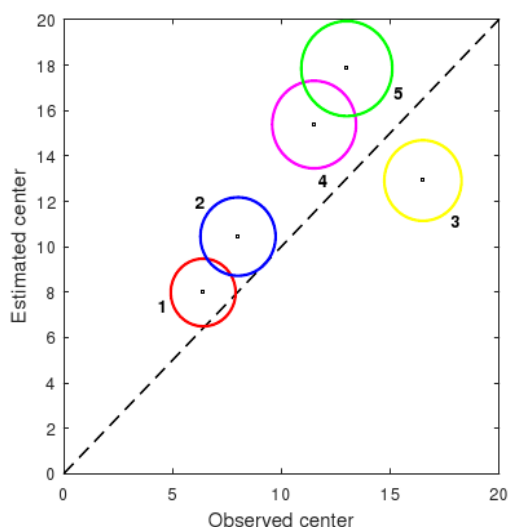
مدل رگرسیون فازی به صورت زیر به دست می‌آید:

$$\begin{aligned} \hat{Y} = & (36/576, 38/260, 39/392)_T + (0/178, 0/183, 0/183)_T \times 10^{-4} x_1 \\ & + (-0/89, -0/106, -0/110)_T \times 10^{-7} x_1^2 \end{aligned}$$



(ب) مدل ارائه شده در [۲۹]

(آ) مدل (۱۹.۲) و (۲۰.۲)



(ج) مدل ارائه شده در [۷۷]

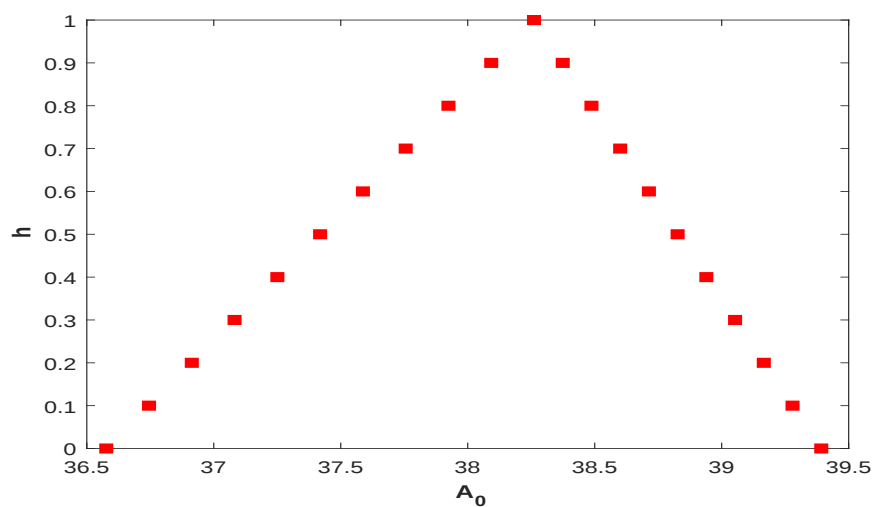
شکل ۱۷.۲: نمودار حبابی فازی در مثال ۵.۱.۲.

خطای به دست آمده از حل این مسئله با به کارگیری روش ارائه شده در [۶۵] و با ۱۵ تکرار از الگوریتم یادگیری در جدول ۱۳.۲ نشان داده شده است. خطای محاسبه شده از روش پیشنهادی برابر است با $۱/۲۰۷۹$ که مشخصاً بهتر است از $۴/۴۴۶۳$ ، که مقدار خطای به دست آمده از حل مدل با روش ارائه شده در [۶۵] است.

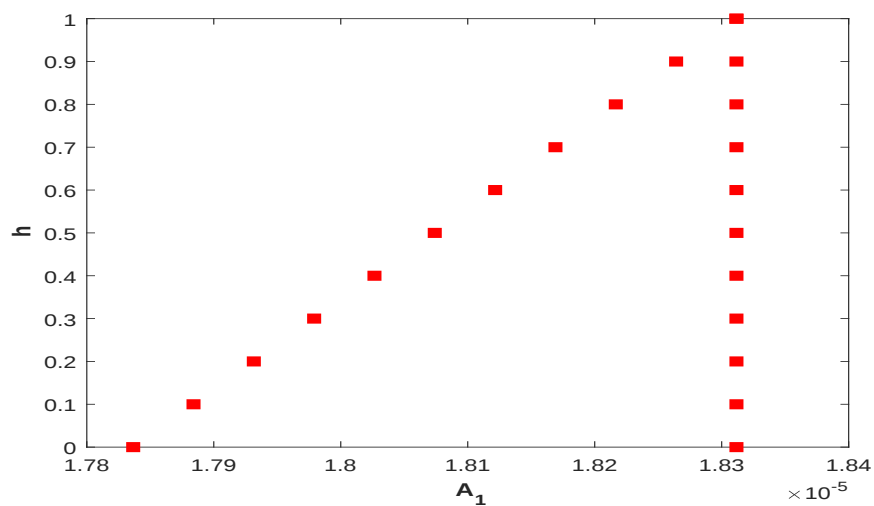
۶.۱.۲ مزایای روش پیشنهادی

قبل از پایان دادن به این بخش، برخی از مزایای روش پیشنهادی را به شرح زیر خلاصه می‌کنیم:

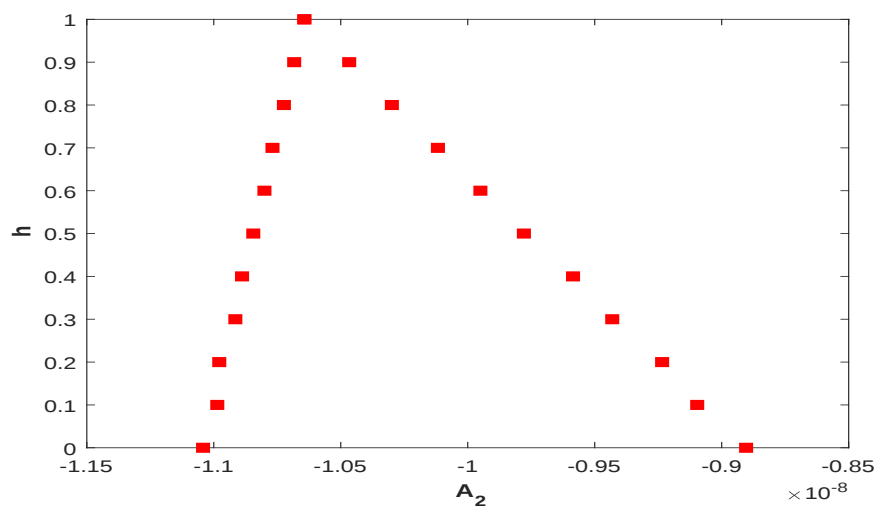
- یک شبکه عصبی مناسب برای پیاده سازی سخت‌افزاری پیشنهاد شده است که جواب خوبی برای مسئله رگرسیون فازی ارائه می‌دهد.



(ا) تقریب ضریب A_0

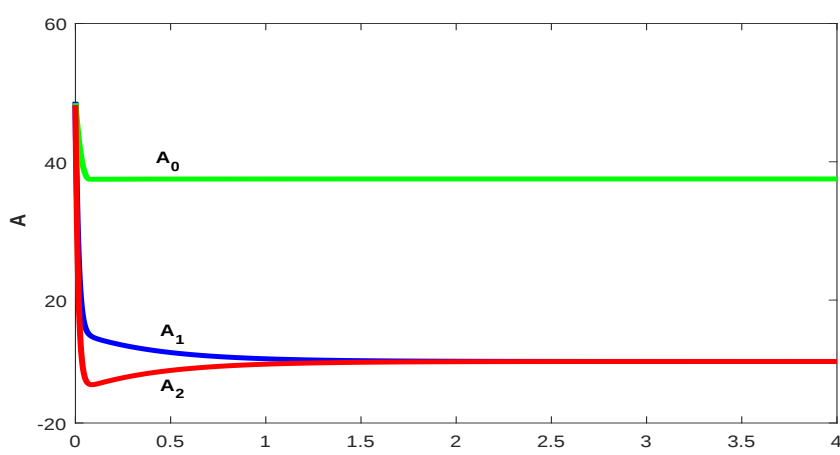


(ب) تقریب ضریب A_1



(ج) تقریب ضریب A_2

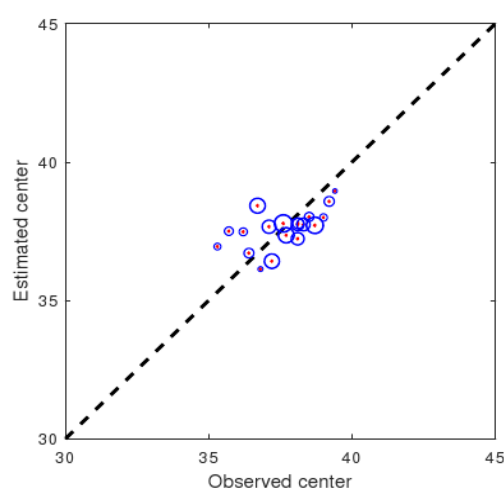
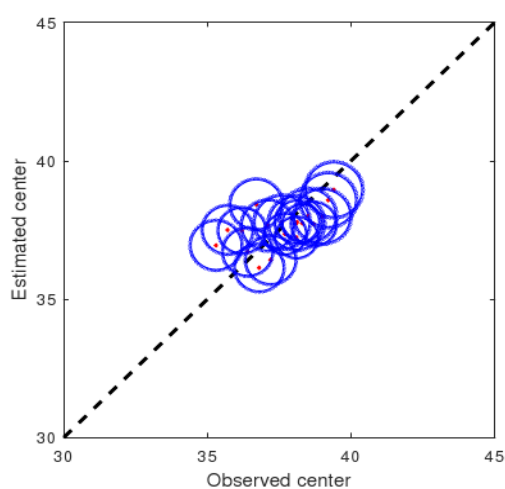
شکل ۱۸.۲: تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۶.۱.۲



شکل ۱۹.۲: رفتار مسیر A از شبکه عصبی (۱۹.۲) و (۲۰.۲) با $h = 0.5$ در مثال ۶.۱.۲.

جدول ۱۴.۲: تقریب ضرایب مدل برای مقادیر مختلف h در مثال ۶.۱.۲

h	$A_0[h]$	$A_1[h]$	$A_2[h]$
۰	$[36/57660, 39/39205]$	$[0/1783654, 0/1831162] \times 10^{-4}$	$[-0/8903841, -0/11042224] \times 10^{-8}$
۰/۱	$[36/74503, 39/27893]$	$[0/1788404, 0/1831162] \times 10^{-4}$	$[-0/9096402, -0/1098623] \times 10^{-8}$
۰/۲	$[36/91346, 39/16580]$	$[0/1793154, 0/1831162] \times 10^{-4}$	$[-0/9234589, -0/1097843] \times 10^{-8}$
۰/۳	$[37/08189, 39/05268]$	$[0/1797904, 0/1831162] \times 10^{-4}$	$[-0/9430992, -0/1091507] \times 10^{-8}$
۰/۴	$[37/25032, 38/93956]$	$[0/1802654, 0/1831162] \times 10^{-4}$	$[-0/9588354, -0/1088869] \times 10^{-8}$
۰/۵	$[37/41875, 38/82641]$	$[0/1807403, 0/1831162] \times 10^{-4}$	$[-0/9778907, -0/1084445] \times 10^{-8}$
۰/۶	$[37/58720, 38/71338]$	$[0/1812159, 0/1831162] \times 10^{-4}$	$[-0/9950173, -0/1080057] \times 10^{-8}$
۰/۷	$[37/75562, 38/60025]$	$[0/1816908, 0/1831162] \times 10^{-4}$	$[-0/1011733, -0/1076884] \times 10^{-8}$
۰/۸	$[37/92406, 38/48715]$	$[0/1821660, 0/1831162] \times 10^{-4}$	$[-0/1029877, -0/1072389] \times 10^{-8}$
۰/۹	$[38/09249, 38/37404]$	$[0/1826411, 0/1831162] \times 10^{-4}$	$[-0/1046680, -0/1068336] \times 10^{-8}$
۱	$[38/26093, 38/26093]$	$[0/1831162, 0/1831162] \times 10^{-4}$	$[-0/1064342, -0/1064343] \times 10^{-8}$



(ب) مدل ارائه شده در [۶۵] (آ) مدل (۱۹.۲) و (۲۰.۲)
شکل ۲۰.۲: نمودار حبابی فازی در مثال ۶.۱.۲.

- شبکه عصبی بازگشتی پیشنهادی برای حل مسئله رگرسیون فازی به پارامتر جریمه نیاز ندارد.
- ثابت شده است که نقطه تعادل شبکه عصبی پیشنهادی، معادل با نقطه کروش-کان-تاکر از مدل بهینه سازی درجه دوم است.
- با ساخت یک تابع لیپانوف مناسب، شرایط کافی برای اطمینان از وجود و پایداری مجانبی سراسری نقطه تعادل یکتای شبکه عصبی بدست می‌آید.
- ثابت شده است که شبکه عصبی پیشنهادی همگرای سراسری به جواب مسئله می‌باشد. بنابراین، همگرایی روش به انتخاب نقطه اولیه مناسب بستگی ندارد.
- بار محاسباتی را می‌توان با استفاده از روش پیشنهادی در مقایسه با روش‌های موجود تا حد زیادی کاهش داد زیرا برای بدست آوردن وزن‌ها (ضرایب)، محاسبه معکوس ماتریس در هر تکرار را نیاز نداریم (برای مثال به [۷۹، ۸۰] مراجعه شود).
- به منظور ارزیابی مدل پیشنهادی، در حالی که برخی از نویسندگان فقط از یک فاصله استفاده می‌کنند (برای مثال به [۲۹، ۶۵، ۷۴] مراجعه شود)؛ در این فصل برای بررسی هرچه دقیق‌تر مدل از سه فاصله مختلف استفاده شده است.
- برای مدل سازی ریاضی مسئله، یک مدل بهینه سازی مقید (برای در نظر گرفتن مثبت بودن پهنای اعداد فازی استفاده شده) ایجاد شده است. در صورتی که نویسندگان در مدل‌های دیگر (برای مثال [۲۹، ۶۵، ۷۴]) معمولاً یک مسئله بهینه سازی نامقید (بدون در نظر گرفتن شرایط لازم برای کار با اعداد فازی) را بررسی می‌کنند. تفاوت در مثال‌های کاربردی به وضوح نشان داده شده است.
- دو مثال جهان واقعی به نام‌های «سطح مقرون به صرفه قیمت خانه‌ها» و «تأمین نیروی کار» بررسی شده است که نشان دهنده این است که روش پیشنهادی به طور کامل قادر به حل مسائل دنیای واقعی در مهندسی و علوم است.
- برای نشان دادن برخی از مزایای مدل پیشنهادی نسبت به دیگر مدل‌ها، از نمودار آماری شناخته شده ای به نام «نمودار حبایی فازی» استفاده شده است.

۲.۲ یک روش شبکه عصبی برای حل مدل رگرسیون چندجمله‌ای فازی

در بخش‌های ۲.۱.۲ و ۳.۱.۲ به حل مدل رگرسیون چندجمله‌ای در دو حالت خاص پرداختیم. اکنون با توجه به مفروضات ارائه شده در ۵.۱.۱ مدل رگرسیون چندجمله‌ای فازی زیر را در نظر می‌گیریم:

$$Y_i = A_0 + \sum_{j=1}^p A_j x_{ij} + \sum_{j=1}^p \sum_{k=1}^p A_{jk} x_{ij} x_{ik} + \dots, \quad i = 1, 2, \dots, n \quad (32.2)$$

مدل (۳۲.۲) را می‌توان به صورت زیر بازنویسی کرد:

$$Y_i = A_0 + A_1 z_{i1} + A_2 z_{i2} + \dots + A_{p'} z_{ip'} \quad i = 1, 2, \dots, n \quad (33.2)$$

که در آن p' تعداد متغیرهای مدل جدید را نشان می‌دهد. برای مثال در حالتی که مدل از مرتبه یک است، با قرار دادن $z_{ij} = x_i^j$ برای $j = 0, 1, \dots, p$ مدل رگرسیون چندجمله‌ای فازی به یک مدل رگرسیون خطی فازی تبدیل می‌شود (در این حالت $p' = p$). برای زمانی که مدل از مرتبه دو باشد، با قرار دادن $z_{ij} = x_{ij}$ برای $j = 0, 1, \dots, p$ و

$$\begin{aligned} z_{i(p+1)} &= x_i^2 \\ z_{i(2p+1)} &= x_i^1 x_i^2, \quad z_{i(p+2)} = x_i^2 \\ z_{i(2p+2)} &= x_i^1 x_i^3, \quad z_{i(3p+1)} = x_i^2 x_i^3, \dots \\ &\vdots \qquad \qquad \qquad \vdots \qquad \qquad \vdots \\ z_{i(3p-1)} &= x_i^1 x_{ip}, \quad z_{i(4p-2)} = x_i^2 x_{ip}, \dots, z_{ip'} = x_{ip}^2 \end{aligned}$$

می‌توانیم مدل رگرسیون چندجمله‌ای فازی با مرتبه دو را به یک مدل رگرسیون خطی فازی تبدیل کنیم. به آسانی می‌توان این روش را برای مراتب بالاتر بسط داد. در حالت خاص، زمانی که هیچ تعاملی بین متغیرها نباشد مدل رگرسیون چندجمله‌ای فازی همان مدل خطی فازی ارائه شده در (۷.۱) می‌باشد.

۱.۲.۲ مدل بهینه سازی

با روشی مشابه روش ارائه شده در ۱.۲.۱ و با استفاده از تعریف فاصله ژو در ۳.۱.۱ تابع خطایی که باید کمینه شود به صورت زیر نوشته می‌شود:

$$C(\mathbf{A}) = \sum_{i=1}^n D_{\Psi}^{\gamma}(\bar{Y}_i[h]_i, Y_i[h])$$

که در آن

$$D_{\mathfrak{Y}}(\bar{Y}_i[h], Y_i[h]) = \left(\int_{\circ}^1 f(h) d^{\mathfrak{Y}}(\bar{Y}_i[h], Y_i[h]) dh \right)^{1/2}, \quad i = 1, 2, \dots, n$$

و $d(\bar{Y}_i[h], Y_i[h])$ برای $i = 1, 2, \dots, n$ ، فاصله میان h -برش‌های عدد فازی $\bar{Y}_i$ و Y_i است. بنابراین مسئله بهینه سازی زیر را داریم:

$$\text{minimize } C(\mathbf{A}) = \sum_{i=1}^n \int_{\circ}^1 f(h) d^{\mathfrak{Y}}(\bar{Y}_i[h], Y_i[h]) dh$$

اکنون با استفاده از تعریف ۷.۱.آ و همچنین با فرض $Y_i = (y_{im}, y_{il}, y_{ir})_T$ و $A_j = (a_{jm}, a_{jl}, a_{jr})_T$ برای $i = 1, \dots, n$ و $j = \circ, 1, \dots, p'$ ، مسئله رگرسیون چندجمله‌ای فازی به شکل زیر تبدیل می‌شود:

$$\bar{Y}_i = \left(\sum_{j=\circ}^{p'} z_{ij} a_{jm}, \sum_{j=\circ}^{p'} z_{ij} a_{jl}, \sum_{j=\circ}^{p'} z_{ij} a_{jr} \right)_T$$

با توجه به نتیجه ۲.۱.آ، مسئله بهینه سازی زیر بدست می‌آید:

(۳۴.۲)

$$\begin{aligned} \text{minimize } C(\mathbf{A}) = & \sum_{i=1}^n \left(\sum_{j=\circ}^{p'} z_{ij} a_{jm} - y_{im} \right)^2 \\ & + \frac{1}{\mathfrak{Y}} \sum_{i=1}^n \left(\sum_{j=\circ}^{p'} z_{ij} a_{jm} - y_{im} \right) \left[\left(\sum_{j=\circ}^{p'} z_{ij} a_{jr} - y_{ir} \right) - \left(\sum_{j=\circ}^{p'} z_{ij} a_{jl} - y_{il} \right) \right] \\ & + \frac{1}{\mathfrak{Y}} \sum_{i=1}^n \left[\left(\sum_{j=\circ}^{p'} z_{ij} a_{jl} - y_{il} \right)^2 + \left(\sum_{j=\circ}^{p'} z_{ij} a_{jr} - y_{ir} \right)^2 \right] \end{aligned}$$

که در آن

$$\mathbf{A} = (a_{\circ m}, a_{1m}, \dots, a_{p'm}, a_{\circ l}, a_{1l}, \dots, a_{p'l}, a_{\circ r}, a_{1r}, \dots, a_{p'r})^T \quad (۳۵.۲)$$

اکنون با فرض $x_{ij} \leq 0$ (در نتیجه $z_{ij} \leq 0$) برای $i = 1, \dots, n$ و $j = 0, 1, \dots, p'$ (۳۴.۲) به شکل تابع درجه دوم زیر بازنویسی می شود:

$$\begin{aligned} C(\mathbf{A}) &= \sum_{i=1}^n \left(\sum_{j=0}^{p'} z_{ij} a_{jm} - y_{im} \right)^2 \\ &+ \frac{1}{\varphi} \sum_{i=1}^n \left[\left(\sum_{j=0}^{p'} z_{ij} a_{jm} - y_{im} \right) \left(\sum_{j=0}^{p'} z_{ij} a_{jr} - y_{ir} \right) \right] + \frac{1}{\varphi} \sum_{i=1}^n \left(\sum_{j=0}^{p'} z_{ij} a_{jl} - y_{il} \right)^2 \\ &- \frac{1}{\varphi} \sum_{i=1}^n \left[\left(\sum_{j=0}^{p'} z_{ij} a_{jm} - y_{im} \right) \left(\sum_{j=0}^{p'} z_{ij} a_{jl} - y_{il} \right) \right] + \frac{1}{\varphi} \sum_{i=1}^n \left(\sum_{j=0}^{p'} z_{ij} a_{jr} - y_{ir} \right)^2 \\ &= \sum_{i=1}^n \left[\left(\sum_{j=0}^{p'} z_{ij} a_{jm} \right)^2 - 2 y_{im} \sum_{j=0}^{p'} z_{ij} a_{jm} + y_{im}^2 \right] \\ &+ \frac{1}{\varphi} \sum_{i=1}^n \left[\left(\sum_{j=0}^{p'} z_{ij} a_{jm} - y_{im} \right) \left(\sum_{j=0}^{p'} z_{ij} a_{jr} - y_{ir} \right) \right] \\ &+ \frac{1}{\varphi} \sum_{i=1}^n \left[\left(\sum_{j=0}^{p'} z_{ij} a_{jl} \right)^2 - 2 y_{il} \sum_{j=0}^{p'} z_{ij} a_{jl} + y_{il}^2 \right] \\ &- \frac{1}{\varphi} \sum_{i=1}^n \left[\left(\sum_{j=0}^{p'} z_{ij} a_{jm} - y_{im} \right) \left(\sum_{j=0}^{p'} z_{ij} a_{jl} - y_{il} \right) \right] \\ &+ \frac{1}{\varphi} \sum_{i=1}^n \left[\left(\sum_{j=0}^{p'} z_{ij} a_{jr} \right)^2 - 2 y_{ir} \sum_{j=0}^{p'} z_{ij} a_{jr} + y_{ir}^2 \right] \\ &= \mathbf{A}^\top \mathbf{Q} \mathbf{A} + \mathbf{B}^\top \mathbf{A} + M \end{aligned}$$

که در آن

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q} & -\frac{1}{\varphi} \mathbf{q} & \frac{1}{\varphi} \mathbf{q} \\ \mathbf{O}_{(p'+1) \times (p'+1)} & \frac{1}{\varphi} \mathbf{q} & \mathbf{O}_{(p'+1) \times (p'+1)} \\ \mathbf{O}_{(p'+1) \times (p'+1)} & \mathbf{O}_{(p'+1) \times (p'+1)} & \frac{1}{\varphi} \mathbf{q} \end{bmatrix} \quad (36.2)$$

و $\mathbf{O}$ یک ماتریس صفر است.

$$\mathbf{B} = (b_0, b_1, \dots, b_{p'}, b_{p'+1}, \dots, b_{2p'}, \dots, b_{2p'+2}, \dots, b_{3p'}, \dots, b_{3p'+3})^\top \quad (37.2)$$

$$M = \sum_{i=1}^n \left(y_{im}^2 + y_{il}^2 + y_{ir}^2 - \frac{1}{\varphi} y_{im} y_{il} + \frac{1}{\varphi} y_{im} y_{ir} \right) \quad (38.2)$$

$$\mathbf{q} = \begin{bmatrix} n & \sum_{i=1}^n z_{i1} & \dots & \sum_{i=1}^n z_{ip'} \\ \sum_{i=1}^n z_{i1} & \sum_{i=1}^n z_{i1}^2 & \dots & \sum_{i=1}^n z_{i1} z_{ip'} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n z_{ip'} & \sum_{i=1}^n z_{ip'} z_{i1} & \dots & \sum_{i=1}^n z_{ip'}^2 \end{bmatrix}$$

$$b_t = \begin{cases} \sum_{i=1}^n x_{ij} \left(-\frac{1}{2} y_{im} + \frac{1}{3} y_{il} - \frac{1}{3} y_{ir} \right) & \text{برای } t = 0, 1, \dots, p' \\ \sum_{i=1}^n x_{ij} \left(\frac{1}{3} y_{im} - \frac{1}{6} y_{il} \right) & \text{برای } t = p' + 1, \dots, 2p' + 2 \\ \sum_{i=1}^n x_{ij} \left(-\frac{1}{3} y_{im} - \frac{1}{6} y_{ir} \right) & \text{برای } t = 2p' + 3, \dots, 3p' + 3 \end{cases}$$

دیده می‌شود که ماتریس هسین $C(A)$ برابر است با Q که $\nabla^2 C(A) = \nabla(\nabla C(A)) = Q$ یک ماتریس معین مثبت است.

با توجه به تعریف اعداد فازی در پیوست آ می‌دانیم که پهنای راست و چپ اعداد فازی $A_0, A_1, \dots, A_{p'}$ باید مثبت باشد بنابراین مسئله بهینه سازی درجه دوم محدب زیر بدست می‌آید:

$$\text{minimize } C(A) = A^T Q A + B^T A + M \quad (39.2)$$

subject to

$$GA \leq 0 \quad (40.2)$$

که در آن

$$G = \begin{bmatrix} 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & -1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & -1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 & -1 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & -1 \end{bmatrix}_{(3p'+3) \times (3p'+3)}$$

A, Q, B و M به ترتیب در (35.2)، (36.2)، (37.2) و (38.2) تعریف شده‌اند.

۲.۲.۲ رویکرد شبکه عصبی بازگشتی

در این قسمت به دنبال معرفی یک شبکه عصبی مناسب برای حل مسئله بهینه سازی درجه دوم محدب (39.2) و (40.2) هستیم. همانند روش اجرا شده در بخش ۴.۳.۱ در این بخش نیز در تلاش می‌کنیم که با استفاده از تکنیک‌های بهینه سازی، مسئله برنامه‌ریزی (39.2) و

(۴۰.۲) را به یک سیستم غیر خطی تبدیل کنیم. برای تمرکز بیشتر بر روی مباحث ارائه شده تمام اثبات های این قسمت به پیوست ب انتقال داده شده است.
 یک جواب بهینه از (۳۹.۲) و (۴۰.۲) است. اگر و تنها اگر وجود داشته باشد $A^* \in \mathbb{R}^{p'+3}$ و $U^* \in \mathbb{R}^{p'+3}$ به طوری که $(A^{*\top}, U^{*\top})^\top$ در شرایط فروش-کان-تاکر زیر صدق کند:

$$\begin{cases} U^* \geq 0, & GA^* \leq 0, & U^{*\top}GA^* = 0 \\ 2QA^* + B + G^\top U^* = 0 \end{cases} \quad (41.2)$$

که در آن A^* نقطه فروش-کان-تاکر مدل (۳۹.۲) و (۴۰.۲) و U^* بردار ضریب لاگرانژ متناظر با A^* است.

فرض کنید $A(\cdot)$ و $U(\cdot)$ دو متغیر وابسته به زمان باشند. آن گاه شبکه عصبی زیر را معرفی می کنیم:

$$\begin{cases} \frac{dy}{dt} = \tau \phi(y) \end{cases} \quad (42.2)$$

$$\begin{cases} y(t_0) = y_0 = (A(t_0)^\top, U(t_0)^\top)^\top \in \mathbb{R}^{p'+6}, \tau > 0 \end{cases} \quad (43.2)$$

که در آن

$$\phi(y) = \begin{pmatrix} -(2QA + B + G^\top(GA)^+) \\ (GA)^+ - U \end{pmatrix} \quad (44.2)$$

و $\tau = 1$.

بررسی همگرایی و پایداری

در این قسمت به بررسی گزاره های مربوط به همگرایی و پایداری مدل (۴۲.۲) و (۴۳.۲) می پردازیم.

قضیه ۱.۲.۲. فرض کنید $y^* = (A^{*\top}, U^{*\top})^\top$ نقطه تعادل مدل شبکه عصبی (۴۲.۲) و (۴۳.۲) باشد. بنابراین A^* یک نقطه فروش-کان-تاکر برای مسئله (۳۹.۲) و (۴۰.۲) است. از سمت دیگر، اگر $A^* \in \mathbb{R}^{p'+3}$ یک جواب مسئله (۳۹.۲) و (۴۰.۲) باشد، بنابراین $U^* \in \mathbb{R}^{p'+3}$ وجود دارد به طوری که $y^* = (A^{*\top}, U^{*\top})^\top$ نقطه تعادل مدل شبکه عصبی (۴۲.۲) و (۴۳.۲) است.

در ادامه ثابت می کنیم که مدل شبکه عصبی پیشنهادی پایدار لیپانوف است.

لم ۱.۲.۲. ژاکوبین ماتریس ϕ که در (۴۴.۲) تعریف شده است، نیمه معین منفی است.

قضیه ۲.۲.۲. تحت مفروضات لم ۱.۲.۲ شبکه عصبی (۴۲.۲) و (۴۳.۲) پایدار لیپانوف است.

برای نشان دادن همگرایی سراسری مدل، لم و قضیه زیر را داریم:

لم ۲.۲.۲. (i) برای هر نقطه ابتدایی $y(t_0) = (A(t_0)^T, U(t_0)^T)^T$ یک جواب پیوسته یکتا $y(t) = (A(t)^T, U(t)^T)^T$ برای سیستم (۴۲.۲) و (۴۳.۲) وجود دارد.
(ii) فرض کنید $y(t) = (A(t)^T, U(t)^T)^T$ یک مسیر از شبکه عصبی (۴۲.۲) و (۴۳.۲) با نقطه ابتدایی $y(t_0) = (A(t_0)^T, U(t_0)^T)^T$ باشد. در این صورت اگر، $U(t_0) \geq 0$ ، بنابراین $U(t) \geq 0$.

قضیه ۳.۲.۲. مسیرهای شبکه عصبی (۴۲.۲) و (۴۳.۲) برای هر نقطه ابتدایی

$$y_0 = (A(t_0)^T, U(t_0)^T)^T \in \mathbb{R}^{p'+6}$$

به نقطه تعادل شبکه عصبی همگرا می‌شود. به طور خاص، شبکه عصبی (۴۲.۲) و (۴۳.۲) با هر نقطه ابتدایی $y_0 = (A(t_0)^T, U(t_0)^T)^T \in \mathbb{R}^{p'+6}$ پایدار مجانبی سراسری است و D^* یک نقطه تعادل یکتا دارد. که در آن D^* نشان دهنده مجموعه نقاط بهینه مسئله اصلی و دوگان است.

قضیه ۴.۲.۲. سرعت همگرایی شبکه عصبی (۴۲.۲) و (۴۳.۲) با افزایش τ افزایش می‌یابد.

۳.۲.۲ مثال‌های عددی

در این قسمت برای بررسی کارایی مدل پیشنهادی در حل مسائل رگرسیون خطی و چندجمله‌ای فازی، به حل چند مثال می‌پردازیم. در بعضی از مثال‌ها، درستی روش پیشنهادی را با چندین نقطه شروع مختلف آزمایش می‌کنیم. همچنین برای تأکید بیشتر بر مدل پیشنهادی، مقایسه‌هایی با مدل‌های موجود در [۲۹، ۶۵، ۷۵-۷۷] به وسیله فاصله‌های فازی D_1 ، D_2 و D_3 انجام شده است.

مثال ۱.۲.۲. داده‌های ارائه شده در جدول ۱۵.۲ را برای یک متغیر فازی پیوسته و دو متغیر مستقل غیر فازی در نظر بگیرید.

جدول ۱۵.۲: مقادیر مشاهده شده در مثال ۱.۲.۲

۶	۵	۴	۲	۱	x_1
۸	۵	۴	۳	۲	x_2
$(6, 13, 17)_T$	$(7, 11, 14)_T$	$(5, 8, 7)_T$	$(3, 6, 8)_T$	$(2, 4, 5)_T$	Y

با استفاده از این داده‌ها در نظر داریم که ضرایب مسئله رگرسیون

$$Y = A_0 + A_1 x_1 + A_2 x_2 \quad (45.2)$$

را با توجه به مدل شبکه عصبی (۴۲.۲) و (۴۳.۲) تقریب بزنیم. واضح است که (۴۵.۲) یک مدل رگرسیون دو متغیره و از مرتبه یک است. پس یک مسئله رگرسیون خطی معمولی داریم.

۷۲ مدل شبکه عصبی بازگشتی برای حل مسائل رگرسیون فازی

شکل ۲۱.۲ نشان می‌دهد که مسیرهای شبکه عصبی (۴۲.۲) و (۴۳.۲) به جواب بهینه مسئله همگرا است. بنابراین، مدل رگرسیون خطی زیر را داریم:

$$\bar{Y} = (1/6861, 1/8327, 1/1841)_T + (1/6731, 1/2260, 0/3514)_{Tx_1} \\ + (-0/7077, 0/4896, 1/7614)_{Tx_2}$$

شکل ۲۲.۲ همگرایی مدل پیشنهادی برای ۵۰ نقطه شروع متفاوت را به تصویر می‌کشد. جدول ۱۶.۲ تفاوت خطای محاسبه شده با استفاده از مدل (۴۲.۲) و (۴۳.۲) و مدل ارائه شده در [۶۵، ۲۹] را نشان می‌دهد.

جدول ۱۶.۲: خطای به‌دست آمده در مثال ۱.۲.۲

خطای به‌دست آمده			
D_2	D_1	D_3	روش / فاصله
0/1506	0/0357	0/0030	مدل ارائه شده در [۶۵، ۲۹]
1/1644	1/8061	0/0872	
3/8811	23/5498	1/2810	
3/3239	17/3727	0/9498	
0/5601	0/4716	0/0250	
1/8160	8/6472	0/4692	خطا
0/1524	0/0377	0/0032	مدل (۴۲.۲) و (۴۳.۲)
1/1662	1/8211	0/0882	
3/8781	23/4431	1/2732	
3/3281	17/4804	0/9584	
0/5523	0/4452	0/0230	
1/8154	8/6455	0/4692	خطا

جدول ۱۷.۲: روش MCV در مثال ۱.۲.۲

MCV			
D_2	D_1	D_3	روش / فاصله
0/3702	0/2297	0/0198	مدل ارائه شده در [۶۵، ۲۹]
1/8536	4/5787	0/2209	
6/8144	72/7682	3/9280	
7/7637	94/8638	5/1903	
16/6623	65/0180	4/1405	
6/6928	47/4917	2/6999	خطا
0/4031	0/2516	0/0211	مدل (۴۲.۲) و (۴۳.۲)
1/8552	4/5852	0/0590	
3/8877	23/4810	1/2723	
7/0229	67/2057	4/4028	
13/4617	56/8305	5/1224	
5/3261	30/4708	2/1755	خطا

واضح است که مدل ارائه شده عملکرد شایسته‌ای بر روی داده‌های مشاهده شده داشته است.

مثال ۲.۲.۲. نویسندگان در [۷۶]، [۲۹] و [۷۷] از داده‌های ارائه شده در جدول ۱۸.۲ برای نشان دادن روش پیشنهادیشان استفاده کرده‌اند.

جدول ۱۸.۲: مقادیر مشاهده شده و خطای به‌دست آمده در مثال ۲.۱.۲

خطای به‌دست آمده				متغیر پاسخ	متغیر مستقل
مدل (۴۲.۲) و (۴۳.۲)	مدل ارائه شده در [۲۹]	مدل ارائه شده در [۷۶]	مدل ارائه شده در [۷۷]		
۱۹/۹۰۸۵	۱۹/۹۰۸۴	۲۱/۲۶۷۱	۶۲/۴۰۷۱	$(۸/۰, ۱/۸, ۱/۸)_T$	۱
۱۹/۹۰۸۵	۱۹/۹۰۸۴	۲۱/۲۶۷۱	۶۲/۴۰۷۱	$(۶/۴, ۲/۲, ۲/۲)_T$	۲
۳/۹۹۸۴	۴/۰۰۱۶	۴/۰۰۲۲	۱۰/۶۶۳۱	$(۹/۵, ۲/۶, ۲/۶)_T$	۳
۳۲/۲۵۶۳	۳۲/۲۲۱۴	۳۲/۱۶۶۷	۲۳/۲۰۳۱	$(۱۳/۵, ۲/۶, ۲/۶)_T$	۴
۲/۹۵۵۸	۲/۹۷۱۹	۲/۹۵۳۵	۲۸/۱۴۲۱	$(۱۳/۰, ۲/۴, ۲/۴)_T$	۵
۱۵/۸۰۱۵	۱۵/۸۰۲۳	۱۶/۳۳۱۳	۳۷/۳۶۴۵	MSPE	

ضرایب رگرسیون به صورت زیر تخمین زده می‌شوند:

$$\bar{Y} = (۴/۹۵۰۰, ۱/۸۴۰۰, ۱/۸۴۰۰)_T + (۱/۷۰۹۹, ۰/۱۶۰۰, ۰/۱۶۰۰)_{Tx}$$

شکل ۲۳.۲ مسیرهای شبکه عصبی که به جواب بهینه مدل همگرا هستند را نشان می‌دهد. در شکل ۲۴.۲ نشان داده شده است که مسیرهای شبکه عصبی پیشنهادی با هر نقطه شروع اولیه دلخواه به جواب بهینه مسئله همگرا می‌شوند. از طرف دیگر، در [۲۹] با ۸ تکرار از الگوریتم یادگیری، مدل رگرسیون زیر بدست می‌آید:

$$\bar{Y} = (۴/۹۴۹۹, ۱/۸۳۹۹, ۱/۸۳۹۸)_T + (۱/۷۱۰۰, ۰/۱۶۰۰, ۰/۱۶۰۱)_{Tx}$$

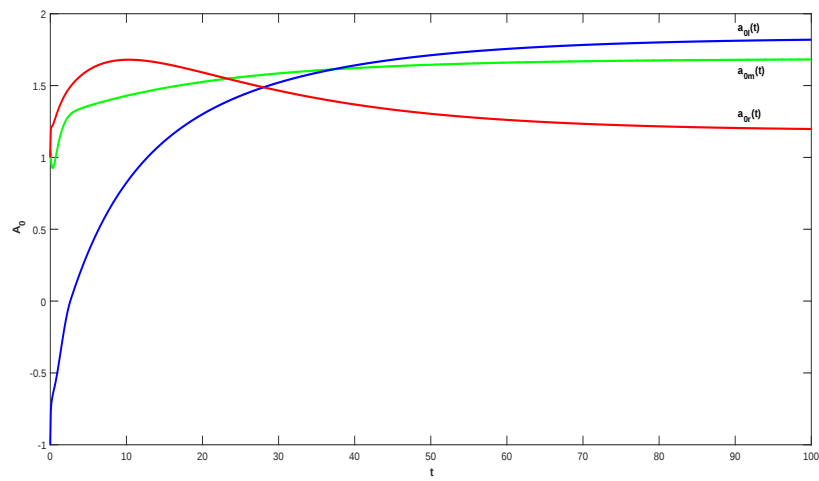
در [۷۷]، نتیجه مسئله کمینه سازی با $h = ۰$ به صورت زیر است:

$$\bar{Y} = ۳/۸۵۰ + ۲/۱۰۰x$$

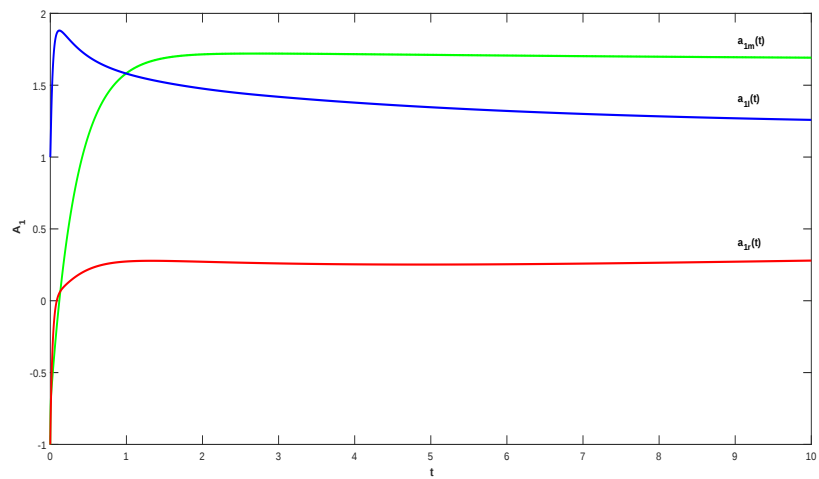
همچنین نویسندگان در [۷۶] مدل زیر را پیشنهاد داده‌اند:

$$\bar{Y} = ۴/۸۰۸ + ۱/۷۱۸x + (۰/۱۱۸, ۲/۳۲, ۲/۳۲)_T$$

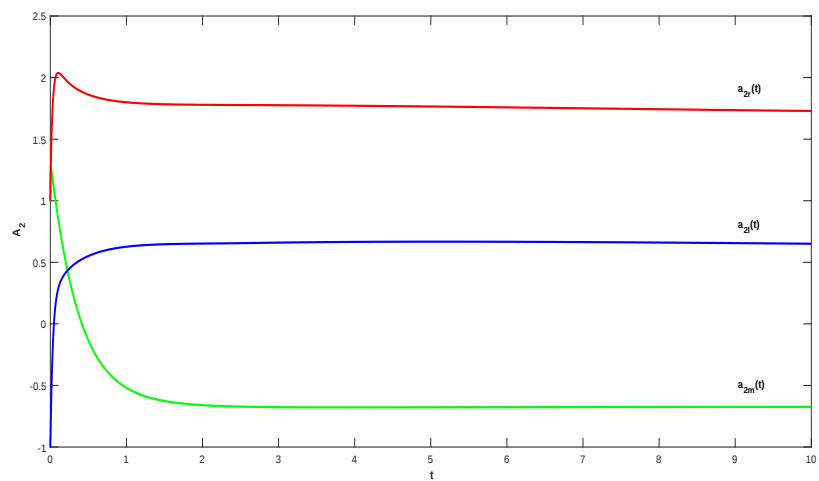
جدول ۱۸.۲ مقایسه ای میان خطاهای به‌دست آمده از این چهار روش را نشان می‌دهد. واضح است که روش شبکه عصبی (۴۲.۲) و (۴۳.۲) تقریب بهتری برای داده ها بوده است.



(ا) مسیر A_0

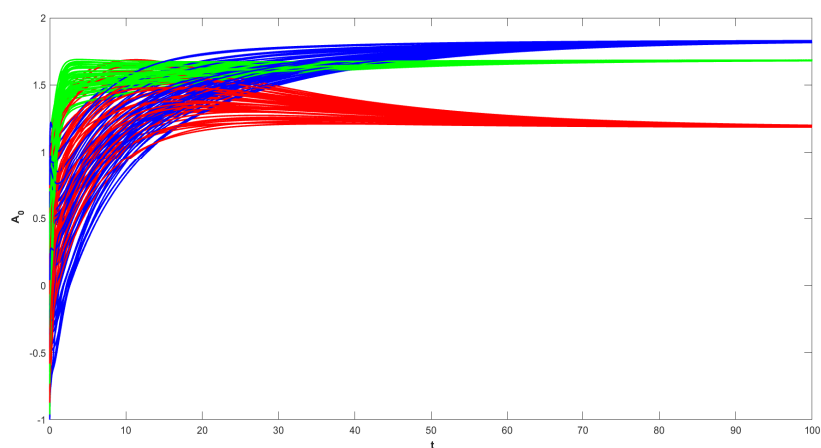


(ب) مسیر A_1

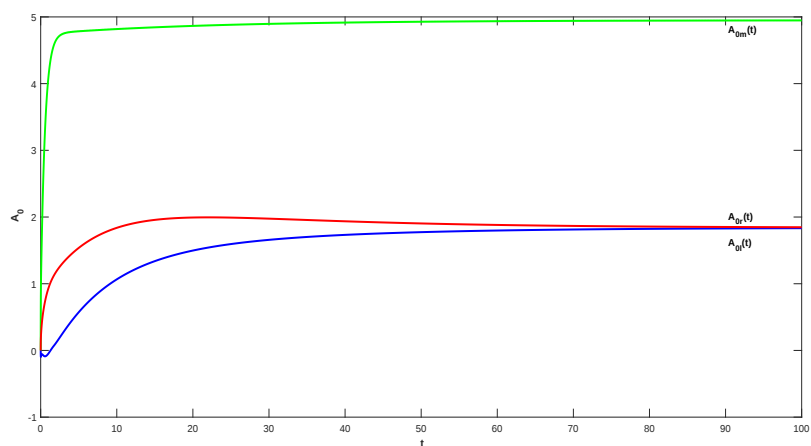


(ج) مسیر A_2

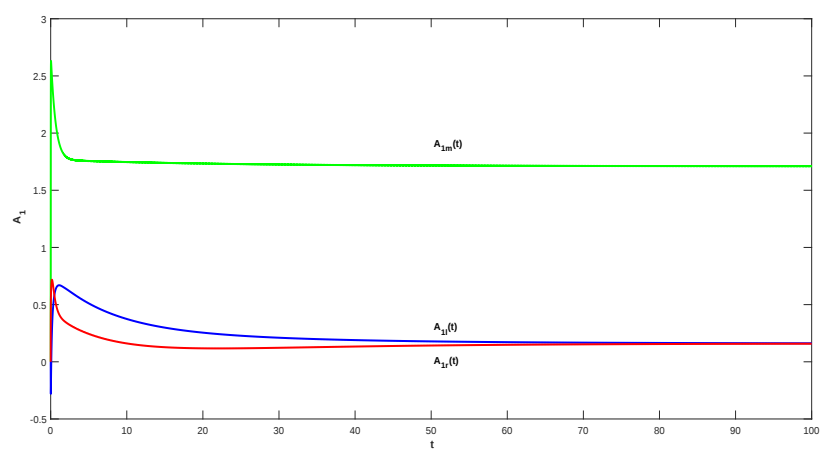
شکل ۲۱.۲: رفتار مسیر A از شبکه عصبی (۴۲.۲) و (۴۳.۲) در مثال ۱.۲.۲.



شکل ۲۲.۲: رفتار مسیر A_0 از شبکه عصبی (۴۲.۲) و (۴۳.۲) با 5° نقطه شروع متفاوت در مثال ۱.۲.۲.

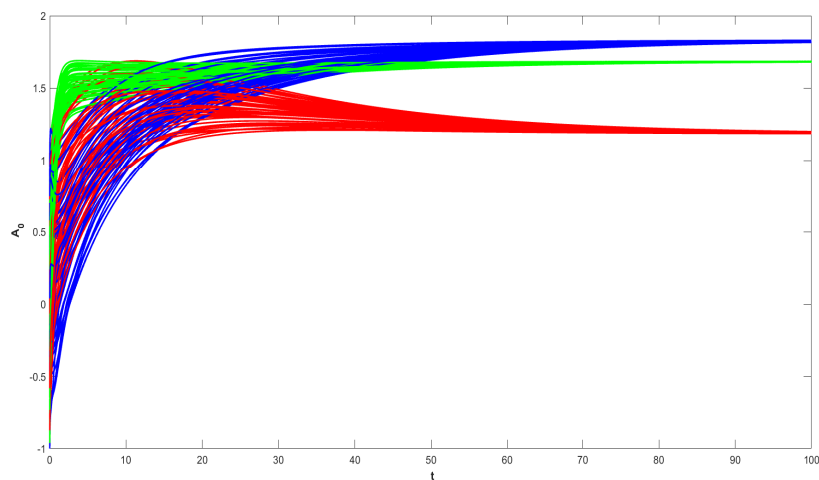


(ا) مسیر A_0



(ب) مسیر A_1

شکل ۲۳.۲: رفتار مسیر A از شبکه عصبی (۴۲.۲) و (۴۳.۲) در مثال ۲.۲.۲.



شکل ۲۴.۲: رفتار مسیر A_1 از شبکه عصبی (۴۲.۲) و (۴۳.۲) با 20° نقطه شروع متفاوت در مثال ۲.۲.۲.

مثال ۳.۲.۲. (سطح مقرون به صرفه قیمت خانه‌ها) تحت مفروضات مثال ۳.۱.۲ و با استفاده از مدل ارائه شده در (۴۲.۲) و (۴۳.۲)، ضرایب مدل رگرسیون خطی (۲۸.۲) را تقریب می‌زنیم. نتایج در جدول ۱۹.۲ نشان داده شده است.

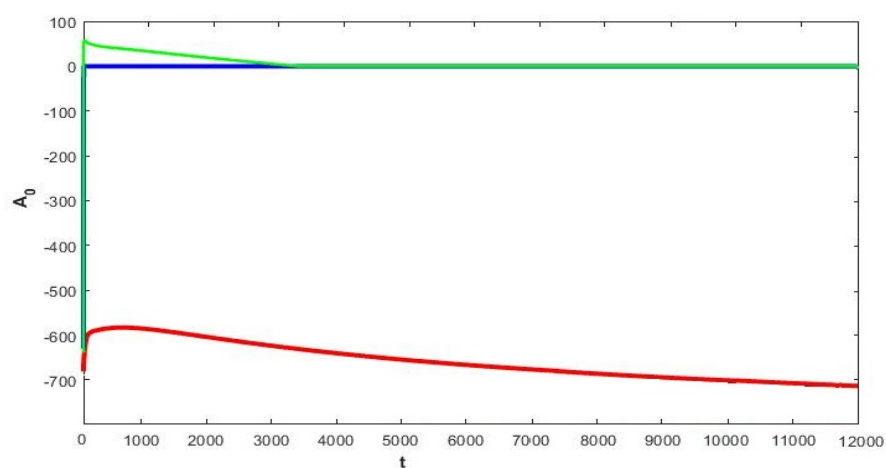
جدول ۱۹.۲: تقریب ضرایب مدل در مثال ۳.۲.۲

روش ارائه شده در [۷۵]	مدل (۴۲.۲) و (۴۳.۲)	
$(13/0.1, 5/29, 10/88)_T$	$(-729/8221, 0, 0)_T$	A_0
$(709/72, 241/56, 95/90)_T$	$(6/5943, 0/6247 \times 10^{-7}, 0/1393 \times 10^{-7})_T$	A_1
$(-20/12, 6/03, 5/15)_T$	$(-17/1313, 0/6162 \times 10^{-7}, 0/6442 \times 10^{-7})_T$	A_2
$(-34/11, 11/54, 6/59)_T$	$(-51/2636, 0/1104 \times 10^{-5}, 0/1265 \times 10^{-5})_T$	A_3
$(-25/7, 3/27, 0/8)_T$	$(-2/4814, 0/6866 \times 10^{-8}, 0/7861 \times 10^{-8})_T$	A_4
$(879/19, 214/27, 155/5)_T$	$(13/3201, 3/9216, 1/7018)_T$	A_5
$(-7/65, 1/87, 3/06)_T$	$(103/2563, 11/0539, 13/8195)_T$	A_6

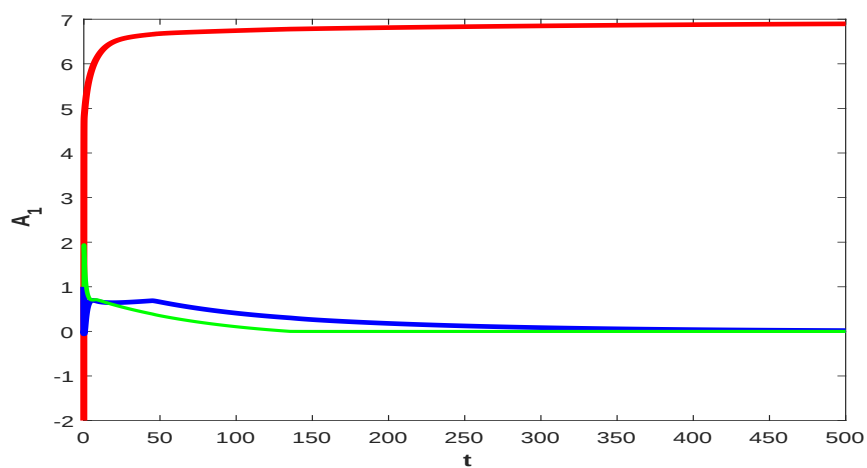
شکل‌های ۲۵.۲-۲۷.۲ مسیرهای همگرایی مدل شبکه عصبی (۴۲.۲) و (۴۳.۲) را نشان می‌دهند. جدول ۲۰.۲ مقایسه‌ای از مقدار خطای به‌دست آمده در روش‌های مختلف را نشان می‌دهد. بنابراین واضح است که مدل پیشنهادی عملکرد بهتری بر روی داده‌ها داشته است.

جدول ۲۰.۲: خطای به‌دست آمده در مثال ۳.۲.۲

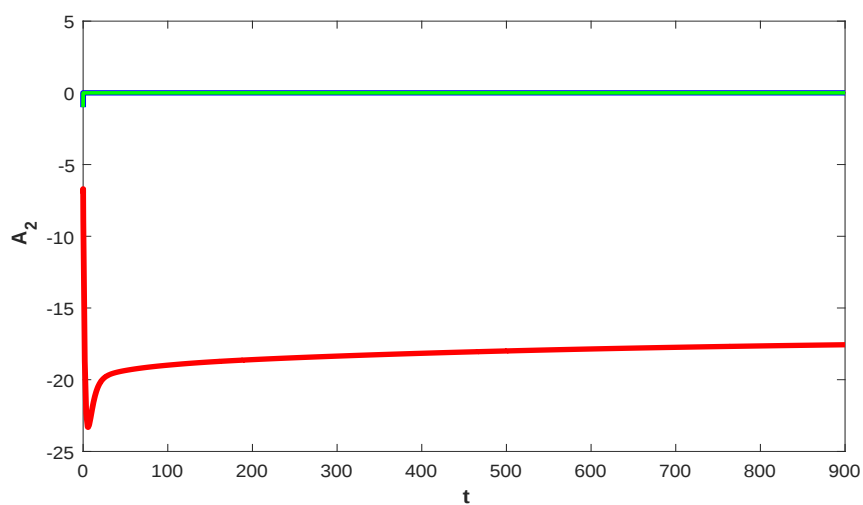
روش	مدل ارائه شده در [۷۵]	مدل (۴۲.۲) و (۴۳.۲)
MSPE	۳۵۷/۶۰۳۱	۱۳/۱۹۰۹



(ا) مسیر A_0

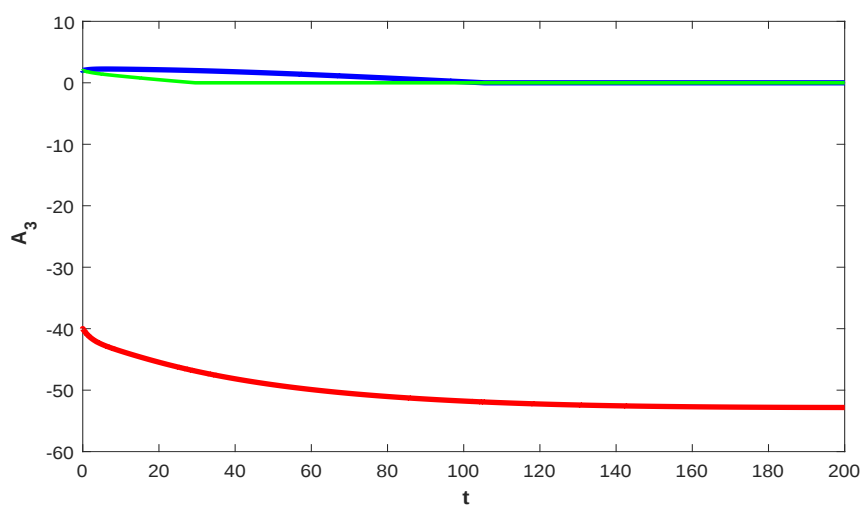


(ب) مسیر A_1

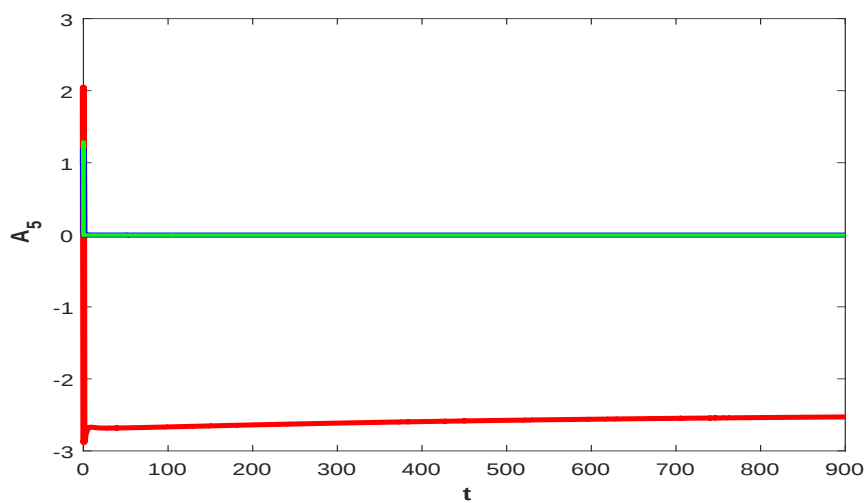


(ج) مسیر A_2

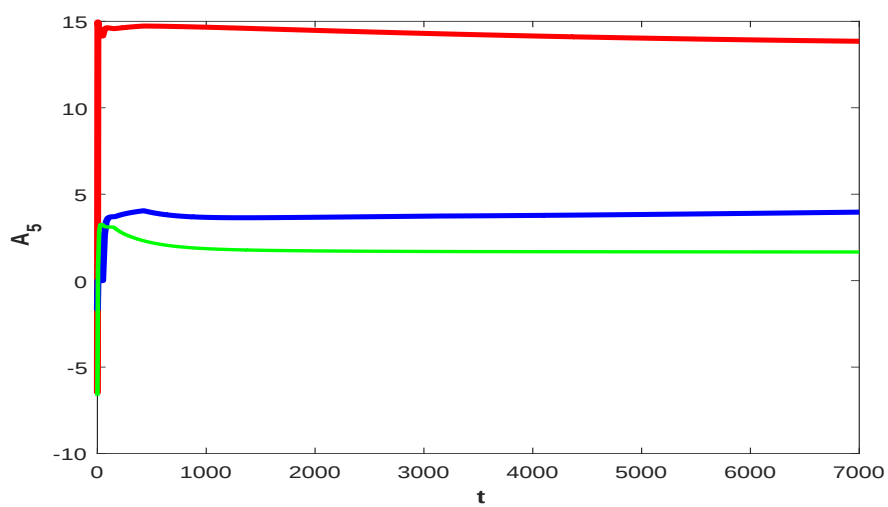
شکل ۲۵.۲: رفتار مسیرهای A_0 ، A_1 و A_2 از شبکه عصبی (۴۲.۲) و (۴۳.۲) در مثال ۳.۲.۲.



(ا) مسیر A_3

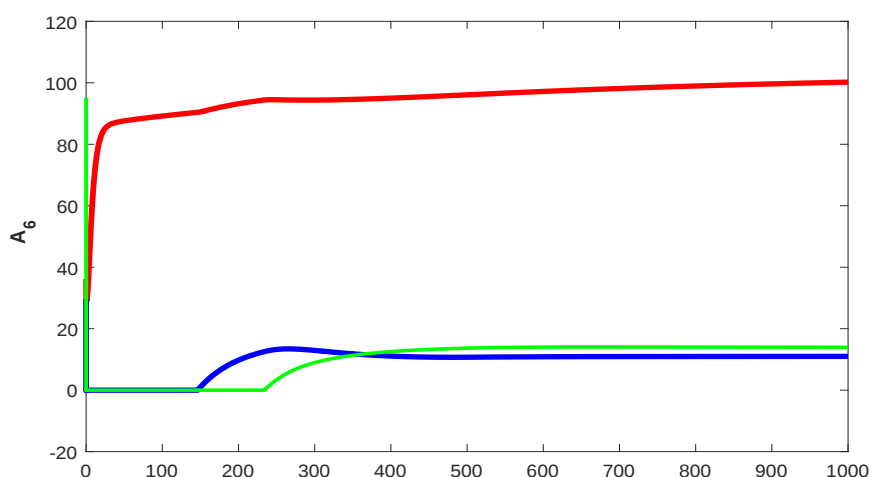


(ب) مسیر A_4



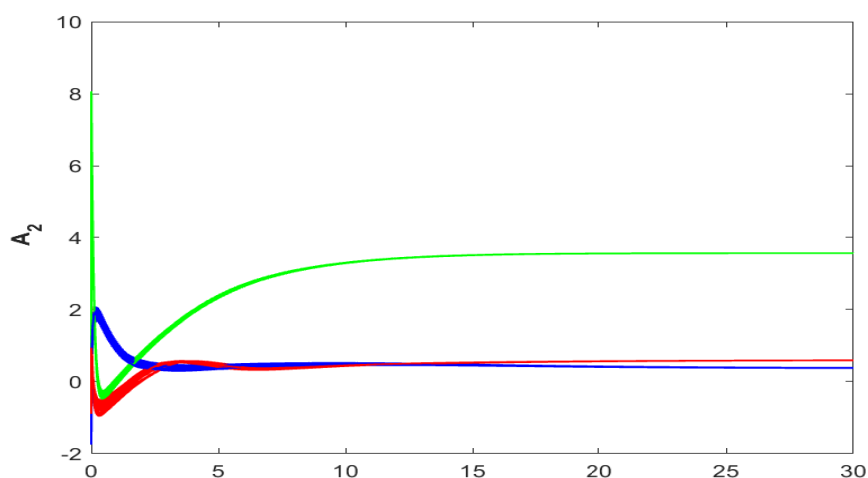
(ج) مسیر A_5

شکل ۲۶.۲: رفتار مسیرهای A_3 ، A_4 و A_5 از شبکه عصبی (۴۲.۲) و (۴۳.۲) در مثال ۳.۲.۲.



شکل ۲۷.۲: رفتار مسیر A_6 از شبکه عصبی (۴۲.۲) و (۴۳.۲) در مثال ۳.۲.۲.

مثال ۴.۲.۲. داده های مثال ۴.۱.۲ را در نظر بگیرید. برای تقریب ضرایب مدل رگرسیون چندجمله‌ای (۲۹.۲) از مدل شبکه عصبی (۴۲.۲) و (۴۳.۲) استفاده می‌کنیم. شکل‌های ۲۸.۲ و ۲۹.۲ رفتار همگرایی مسیرهای مدل شبکه عصبی را نشان می‌دهد.



شکل ۲۸.۲: رفتار مسیر A_2 از شبکه عصبی (۴۲.۲) و (۴۳.۲) با z_0 نقطه شروع متفاوت در مثال ۴.۲.۲.

با فرض $z_1 = x$ و $z_2 = x^2$ ضرایب مدل به صورت زیر تقریب زده می‌شوند:

$$\bar{Y} = (72/4603, 0/2829, 1/8738)_T + (5/2663, 1/7929, 0/0012)_T z_1 + (3/5575, 0/3739, 0/5952)_T z_2$$

جدول‌های ۲۱.۲ و ۲۲.۲ خطای محاسبه شده برای روش پیشنهادی را با روش موجود در [۶۵] مقایسه می‌کنند. در هر دو جدول خطای محاسبه شده برای روش ارائه شده در این فصل خطای کمتری نسبت به روش دیگر دارد.

۸۰ مدل شبکه عصبی بازگشتی برای حل مسائل رگرسیون فازی

جدول ۲۱.۲: MSPE در مثال ۴.۲.۲

خطای به دست آمده			روش / فاصله
D_2	D_1	D_3	
۱/۰۲۷۷	۳/۰۴۴۴	۰/۲۱۶۱	مدل ارائه شده در [۶۵]
۱/۹۱۵۰	۱۲/۷۹۰۲	۰/۹۴۸۴	
۱/۴۵۸۷	۱/۵۹۸۹	۰/۰۸۷۳	
۲/۴۸۵۴	۱۲/۳۵۵۸	۰/۷۱۶۳	
۱/۱۷۱۵	۲/۹۵۹۴	۰/۱۸۴۳	
۱/۶۱۱۷	۶/۵۴۹۷	۰/۴۳۰۵	خطای کل
۱/۰۳۴۶	۳/۰۹۱۰	۰/۲۲۱۰	مدل (۴۲.۲) و (۴۳.۲)
۱/۸۸۴۴	۱۲/۶۹۳۶	۰/۹۴۷۸	
۱/۵۳۷۶	۱/۷۷۰۹	۰/۰۸۵۹	
۲/۴۱۰۴	۱۲/۰۱۱۸	۰/۷۰۷۳	
۱/۱۰۵۸	۲/۸۴۱۷	۰/۱۸۲۹	
۱/۵۹۴۶	۶/۴۸۱۸	۰/۴۲۹۰	خطای کل

جدول ۲۲.۲: روش MCV در مثال ۴.۲.۲

MCV			روش / فاصله
D_2	D_1	D_3	
۰/۱۴۲۷	۰/۲۲۹۷	۰/۰۱۹۸	مدل ارائه شده در [۶۵]
۱/۸۵۳۶	۴/۵۷۸۷	۰/۲۲۰۹	
۶/۸۱۴۴	۷۲/۷۶۸۲	۳/۹۲۸۰	
۷/۷۶۳۷	۹۴/۸۶۳۸	۵/۱۹۰۳	
۶/۶۶۲۳	۶۵/۰۱۸۰	۴/۱۴۰۵	
۴/۶۴۷۳	۴۷/۴۹۱۷	۲/۶۹۹۹	خطای کل
۰/۴۰۳۱	۰/۲۴۵۷	۰/۰۲۱۰	مدل (۴۲.۲) و (۴۳.۲)
۱/۸۵۵۲	۴/۵۹۵۸	۰/۲۲۲۰	
۳/۸۸۷۷	۲۳/۴۸۰۸	۱/۲۷۱۸	
۷/۰۲۲۹	۶۷/۲۰۵۷	۴/۴۰۱۹	
۳/۴۶۱۷	۵۶/۸۳۰۵	۵/۱۷۹۱	
۳/۳۲۶۱	۳۰/۴۷۱۷	۲/۲۱۹۱	خطای کل

مثال ۵.۲.۲. داده‌های جدول ۲۳.۲ را در نظر بگیرید.

جدول ۲۳.۲: مقادیر مشاهده شده در مثال ۵.۲.۲

۶	۴	۳	۲	۱	x_1
۴	۵	۳	۳	۱	x_2
$(۸۴, ۹, ۶)_T$	$(۶۷, ۵, ۳)_T$	$(۳۳, ۳, ۲)_T$	$(۲۳, ۳, ۱)_T$	$(۵, ۱, ۱)_T$	Y

هدف تقریب ضرایب مدل رگرسیون چندجمله‌ای زیر با استفاده از این داده‌ها است.

$$Y = A_0 x_1 + A_1 x_1 x_2$$

با استفاده از روش پیشنهادی و با فرض $z_1 = x_1$ و $z_2 = x_1 x_2$ مدل رگرسیون زیر بدست می‌آید:

$$Y = (۲/۶۰۱۳, ۱/۱۵۳۱, ۰/۷۵۸۴)_T z_1 + (۲/۸۴۳۲, ۰/۰۵۶۰, ۰/۰۲۳۰)_T z_2$$

همگرایی مسیرهای شبکه عصبی در شکل ۳۰.۲ نشان داده شده است. در جدول ۲۴.۲ نتایج عددی حاصل از مقایسه روش پیشنهادی و روش موجود در [۶۵] نشان داده شده است.

جدول ۲۴.۲: خطای به دست آمده در مثال ۵.۲.۲

خطای به دست آمده			
روش / فاصله	D_3	D_1	D_2
روش ارائه شده در [۶۵]	۰/۱۴۱۸	۱/۳۰۶۵	۰/۸۶۴۵
	۰/۳۴۳۲	۲/۹۴۱۰	۱/۷۹۷۴
	۰/۱۸۷۹	۲/۹۰۳۸	۱/۸۲۷۱
	۰/۱۱۶۰	۱/۹۸۰۹	۱/۴۷۰۸
	۰/۱۴۴۹	۲/۹۴۶۰	۱/۸۱۹۸
خطای کل	۰/۱۸۶۵	۲/۴۱۵۶	۱/۵۵۶۰
مدل (۴۲.۲) و (۴۳.۲)	۰/۱۴۱۸	۱/۳۰۳۹	۰/۸۷۲۲
	۰/۳۴۲۲	۲/۹۵۴۹	۱/۷۵۰۸
	۰/۱۸۷۱	۲/۸۹۱۱	۱/۸۳۸۲
	۰/۱۱۵۳	۱/۹۴۵۲	۱/۴۹۵۲
	۰/۱۴۴۸	۲/۰۰۰۲	۱/۷۹۰۴
خطای کل	۰/۱۸۶۴	۲/۲۱۹۰	۱/۵۴۹۴

جدول برتری روش پیشنهادی را نسبت به دیگر مدل ارائه شده در [۶۵] نشان می‌دهد.

مثال ۶.۲.۲. (تأمین نیروی کار) تحت مفروضات مثال ۶.۱.۲ و با استفاده از روش (۴۲.۲) و (۴۳.۲) ضرایب مدل رگرسیون (۳۱.۲) را با فرض $z_1 = x_1$ و $z_2 = x_2^2$ به صورت زیر تقریب می‌زنیم:

$$\bar{Y} = (38/1789, 0/8158, 0/7547)_T + (1/8153 \times 10^{-5}, 2/4033 \times 10^{-13}, 1/7276 \times 10^{-13})_T z_1 + (-1/0420 \times 10^{-8}, 1/0383 \times 10^{-15}, 4/8664 \times 10^{-17})_T z_2$$

در [۶۵]، با ۱۵ تکرار از الگوریتم یادگیری مدل زیر را داریم:

$$\bar{Y} = (38/261, 2/2345, 1/0203)_T + (1/8312, 0/2233, 0/3124)_T \times 10^{-5} x_1 + (-1/0643, 0/0021, 0/061)_T \times 10^{-8} x_2^2$$

جدول ۲۵.۲ تفاوت میان خطای به دست آمده با این دو روش را نشان می‌دهد. واضح است که روش پیشنهادی عملکرد بهتری بر روی مقادیر مشاهده شده داشته است.

۴.۲.۲ مزایای روش پیشنهادی

علاوه بر مزایای شماره ۱ تا ۹ بیان شده در ۳۵.۲، برخی دیگر از مزایای روش پیشنهادی را به شرح زیر خلاصه می‌کنیم:

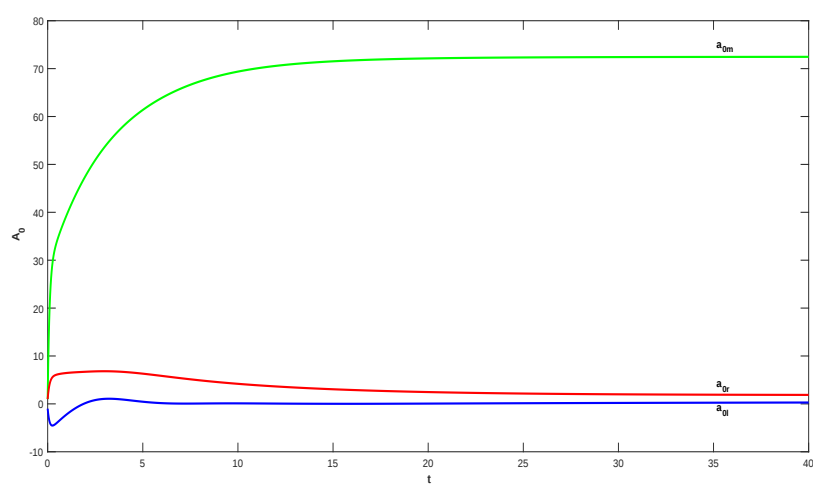
- برای نشان دادن برخی از مزایای مدل پیشنهادی نسبت به دیگر مدل‌ها، از یک روش آماری معروف به نام «اعتبار سنجی متقابل» استفاده می‌کنیم.

۸۲ مدل شبکه عصبی بازگشتی برای حل مسائل رگرسیون فازی

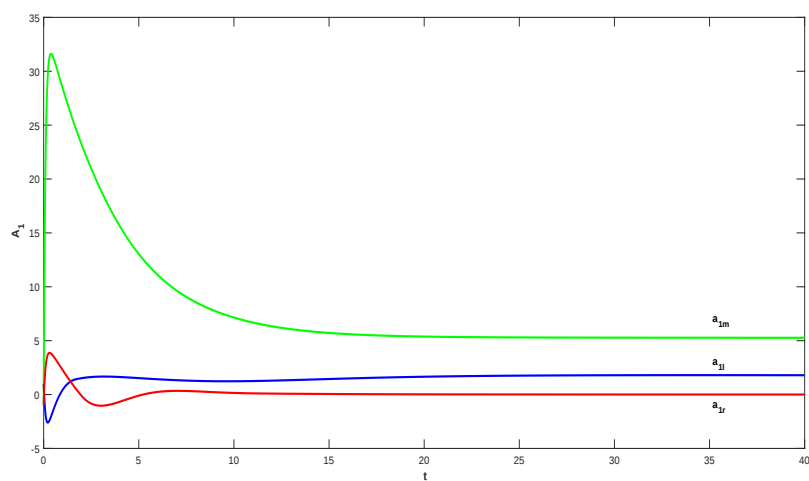
جدول ۲۵.۲: خطای به دست آمده در مثال ۶.۲.۲

خطای به دست آمده	
مدل ارائه شده در [۶۵]	مدل (۴۲.۲) و (۴۳.۲)
۳/۹۴۲۵	۱/۴۵۸۹
۳/۱۳۸۳	۰/۴۹۲۶
۲/۷۵۱۹	۱/۵۰۹۹
۳/۴۶۸۶	۱/۶۰۸۶
۲/۸۹۳۴	۱/۱۸۳۴
۴/۴۱۴۴	۱/۱۰۸۴
۴/۱۲۶۳	۰/۷۱۱۰
۳/۸۷۷۴	۰/۸۷۸۵
۵/۴۱۴۲	۲/۱۰۳۲
۵/۴۲۹۷	۱/۹۸۸۳
۴/۵۱۰۷	۱/۷۰۷۰
۵/۶۳۰۷	۲/۰۶۴۹
۴/۵۱۳۹	۰/۹۰۲۷
۴/۵۹۲۲	۰/۷۵۱۹
۴/۸۲۸۶	۰/۸۱۴۶
۵/۰۰۱۸	۰/۸۴۷۳
۵/۶۴۳۴	۱/۳۶۴۵
۵/۱۶۴۷	۰/۷۸۹۴
۵/۱۳۶۳	۰/۶۱۷۱
۴/۴۴۶۳	۱/۲۰۵۴

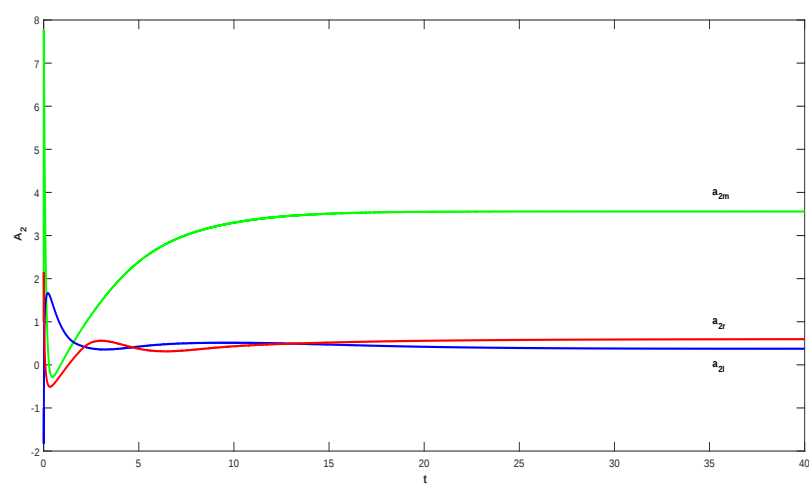
- در این بخش حالت کامل مدل رگرسیون چندجمله‌ای فازی با تعامل میان متغیرها در نظر گرفته شده. در حالی که، در برخی مقالات نویسندگان بیشتر سعی می‌کنند مسئله را برای مرتبه اول و دوم حل کنند. (برای مثال به [۲۹] مراجعه شود).



(آ) مسیر A_0

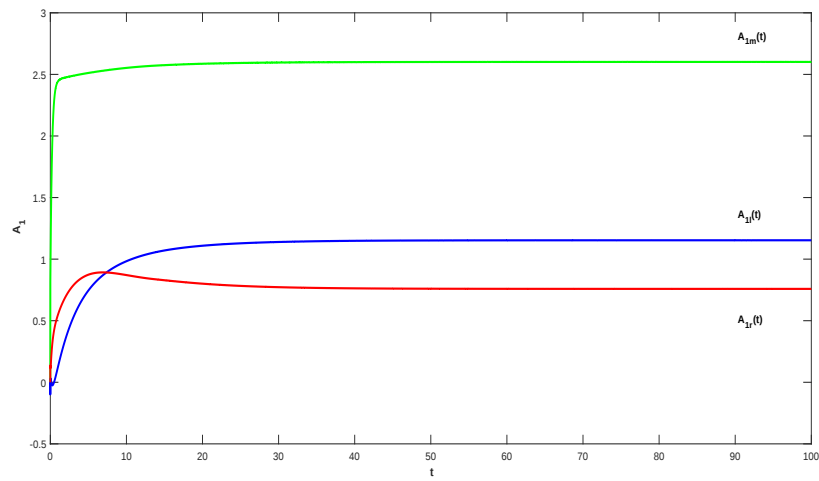


(ب) مسیر A_1

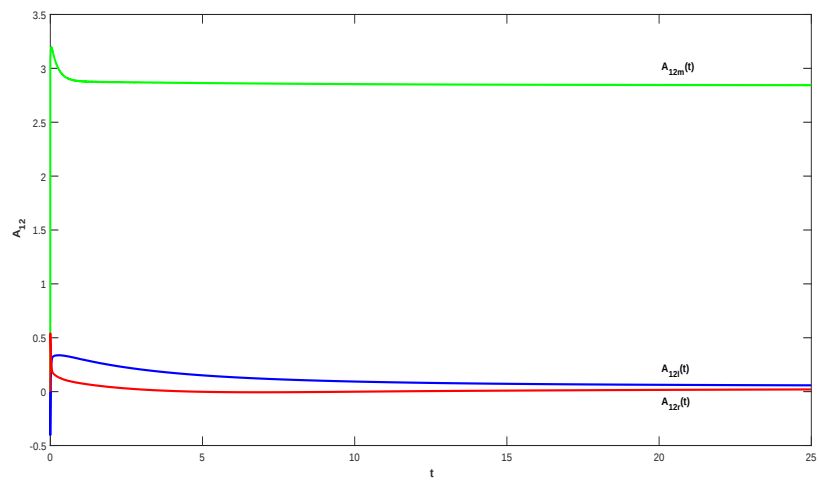


(ج) مسیر A_2

شکل ۲۹.۲: رفتار مسیر A از شبکه عصبی (۴۲.۲) و (۴۳.۲) در مثال ۴.۲.۲.



(آ) مسیر A_0



(ب) مسیر A_1

شکل ۳۰.۲: رفتار مسیر A از شبکه عصبی (۴۲.۲) و (۴۳.۲) در مثال ۵.۲.۲.

فصل ۳

مدل شبکه عصبی بازگشتی برای حل مسائل رگرسیون ریج فازی

همانند آنچه در ۳.۱.۱ برای مدل‌های رگرسیون غیرفازی بیان شد، در مدل‌های رگرسیون فازی نیز محققان اغلب با مشکل هم‌خطی که به دلیل وجود وابستگی خطی میان ستون‌های ماتریس طرح می‌باشد، مواجه می‌شوند (برای مثال به [۱۲، ۸۲، ۸۳] مراجعه شود). اتخاذ روش رگرسیون ریج فازی برای غلبه بر مشکل هم‌خطی به طور گسترده‌ای در [۸۲، ۸۴] مورد مطالعه قرار گرفته است. مشاهده شده است که این برآوردگر عملکرد بهتری نسبت به سایر برآوردگرهای متکی به روش کمترین توان‌های کلاسیک دارد.

با توجه به مباحث ارائه شده در فصل ۲، شبکه‌های عصبی مختلفی برای حل مسائل رگرسیون خطی و چندجمله‌ای فازی ارائه شده‌اند. با این حال تنها یک گزارش از حل مدل رگرسیون فازی ریج با شبکه عصبی وجود دارد (برای اطلاعات بیشتر به [۸۵] مراجعه شود). در مرجع اشاره شده نویسندگان یک الگوریتم یادگیری برای شبکه عصبی فازی پیشنهاد دادند و نشان دادند که الگوریتم قادر به کاهش اثر داده‌های پرت است. مهمترین ایراد این روش اشاره نکردن به قضایا و گزاره‌های مربوط به همگرایی و پایداری شبکه عصبی پیش‌رونده بیان شده است. از طرف دیگر، در این روش برای به‌دست آوردن وزن‌ها از یک الگوریتم تکراری استفاده شده است که نیازمند محاسبه ماتریس وارون در هر تکرار است. واضح است که هر چه بعد مسئله بالاتر برود پیچیدگی محاسباتی بیشتر می‌شود. علاوه بر این، در روش پیشنهادی

همگرایی شبکه عصبی به انتخاب یک نقطه اولیه مناسب بستگی دارد. بنابراین هدف این بخش، پیشنهاد یک شبکه عصبی مناسب برای حل مسائل رگرسیون فازی ریج بدون نیاز به محاسبه پارامتر جریمه و با تجزیه و تحلیل دقیق پایداری و همگرایی است.

۱.۳ مدل رگرسیون ریج فازی

هدف این بخش مقابله با اثر هم خطی در مسائل رگرسیون فازی است. برای دستیابی به این هدف نویسندگان در [۸۲] پیشنهاد داده‌اند که مجموع مربعات باقیمانده $\sum_{i=1}^n (Y_i - \sum_{j=0}^p A_j x_{ij})^2$ را تحت قید $\sum_{j=0}^p (A_j)^2 \leq t$ ، برای $t > 0$ به حداقل برسانیم. به عبارت دیگر با توجه به (۶.۱)، مسئله رگرسیون ریج فازی به صورت زیر مدل سازی می شود:

$$\text{minimize } \sum_{i=1}^n D^2(\hat{Y}_i, Y_i) + \lambda \sum_{j=0}^p D^2(A_j, 0) \quad (1.3)$$

که در آن پارامتر مثبت لاگرانژی λ به عنوان پارامتر ریج شناخته می شود (زمانی که $\lambda = 0$ مدل حداقل مربعات معمولی به دست می آید).

با توجه به فاصله تعریف شده در (۳.۱.آ) مدل (۱.۳) به صورت زیر بازنویسی می شود:

$$\begin{aligned} \text{minimize } \sum_{i=1}^n & \left[(y_{im} - \hat{y}_{im})^2 + ((y_{im} - \gamma y_{il}) - (\hat{y}_{im} - \gamma \hat{y}_{il}))^2 \right. \\ & \left. + ((y_{im} - \rho y_{iu}) - (\hat{y}_{im} - \rho \hat{y}_{iu}))^2 \right] \\ & + \lambda \sum_{j=0}^p \left[a_{jm}^2 + (a_{jm} - \gamma a_{jl})^2 + (a_{jm} - \rho a_{ju})^2 \right] \end{aligned} \quad (2.3)$$

پارامترهای $\gamma = \int_0^1 L^{-1}(\omega) d\omega$ و $\rho = \int_0^1 R^{-1}(\omega) d\omega$ تنوع تابع عضویت در مدل را در نظر می گیرند و هم زمان تأثیر پهنای در مدل را کم می کنند (برای اطلاعات بیشتر به پیوست آ مراجعه شود).

مسئله (۲.۳) یک مدل بهینه سازی است که روش ها متفاوتی برای حل آن پیشنهاد شده است. هر یک از این روش ها نقاط ضعف و قوت خاص خود را دارند. در این بخش، ما تلاش می کنیم که با استفاده از شبکه های عصبی به حل مسئله بهینه سازی (۲.۳) بپردازیم.

۲.۳ مدل بهینه سازی

بر اساس مفاهیم بخش قبل، تابع هزینه $C(\mathbf{A})$ به صورت زیر تعریف می شود:

$$\begin{aligned} C(\mathbf{A}) = & \sum_{i=1}^n \left[\left(y_{im} - \left(\sum_{j=\circ}^p a_{jm} x_{ij} \right) \right)^2 \right. \\ & + \left((y_{im} - \gamma y_{il}) - \left(\sum_{j=\circ}^p a_{jm} x_{ij} - \gamma \sum_{j=\circ}^p a_{jl} x_{ij} \right) \right)^2 \\ & + \left. \left((y_{im} - \rho y_{iu}) - \left(\sum_{j=\circ}^p a_{jm} x_{ij} - \rho \sum_{j=\circ}^p a_{ju} x_{ij} \right) \right)^2 \right] \\ & + \lambda \sum_{j=\circ}^p \left[a_{jm}^2 + (a_{jm} - \gamma a_{jl})^2 + (a_{jm} - \rho a_{ju})^2 \right] \end{aligned} \quad (۳.۳)$$

که در آن $\mathbf{A} = (a_{\circ m}, \dots, a_{pm}, a_{\circ l}, \dots, a_{pl}, a_{\circ u}, \dots, a_{pu})^\top$ ، یک بردار ستونی $(3p+3)$ -بعدی است. بنابراین (۳.۳) معادل است با

$$\begin{aligned} C(\mathbf{A}) = & \sum_{i=1}^n \left[\left((y_{im})^2 - 2 y_{im} \left(\sum_{j=\circ}^p a_{jm} x_{ij} \right) + \left(\sum_{j=\circ}^p a_{jm} x_{ij} \right)^2 \right) \right. \\ & + \left((y_{im} - \gamma y_{il})^2 - 2 (y_{im} - \gamma y_{il}) \left(\sum_{j=\circ}^p a_{jm} x_{ij} \right) \right. \\ & + 2 \gamma (y_{im} - \gamma y_{il}) \left(\sum_{j=\circ}^p a_{jl} x_{ij} \right) + \left. \left(\sum_{j=\circ}^p a_{jm} x_{ij} \right)^2 \right. \\ & - \left. 2 \gamma \left(\sum_{j=\circ}^p a_{jm} x_{ij} \right) \left(\sum_{j=\circ}^p a_{jl} x_{ij} \right) + \left(\sum_{j=\circ}^p a_{jl} x_{ij} \right)^2 \right) \\ & + \left((y_{im} - \rho y_{iu})^2 - 2 (y_{im} - \rho y_{iu}) \left(\sum_{j=\circ}^p a_{jm} x_{ij} \right) \right. \\ & + 2 \rho (y_{im} - \rho y_{iu}) \left(\sum_{j=\circ}^p a_{ju} x_{ij} \right) + \left. \left(\sum_{j=\circ}^p a_{jm} x_{ij} \right)^2 \right. \\ & - \left. 2 \rho \left(\sum_{j=\circ}^p a_{jm} x_{ij} \right) \left(\sum_{j=\circ}^p a_{ju} x_{ij} \right) + \left(\sum_{j=\circ}^p a_{ju} x_{ij} \right)^2 \right) \\ & + \lambda \sum_{j=\circ}^p \left[a_{jm}^2 + (a_{jm} - \gamma a_{jl})^2 + (a_{jm} - \rho a_{ju})^2 \right] \\ & = \mathbf{A}^\top \mathbf{Q} \mathbf{A} + \mathbf{B}^\top \mathbf{A} + M \end{aligned} \quad (۴.۳)$$

که در آن

$$\mathbf{Q} = \begin{bmatrix} 2(\mathbf{q} + \lambda I) & -2\gamma(\mathbf{q} + \lambda I) & -2\rho(\mathbf{q} + \lambda I) \\ \mathbf{0} & \gamma^2(\mathbf{q} + \lambda I) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \rho^2(\mathbf{q} + \lambda I) \end{bmatrix} \quad (۵.۳)$$

و O یک ماتریس $(n+1) \times (n+1)$ صفر است. همچنین

$$\mathbf{B} = (b_{1\circ}, b_{11}, \dots, b_{1p}, b_{2\circ}, \dots, b_{2p}, b_{3\circ}, \dots, b_{3p})^\top \quad (۶.۳)$$

$$M = \sum_{i=1}^n \left(y_{im}^2 + (y_{im} - \gamma y_{il})^2 + (y_{im} - \rho y_{iu})^2 \right) \quad (۷.۳)$$

$$\mathbf{q} = \begin{bmatrix} \sum_{i=1}^n x_{i\circ} & \sum_{i=1}^n x_{i1} & \dots & \sum_{i=1}^n x_{ip} \\ \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1}^2 & \dots & \sum_{i=1}^n x_{i1}x_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n x_{ip} & \sum_{i=1}^n x_{ip}x_{i1} & \dots & \sum_{i=1}^n x_{ip}^2 \end{bmatrix}$$

که در آن

$$b_{1j} = -2 \sum_{i=1}^n x_{ij} (\gamma y_{im} - \gamma y_{il} - \rho y_{iu})$$

$$b_{2j} = 2\gamma \sum_{i=1}^n x_{ij} (y_{im} - \gamma y_{il})$$

$$b_{3j} = 2\rho \sum_{i=1}^n x_{ij} (y_{im} - \rho y_{iu})$$

با توجه به تعریف اعداد فازی LR در پیوست A ، گستره سمت چپ و راست اعداد فازی A_j برای $j = \circ, 1, \dots, p$ باید مثبت باشند. بنابراین (۲.۳) به شکل درجه دوم مقید زیر تبدیل می‌شود:

$$\text{minimize } C(\mathbf{A}) = \mathbf{A}^\top \mathbf{Q} \mathbf{A} + \mathbf{B}^\top \mathbf{A} + M \quad (۸.۳)$$

subject to

$$\mathbf{G} \mathbf{A} \leq \circ \quad (۹.۳)$$

که در آن

$$\mathbf{G} = \begin{bmatrix} \circ & \dots & \circ & \circ & \dots & \circ & \circ & \dots & \circ \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \circ & \dots & \circ & \circ & \dots & \circ & \circ & \dots & \circ \\ \circ & \dots & \circ & -1 & \dots & \circ & \circ & \dots & \circ \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \circ & \dots & \circ & \circ & \dots & -1 & \circ & \dots & \circ \\ \circ & \dots & \circ & \circ & \dots & \circ & -1 & \dots & \circ \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \circ & \dots & \circ & \circ & \dots & \circ & \circ & \dots & -1 \end{bmatrix}_{(\gamma p + \gamma) \times (\gamma p + \gamma)}$$

و Q, B و M به ترتیب در (۵.۳)، (۶.۳) و (۷.۳) تعریف شده اند. محاسبه می شود که هسین ماتریس $C(A)$ یک ماتریس معین مثبت ($\nabla^2 C(A) = Q$) است. بنابراین مسئله (۸.۳) و (۹.۳) یک مسئله محدب با قیود نامساوی است.

۳.۳ رویکرد شبکه عصبی بازگشتی

روش شبکه عصبی مناسب برای حل این مسئله مشابه با مباحث ارائه شده در ۴.۳.۱ می باشد و تمامی اثبات ها و نتایج با جایگذاری $p = p'$ به دست می آیند. بنابراین در این بخش تنها به مرور مباحث مربوطه می پردازیم. با فرض این که $A(\cdot)$ و $U(\cdot)$ دو متغیر وابسته به زمان باشند، شبکه عصبی زیر را معرفی می کنیم:

$$\begin{cases} \frac{dy}{dt} = \tau \phi(y) \\ y(t_0) = y_0 = (A(t_0)^T, U(t_0)^T)^T \in \mathbb{R}^{p+q}, \tau > 0 \end{cases} \quad (10.3)$$

$$(11.3)$$

که در آن

$$\phi(y) = \begin{pmatrix} -(\tau QA + B + G^T(GA)^+) \\ (GA)^+ - U \end{pmatrix} \quad (12.3)$$

قضیه ۱۰.۳.۳. فرض کنید $y^* = (A^{*T}, U^{*T})^T$ نقطه تعادل مدل شبکه عصبی (۱۰.۳) و (۱۱.۳) باشد. بنابراین A^* یک نقطه کروش-کان-تاکر برای مسئله (۸.۳) و (۹.۳) است. از سمت دیگر، اگر $A^* \in \mathbb{R}^{p+q}$ یک جواب مسئله (۸.۳) و (۹.۳) باشد، بنابراین $U^* \in \mathbb{R}^{p+q}$ وجود دارد به طوری که $y^* = (A^{*T}, U^{*T})^T$ نقطه تعادل مدل شبکه عصبی (۱۰.۳) و (۱۱.۳) است.

لم ۱۰.۳.۳. ژاکوبین ماتریس ϕ که در (۱۲.۳) تعریف شده است، نیمه معین منفی است.

قضیه ۲.۳.۳. تحت مفروضات لم ۱.۳.۳ شبکه عصبی (۱۰.۳) و (۱۱.۳) پایدار لیاپانوف است.

لم ۲.۳.۳. (i) برای هر نقطه ابتدایی $y(t_0) = (A(t_0)^T, U(t_0)^T)^T$ یک جواب پیوسته یکتا $y(t) = (A(t)^T, U(t)^T)^T$ برای سیستم (۱۰.۳) و (۱۱.۳) وجود دارد.

(ii) فرض کنید $y(t) = (A(t)^T, U(t)^T)^T$ یک مسیر از شبکه عصبی (۱۰.۳) و (۱۱.۳) با نقطه ابتدایی $y(t_0) = (A(t_0)^T, U(t_0)^T)^T$ باشد. در این صورت اگر، $U(t_0) \geq 0$ ، بنابراین $U(t) \geq 0$.

قضیه ۳.۳.۳. مسیرهای شبکه عصبی (۱۰.۳) و (۱۱.۳) برای هر نقطه ابتدایی

$$y_0 = (A(t_0)^T, U(t_0)^T)^T \in \mathbb{R}^{p+q}$$

به نقطه تعادل شبکه عصبی همگرا می‌شود. به طور خاص، شبکه عصبی (۱۰.۳) و (۱۱.۳) با هر نقطه ابتدایی $y_0 = (A(t_0)^T, U(t_0)^T)^T \in \mathbb{R}^{p+6}$ پایدار مجانبی سراسری است و D^* یک نقطه تعادل یکتا دارد. که در آن D^* نشان دهنده مجموعه نقاط بهینه مسئله اصلی و دوگان است.

قضیه ۴.۳.۳. سرعت همگرایی شبکه عصبی (۱۰.۳) و (۱۱.۳) با افزایش τ افزایش می‌یابد.

۴.۳ مثال‌های عددی

به منظور نشان دادن کارایی مدل پیشنهادی، در این بخش به حل دو مثال کاربردی می‌پردازیم و همگرایی مدل را با نقاط ابتدایی مختلف آزمایش می‌کنیم. برای نشان دادن بیشتر اثربخشی این شبکه عصبی، مقایسه‌هایی با مدل‌های موجود در [۸۶] و [۸۷] با فاصله‌های مختلف D_3 و D_4 انجام شده است. علاوه بر این، نمودار تیلور را ارائه می‌دهیم که اطلاعات مربوط به چندین شاخص آماری را برای ارزیابی صحت شبکه عصبی فراهم می‌کند.

مثال ۱.۴.۳. مثال مونت کارلو^۱ ارائه شده در [۸۸] را در نظر بگیرید. برای $n = 3$ ، $p = 50$ و $\alpha = 0.99$ داده‌های جدول ۱.۳ را داریم (داده‌ها استاندارد شده‌اند). هدف، به‌دست آوردن ضرایب فازی مدل رگرسیون خطی

$$Y = A_1 x_1 + A_2 x_2 + A_3 x_3 \quad (13.3)$$

با استفاده از داده‌های ارائه شده می‌باشد.

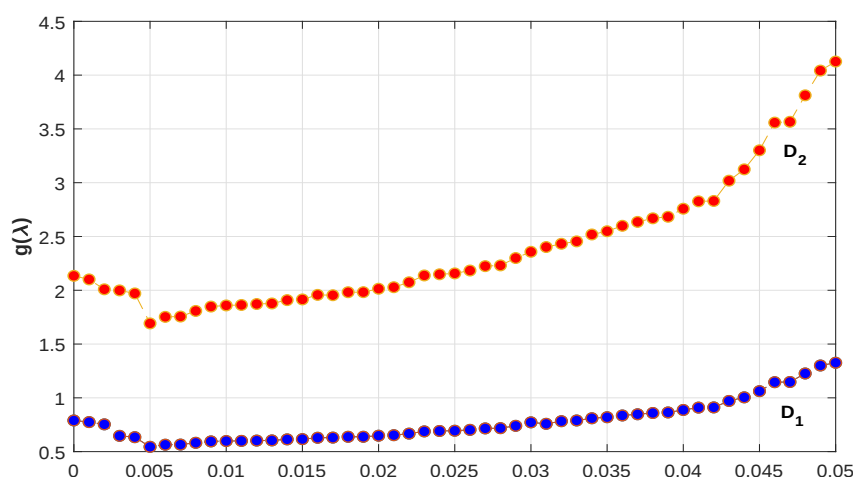
برای اطلاع از وجود هم‌خطی در داده‌ها از معیار تورم واریانس یا VIF_j برای j امین ضریب رگرسیون (A_j) استفاده می‌کنیم (برای اطلاعات بیشتر به پیوست ۳.۱.۱ مراجعه شود). با توجه به این توضیحات در این مثال داریم:

$$VIF_1 = 33/5131, \quad VIF_2 = 41/0900, \quad VIF_3 = 44/8060.$$

VIF ها به‌دست آمده در بازه قابل قبول نیستند. بنابراین به دلیل وجود هم‌خطی در مدل، استفاده از یک مدل رگرسیون ریج مناسب، الزامی می‌باشد. از طرف دیگر، برای تقریب مقدار Y با استفاده از مدل پیشنهادی، به مقدار بهینه پارامتر ریج λ نیاز داریم. در شکل ۱.۳ با استفاده از دو فاصله مختلف مقدار خطا را برای مقادیر مختلف λ به‌دست آورده ایم (برای این منظور در این فصل از روش ارایه در [۸۹] استفاده شده است).

مشاهده می‌شود که $\lambda = 0.005$ کمترین مقدار خطا برای هر دو فاصله استفاده شده را دارد (شکل ۱.۳ مشاهده شود). بنابراین مدل رگرسیون ریج فازی به صورت زیر به‌دست می‌آید:

^۱Monte Carlo



شکل ۱.۳: مقادیر مختلف λ با دو فاصله متفاوت D_3 و D_4 با استفاده از روش ارایه شده در [۸۹] در مثال ۱.۴.۳.

$$\hat{Y}_{Ridge} = (13/3712, 1/6140, 2/2576)Tx_1 + (-5/1237, 0, 0)Tx_2 + (2/1421, 0, 0)Tx_3. \quad (14.3)$$

در این حالت مقدار $MSPE$ با دو فاصله متفاوت D_3 و D_4 به ترتیب برابر است با ۵۴۴۵٪ و ۱/۶۹۱۸ و همچنین $\bar{R}^2 = ۰/۸۰۶۶$. که نشان می‌دهند مدل پیشنهادی مشاهدات را به درستی در بر گرفته است. شکل ۲.۳ نشان می‌دهد که مسیرهای شبکه عصبی به جواب بهینه مسئله همگرا می‌شوند. همچنین در شکل ۲.۳ می‌بینیم که برای هر نقطه شروع اولیه مسئله همواره به جواب بهینه مدل همگرا است.

عملکرد گرافیکی مدل توسط نمودار تیلور فازی در شکل‌های ۴.۳-۵.۳ به تصویر کشیده شده است. نتایج نشان می‌دهد که مدل پیشنهادی (نشان داده شده با مربع) در هر سه شکل، نزدیک به داده‌های مشاهده شده واقع شده است که به معنی عملکرد مناسب مدل بر روی داده‌ها می‌باشد.

برای بررسی بیشتر عملکرد مدل به دست آمده، توانایی پیش‌بینی مدل در جدول ۲.۳ محاسبه می‌شود. مقادیر MCV محاسبه شده از تعریف (۹.۱) تقریباً برابر با مقادیر $MSPE$ محاسبه شده است. بنابراین، نتیجه گرفته می‌شود که مدل پیشنهادی مورد استفاده در این بخش از قدرت برآزش خوبی برخوردار است و از این رو در عمل از شایستگی خاصی برخوردار است.

مثال ۲.۴.۳. در این مثال داده‌های نشان داده شده در جدول (۳.۳) برای متغیرهای ورودی غیرفازی و متغیر خروجی نرمال از [۸۷] گرفته شده است (داده‌ها استاندارد شده اند). این داده‌ها نمونه‌ای از وضعیتی هستند که در آن یک متغیر ترکیب خطی کامل از متغیر دیگر است. همانند آنچه در مثال ۱.۴.۳ دیدیم VIF_j ها برای نشان دادن هم‌خطی در مدل استفاده

می شوند. در این مثال داریم:

$$VIF_1 = VIF_2 = 2683/3,$$

که نشان دهنده وجود هم خطی در مدل است. همچنین به دلیل نرمال بودن خروجی، در مدل پیشنهادی قرار می دهیم: $\gamma = \rho = \frac{\sqrt{11}}{4}$. مقادیر تقریب زده شده برای پارامتر ريج λ در شکل ۶.۳ نشان داده شده است (با استفاده از مدل ارایه شده در [۸۹]). بنابراین با فرض $\lambda = 0/01$ مدل رگرسیون ريج فازی زیر را داریم:

$$\hat{Y}_{Ridge} = (20/5990, 0/1556 \times 10^{-7}, 0/1783 \times 10^{-6})_{Nx1} \quad (15.3) \\ + (21/5685, 1/4852, 1/4852)_{Nx2}.$$

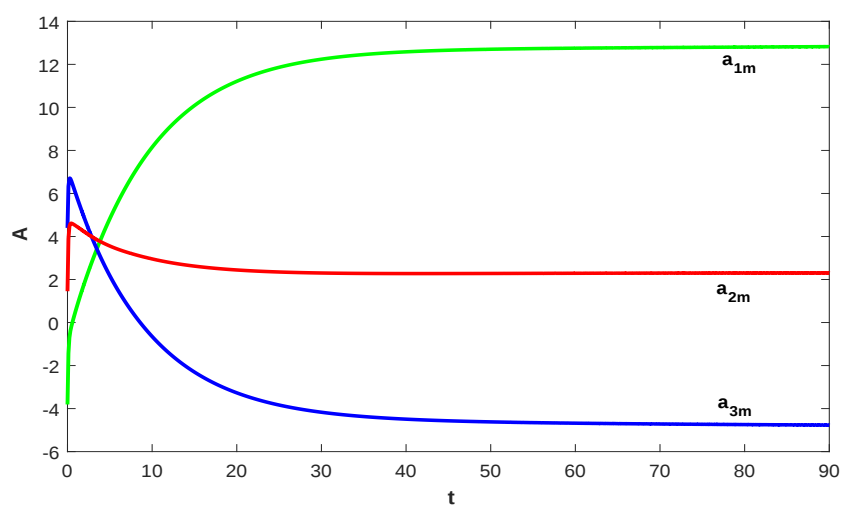
مقادیر $MSPE$ با دو فاصله متفاوت D_3 و D_4 به ترتیب برابر است با $0/157$ و $0/1613$ و همچنین $\bar{R}^2 = 0/7701$. شکل ۷.۳ همگرایی مسیرهای شبکه عصبی پیشنهادی به جواب بهینه مسئله را نشان می دهد. همچنان همگرایی مدل برای چندین نقطه ابتدایی متفاوت، در شکل ۸.۳ نشان داده شده است.

برای مقایسه، این مدل با روش های ارائه شده در [۸۶] و [۸۷] نیز حل شده است. جدول ۵.۳ و ۴.۳ خطای محاسبه شده و قدرت پیش بینی این سه روش را نشان می دهند. به وضوح خطای محاسبه شده برای روش پیشنهادی از دیگر روش های موجود کمتر می باشد.

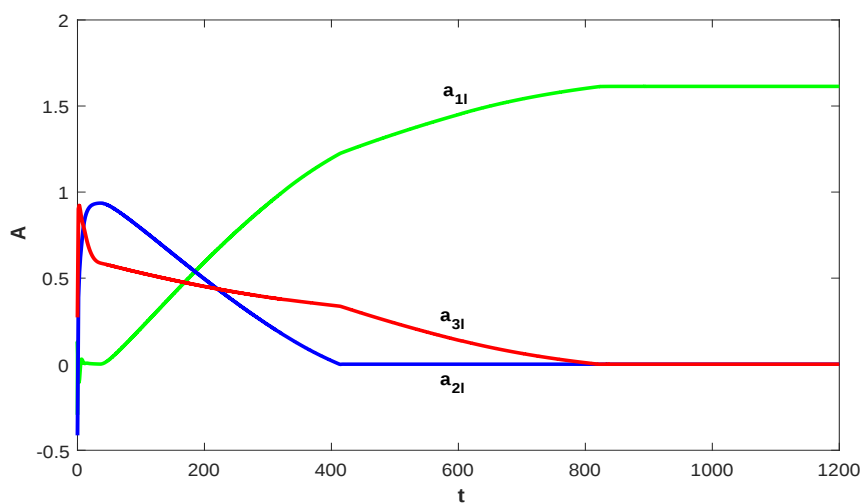
در نمودار تیلور (شکل (۱۰.۳))، مقایسه ای میان روش پیشنهادی و مدل های ارائه شده در [۸۶] و [۸۷] برای حل مدل رگرسیون ريج فازی صورت گرفته است. موقعیت هر مستطیل روی نمودار نشان دهنده نتایج به دست آمده برای یک مدل متفاوت است که توسط ضریب همبستگی و انحراف معیار تعیین می شود. روشن است که شبکه عصبی پیشنهادی در مقایسه با دو مدل دیگر دارای قابلیت بهتر برای حل مسائل این چنینی است.

جدول ۱.۳: مقادیر مشاهده شده در مثال ۱.۴.۳

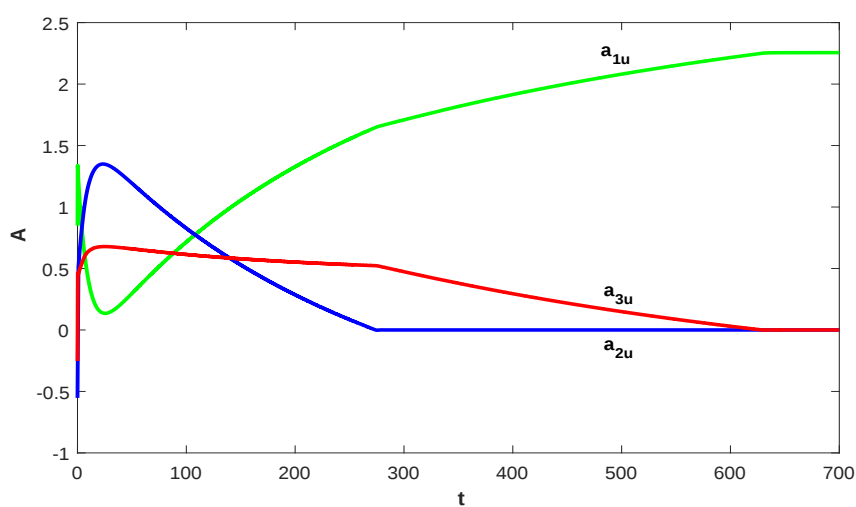
شماره	x_1	x_2	x_3	Y
۱	۰/۴۲۴۶	۰/۴۶۸۷	۰/۵۱۵۹	$(-۰/۸۱۲۶, ۳/۳۴۰۵, ۳/۲۹۱۳)_T$
۲	۰/۱۲۶۲	۰/۲۰۸۴	۰/۱۸۸۷	$(-۵/۱۷۳۶, ۲/۶۸۵۵, ۲/۳۷۵۳)_T$
۳	۰/۳۵۷۵	۰/۴۴۷۲	۰/۴۲۷۷	$(-۰/۹۳۵۶, ۳/۳۲۱۵, ۳/۲۶۵۳)_T$
۴	۰/۴۳۴۶	۰/۴۸۵۳	۰/۵۱۵۶	$(۰/۹۹۴۴, ۳/۶۱۱۵, ۳/۶۷۰۳)_T$
۵	۰/۶۷۱۱	۰/۶۹۶۲	۰/۷۷۰۸	$(۲/۷۲۹۴, ۳/۸۷۱۵, ۴/۰۳۵۳)_T$
۶	۰/۴۱۵۳	۰/۴۷۸۳	۰/۵۰۴۴	$(-۰/۴۴۹۶, ۳/۳۹۴۵, ۳/۳۶۷۳)_T$
۷	۰/۳۵۲۴	۰/۳۹۸۵	۰/۴۴۹۷	$(۰/۱۶۹۴, ۳/۴۸۷۵, ۳/۴۹۷۳)_T$
۸	۰/۴۱۶۶	۰/۵۰۳۲	۰/۵۳۰۴	$(-۰/۶۱۷۶, ۳/۳۶۹۵, ۳/۳۳۲۳)_T$
۹	۰/۵۶۴۵	۰/۶۴۳۷	۰/۶۸۳۹	$(۰/۴۳۴۴, ۳/۵۲۷۵, ۳/۵۵۳۳)_T$
۱۰	۰/۸۸۲۸	۰/۹۳۳۵	۱/۰۲۷۸	$(۴/۳۲۹۴, ۴/۱۱۱۵, ۴/۳۷۱۳)_T$
۱۱	۰/۳۳۲۱	۰/۳۷۵۷	۰/۴۰۳۶	$(-۰/۸۴۶۶, ۳/۳۳۴۵, ۳/۲۸۴۳)_T$
۱۲	۰/۳۸۸۰	۰/۴۰۶۰	۰/۴۲۱۶	$(-۱/۱۴۰۶, ۳/۲۹۰۵, ۳/۲۲۲۳)_T$
۱۳	۰/۴۰۳۵	۰/۴۶۳۰	۰/۴۹۸۵	$(۰/۹۴۴۴, ۳/۶۰۳۵, ۳/۶۶۰۳)_T$
۱۴	۰/۳۰۷۸	۰/۳۵۴۰	۰/۳۶۷۵	$(-۲/۱۴۳۶, ۳/۱۳۱۵, ۳/۰۱۱۳)_T$
۱۵	۰/۴۳۸۹	۰/۴۷۳۶	۰/۵۵۶۷	$(-۱/۰۵۲۶, ۳/۳۰۴۵, ۳/۲۴۰۳)_T$
۱۶	۰/۵۵۵۸	۰/۵۷۰۳	۰/۶۳۳۳	$(۱/۶۱۷۴, ۳/۷۰۴۵, ۳/۸۰۱۳)_T$
۱۷	۰/۴۵۷۹	۰/۴۸۱۵	۰/۵۳۷۶	$(۰/۲۳۹۴, ۳/۴۹۷۵, ۳/۵۱۲۳)_T$
۱۸	۰/۲۳۷۵	۰/۲۸۸۰	۰/۲۹۴۵	$(-۱/۷۷۴۶, ۳/۱۹۵۵, ۳/۰۸۹۳)_T$
۱۹	۰/۳۶۵۰	۰/۴۴۹۲	۰/۴۶۵۸	$(-۰/۹۸۱۶, ۳/۳۱۳۵, ۳/۲۵۴۳)_T$
۲۰	۰/۵۷۸۰	۰/۶۱۵۰	۰/۶۷۰۵	$(۲/۱۵۵۴, ۳/۷۸۵۵, ۳/۹۱۴۳)_T$
۲۱	۰/۲۴۹۸	۰/۳۰۸۱	۰/۳۰۷۹	$(-۲/۰۵۷۶, ۳/۱۵۳۵, ۳/۰۲۹۳)_T$
۲۲	۰/۴۹۶۹	۰/۵۱۵۸	۰/۵۳۳۷	$(۰/۲۲۷۴, ۳/۴۹۶۵, ۳/۵۰۹۳)_T$
۲۳	۰/۱۰۰۰	۰/۱۰۰۰	۰/۱۰۰۰	$(-۳/۸۲۰۶, ۲/۸۸۸۵, ۲/۶۵۹۳)_T$
۲۴	۰/۴۹۲۱	۰/۵۹۳۴	۰/۵۹۵۳	$(-۰/۲۳۶۶, ۳/۴۲۶۵, ۳/۴۱۲۳)_T$
۲۵	۰/۶۱۳۰	۰/۶۳۶۷	۰/۷۰۳۰	$(۱/۶۱۱۴, ۳/۷۰۳۵, ۳/۸۰۰۳)_T$
۲۶	۰/۴۲۰۶	۰/۴۳۵۵	۰/۴۵۴۶	$(۰/۳۷۷۴, ۳/۵۱۸۵, ۳/۵۴۱۳)_T$
۲۷	۰/۳۹۴۷	۰/۴۴۳۷	۰/۴۴۶۵	$(۰/۱۴۳۴, ۳/۴۸۳۵, ۳/۴۹۲۳)_T$
۲۸	۰/۵۶۵۶	۰/۶۳۸۱	۰/۶۸۲۳	$(۱/۴۶۷۴, ۳/۶۸۲۵, ۳/۷۷۰۳)_T$
۲۹	۰/۵۱۳۰	۰/۶۱۹۸	۰/۶۴۲۱	$(۰/۱۱۷۴, ۳/۴۷۹۵, ۳/۴۸۶۳)_T$
۳۰	۰/۶۱۶۷	۰/۶۷۸۸	۰/۷۴۴۸	$(۰/۰۸۹۴, ۳/۴۷۵۵, ۳/۴۸۰۳)_T$
۳۱	۰/۴۲۳۱	۰/۴۵۰۴	۰/۴۸۴۴	$(-۱/۰۳۳۶, ۳/۳۰۶۵, ۳/۲۴۴۳)_T$
۳۲	۰/۴۱۳۴	۰/۴۹۲۹	۰/۵۱۶۲	$(-۱/۲۷۸۶, ۳/۲۷۰۵, ۳/۱۹۳۳)_T$
۳۳	۰/۴۸۸۶	۰/۵۰۱۷	۰/۵۷۰۴	$(-۰/۲۲۰۶, ۳/۴۲۸۵, ۳/۴۱۵۳)_T$
۳۴	۰/۴۵۲۵	۰/۴۸۴۷	۰/۵۳۰۴	$(۰/۵۶۲۴, ۳/۵۴۶۵, ۳/۵۸۰۳)_T$
۳۵	۰/۵۶۷۸	۰/۶۴۰۳	۰/۶۹۶۵	$(۱/۳۴۰۴, ۳/۶۶۲۵, ۳/۷۴۳۳)_T$
۳۶	۰/۳۶۷۵	۰/۴۰۹۵	۰/۴۰۹۵	$(-۱/۱۱۳۶, ۳/۲۹۴۵, ۳/۲۲۸۳)_T$
۳۷	۰/۷۰۲۴	۰/۷۱۸۷	۰/۷۸۶۳	$(۳/۳۲۴۴, ۳/۹۶۰۵, ۴/۱۶۰۳)_T$
۳۸	۰/۵۳۲۲	۰/۵۹۴۷	۰/۶۴۰۵	$(۰/۵۸۲۴, ۳/۵۴۹۵, ۳/۵۸۴۳)_T$
۳۹	۰/۲۴۶۸	۰/۳۲۵۸	۰/۳۲۶۰	$(-۱/۰۷۲۶, ۳/۳۰۱۵, ۳/۲۳۶۳)_T$
۴۰	۰/۴۸۲۰	۰/۵۴۳۰	۰/۵۹۶۶	$(۱/۷۲۰۴, ۳/۷۲۰۵, ۳/۸۲۳۳)_T$
۴۱	۰/۵۹۱۹	۰/۶۱۶۱	۰/۶۷۳۳	$(۲/۰۵۹۴, ۳/۷۷۰۵, ۳/۸۹۴۳)_T$
۴۲	۰/۴۱۹۱	۰/۵۰۴۹	۰/۵۱۹۷	$(-۰/۶۶۲۶, ۳/۳۶۲۵, ۳/۳۲۲۳)_T$
۴۳	۰/۵۲۸۵	۰/۵۶۸۹	۰/۶۱۱۷	$(۱/۷۶۴۴, ۳/۷۲۶۵, ۳/۸۳۲۳)_T$
۴۴	۰/۴۵۳۹	۰/۴۵۵۹	۰/۵۲۹۸	$(-۰/۱۷۹۶, ۳/۴۳۵۵, ۳/۴۲۴۳)_T$
۴۵	۰/۴۴۶۴	۰/۴۸۴۳	۰/۵۴۰۰	$(۰/۹۵۳۴, ۳/۶۰۵۵, ۳/۶۶۲۳)_T$
۴۶	۰/۳۶۰۰	۰/۳۶۶۸	۰/۳۹۷۳	$(-۱/۰۹۳۶, ۳/۲۹۷۵, ۳/۲۳۲۳)_T$
۴۷	۰/۴۰۶۵	۰/۴۳۷۶	۰/۴۶۸۰	$(-۱/۳۱۴۶, ۳/۲۶۴۵, ۳/۱۸۵۳)_T$
۴۸	۰/۲۹۴۷	۰/۳۴۸۹	۰/۳۶۹۳	$(-۲/۳۴۲۶, ۳/۱۱۰۵, ۲/۹۷۰۳)_T$
۴۹	۰/۶۴۸۸	۰/۷۰۳۴	۰/۷۶۶۲	$(۱/۸۰۰۴, ۳/۷۳۲۵, ۳/۸۴۰۳)_T$
۵۰	۰/۶۰۶۳	۰/۶۲۷۶	۰/۶۷۳۸	$(۰/۶۰۵۴, ۳/۵۵۲۵, ۳/۵۸۹۳)_T$



(الف) مسیرهای a_{1m} ، a_{2m} و a_{3m}

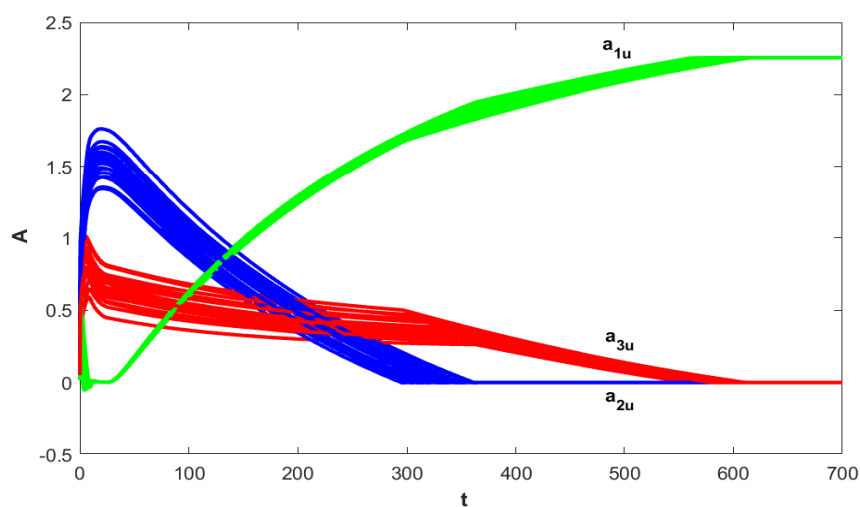


(ب) مسیرهای a_{1l} ، a_{2l} و a_{3l}

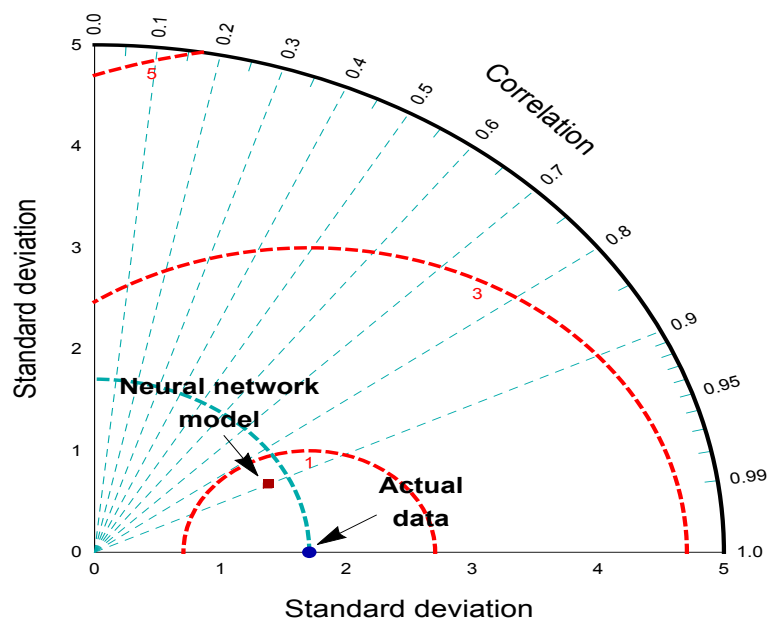


(ج) مسیرهای a_{1u} ، a_{2u} و a_{3u}

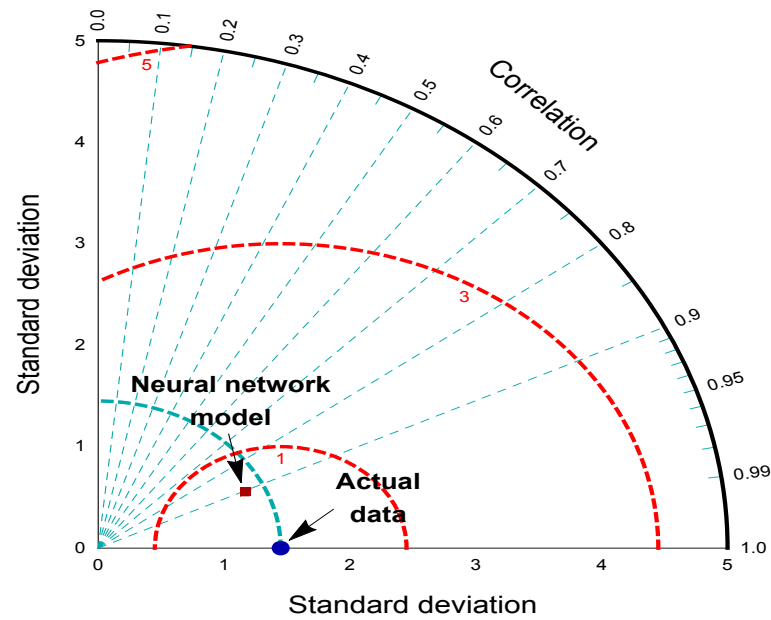
شکل ۲.۳: رفتار مسیرهای A_1 ، A_2 و A_3 از شبکه عصبی (۱۰.۳) و (۱۱.۳) در مثال ۱.۴.۳.



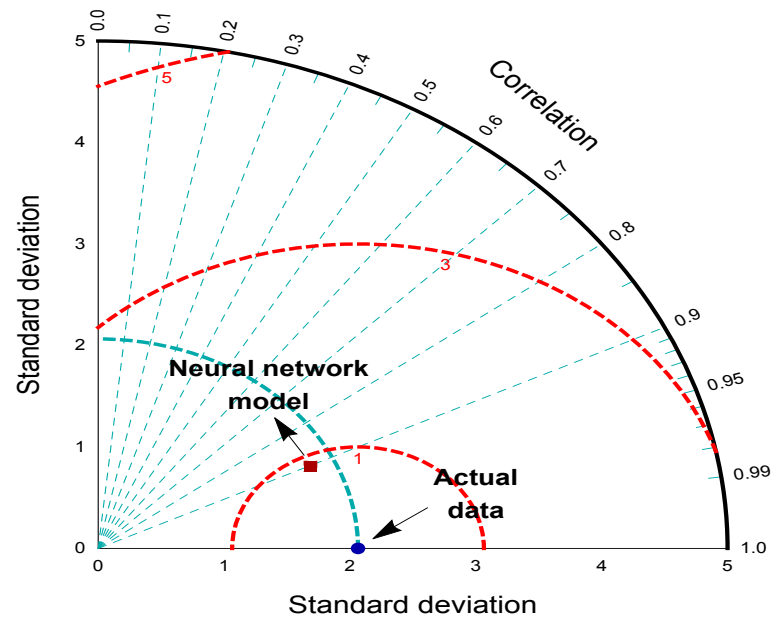
شکل ۳.۳: رفتار مسیرهای a_{1u} ، a_{2u} و a_{3u} با φ_0 نقطه شروع متفاوت از شبکه عصبی (۱۰.۳) و (۱۱.۳) در مثال ۱.۴.۳.



شکل ۴.۳: نمودار تیلور برای مراکز در مثال ۱.۴.۳.



(آ) نمودار تیلور برای مقادیر سمت چپ اعداد فازی.



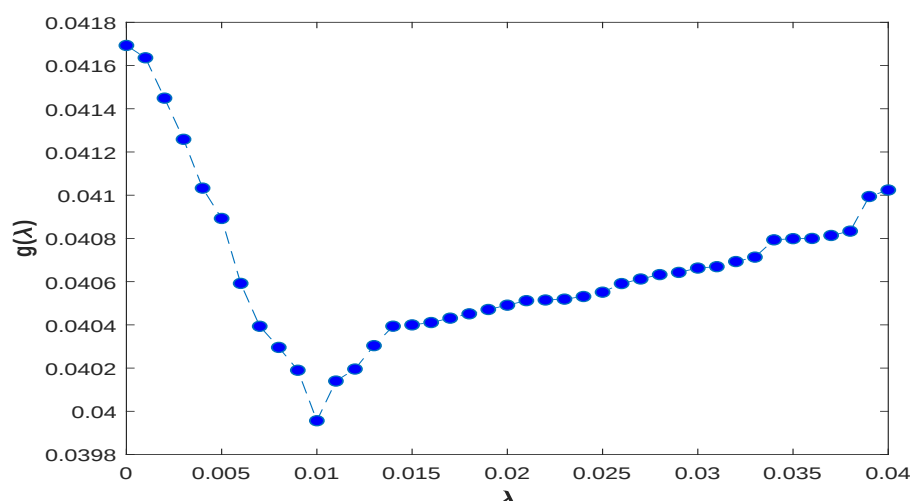
(ب) نمودار تیلور برای مقادیر سمت راست اعداد فازی.
شکل ۵.۳: نمودار تیلور در مثال ۱.۱.۲.

جدول ۲.۳: تخمین مقادیر MCV برای داده‌های مثال ۱.۴.۳

شماره	D_+	D_+	شماره	D_+	D_+
۱	۱/۶۸۷	۰/۵۴۳	۲۶	۱/۶۹۰	۰/۵۴۴
۲	۱/۷۲۰	۰/۵۵۳	۲۷	۱/۶۸۹	۰/۵۴۴
۳	۱/۶۸۷	۰/۵۴۳	۲۸	۱/۶۸۷	۰/۵۴۳
۴	۱/۶۸۸	۰/۵۴۳	۲۹	۱/۶۸۸	۰/۵۴۳
۵	۱/۶۸۷	۰/۵۴۳	۳۰	۱/۷۰۰	۰/۵۴۷
۶	۱/۶۸۷	۰/۵۴۳	۳۱	۱/۶۸۷	۰/۵۴۳
۷	۱/۶۹۷	۰/۵۴۶	۳۲	۱/۶۸۹	۰/۵۴۳
۸	۱/۶۸۷	۰/۵۴۳	۳۳	۱/۶۸۸	۰/۵۴۳
۹	۱/۶۸۹	۰/۵۴۳	۳۴	۱/۶۸۷	۰/۵۴۳
۱۰	۱/۶۸۸	۰/۵۴۳	۳۵	۱/۶۸۷	۰/۵۴۳
۱۱	۱/۶۸۷	۰/۵۴۳	۳۶	۱/۶۸۷	۰/۵۴۳
۱۲	۱/۶۸۸	۰/۵۴۳	۳۷	۱/۶۹۰	۰/۵۴۴
۱۳	۱/۶۹۱	۰/۵۴۴	۳۸	۱/۶۸۷	۰/۵۴۳
۱۴	۱/۶۸۸	۰/۵۴۳	۳۹	۱/۶۹۴	۰/۵۴۵
۱۵	۱/۷۰۴	۰/۵۴۸	۴۰	۱/۶۹۳	۰/۵۴۵
۱۶	۱/۶۸۸	۰/۵۴۳	۴۱	۱/۶۸۸	۰/۵۴۳
۱۷	۱/۶۸۷	۰/۵۴۳	۴۲	۱/۶۸۷	۰/۵۴۳
۱۸	۱/۶۸۸	۰/۵۴۳	۴۳	۱/۶۸۸	۰/۵۴۳
۱۹	۱/۶۸۶	۰/۵۴۳	۴۴	۱/۶۸۷	۰/۵۴۳
۲۰	۱/۶۸۸	۰/۵۴۳	۴۵	۱/۶۹۰	۰/۵۴۴
۲۱	۱/۶۸۷	۰/۵۴۳	۴۶	۱/۶۸۶	۰/۵۴۳
۲۲	۱/۶۸۷	۰/۵۴۳	۴۷	۱/۶۸۸	۰/۵۴۳
۲۳	۱/۶۸۷	۰/۵۴۳	۴۸	۱/۶۸۸	۰/۵۴۳
۲۴	۱/۶۸۹	۰/۵۴۴	۴۹	۱/۶۸۷	۰/۵۴۳
۲۵	۱/۶۸۷	۰/۵۴۳	۵۰	۱/۶۹۴	۰/۵۴۵
MCV			۱/۶۸۹ ۰/۵۴۳		

جدول ۳.۳: مقادیر مشاهده شده در مثال ۲.۴.۳

شماره	x_1	x_2	Y
۱	۰/۰۱۰۰	۰/۰۱۰۰	$(-۱۸/۸, ۱۰/۳۳۳۳, ۱۰/۳۳۳۳)_N$
۲	۰/۱۲۷۸	۰/۱۱۶۶	$(-۱۴/۳, ۱۰/۳۳۳۳, ۱۰/۳۳۳۳)_N$
۳	۰/۱۶۳۲	۰/۱۵۲۲	$(-۱۲/۸, ۱۰/۳۳۳۳, ۱۰/۳۳۳۳)_N$
۴	۰/۲۴۵۶	۰/۲۳۵۱	$(-۹/۳, ۱۰/۳۳۳۳, ۱۰/۳۳۳۳)_N$
۵	۰/۲۹۲۸	۰/۲۸۲۵	$(-۷/۳, ۱۰/۳۳۳۳, ۱۰/۳۳۳۳)_N$
۶	۰/۳۵۱۷	۰/۳۴۱۷	$(-۴/۸, ۱۰/۳۳۳۳, ۱۰/۳۳۳۳)_N$
۷	۰/۳۸۷۰	۰/۳۷۷۲	$(-۳/۳, ۱۰/۳۳۳۳, ۱۰/۳۳۳۳)_N$
۸	۰/۴۶۹۵	۰/۴۶۰۱	$(۰/۲, ۱۰/۳۳۳۳, ۱۰/۳۳۳۳)_N$
۹	۰/۵۱۶۶	۰/۵۰۷۵	$(۲/۲, ۱۰/۳۳۳۳, ۱۰/۳۳۳۳)_N$
۱۰	۰/۵۹۹۱	۰/۵۹۰۴	$(۵/۷, ۱۰/۳۳۳۳, ۱۰/۳۳۳۳)_N$
۱۱	۰/۶۳۴۴	۰/۶۲۶۰	$(۷/۲, ۱۱/۳۳۳۳, ۱۱/۳۳۳۳)_N$
۱۲	۰/۷۰۵۱	۰/۶۹۷۰	$(۱۰/۲, ۱۱/۳۳۳۳, ۱۱/۳۳۳۳)_N$
۱۳	۰/۷۵۲۳	۰/۷۴۴۴	$(۱۲/۲, ۱۱/۳۳۳۳, ۱۱/۳۳۳۳)_N$
۱۴	۰/۸۲۳۰	۰/۸۱۵۵	$(۱۵/۲, ۱۱/۳۳۳۳, ۱۱/۳۳۳۳)_N$
۱۵	۰/۸۸۱۹	۰/۸۷۴۷	$(۱۷/۷۱۱/۳۳۳۳, ۱۱/۳۳۳۳)_N$



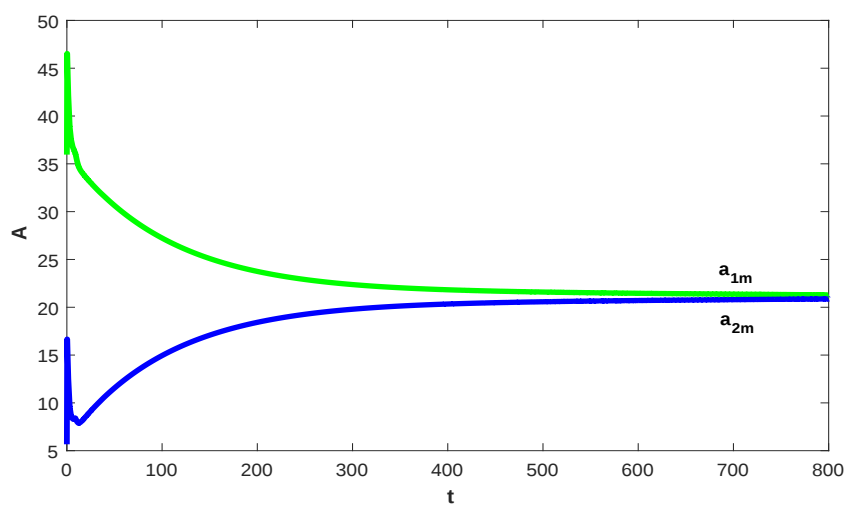
شکل ۶.۳: مقادیر مختلف λ با فاصله D_1 با استفاده از روش ارایه شده در [۸۹] در مثال ۲.۴.۳.

جدول ۴.۳: تخمین مقادیر MCV برای داده‌های مثال ۱.۴.۳

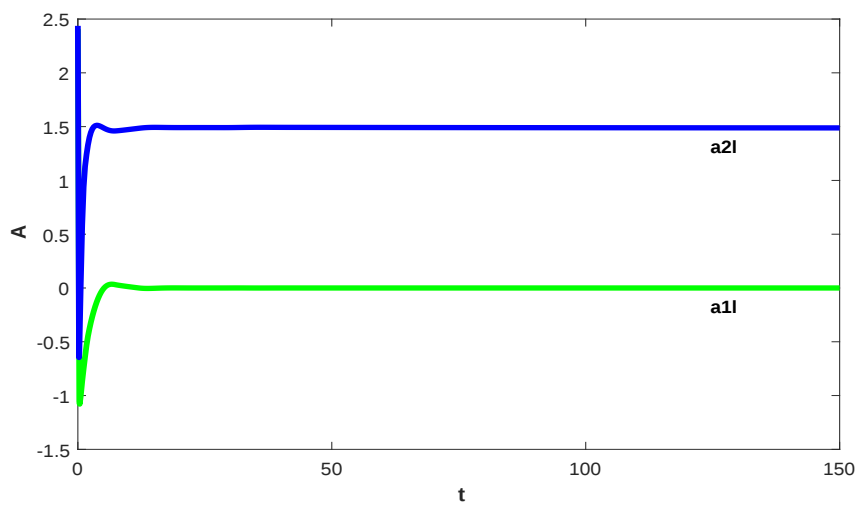
شماره	Y	فاصله D_1 میان Y_i و $\hat{Y}_{(-i)}(x_i)$
۱	$(12/5, 5, 5)_N$	۰/۰۴۲۹
۲	$(17, 5, 5)_N$	۰/۰۴۰۰
۳	$(18/5, 5, 5)_N$	۰/۰۳۹۷
۴	$(22, 5, 5)_N$	۰/۰۳۹۳
۵	$(24, 5, 5)_N$	۰/۰۳۹۲
۶	$(26/5, 5, 5)_N$	۰/۰۳۹۲
۷	$(28, 5, 5)_N$	۰/۰۳۹۲
۸	$(31/5, 5, 5)_N$	۰/۰۳۹۴
۹	$(33/5, 5, 5)_N$	۰/۰۳۹۶
۱۰	$(37, 5, 5)_N$	۰/۰۴۰۳
۱۱	$(38/5, 6, 6)_N$	۰/۰۳۹۸
۱۲	$(41/5, 6, 6)_N$	۰/۰۳۹۶
۱۳	$(43/5, 6, 6)_N$	۰/۰۳۹۵
۱۴	$(46/5, 6, 6)_N$	۰/۰۳۹۳
۱۵	$(49, 6, 6)_N$	۰/۰۳۹۲
MCV		۰/۰۳۹۷

جدول ۵.۳: تخمین مقادیر $MSPE$ برای داده‌های مثال ۲.۴.۳

$MSPE$	مدل ارائه شده در [۸۶]	مدل ارائه شده در [۸۷]	مدل پیشنهادی
D_1	۰/۳۳۲۰	۳۰/۴۶۰۰	۰/۰۱۵۷
D_2	۲/۰۳۶۸	۳۸/۷۷۴۰	۰/۰۱۶۱۳

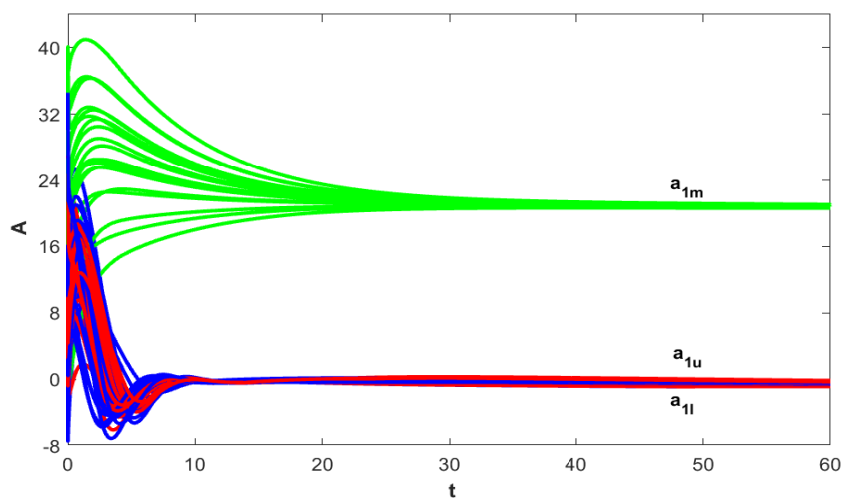


(ا) مسیر A_1

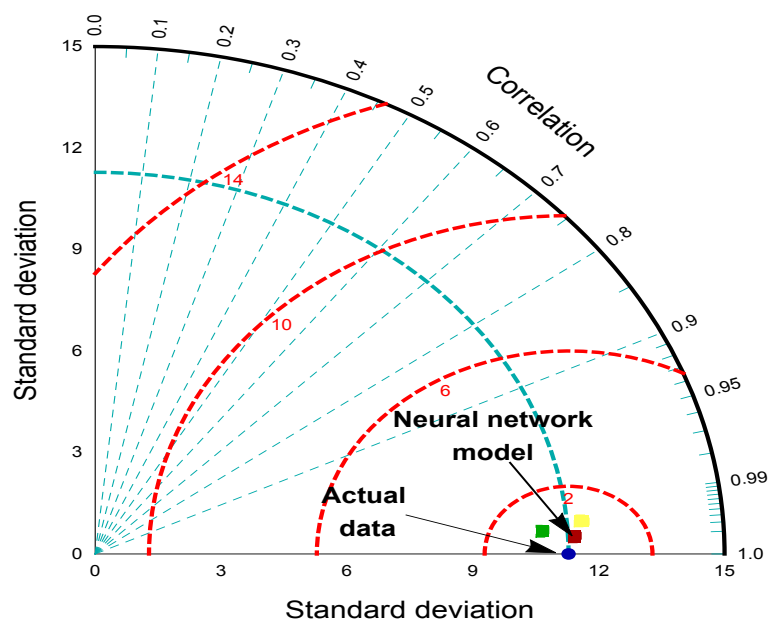


(ب) مسیر A_2

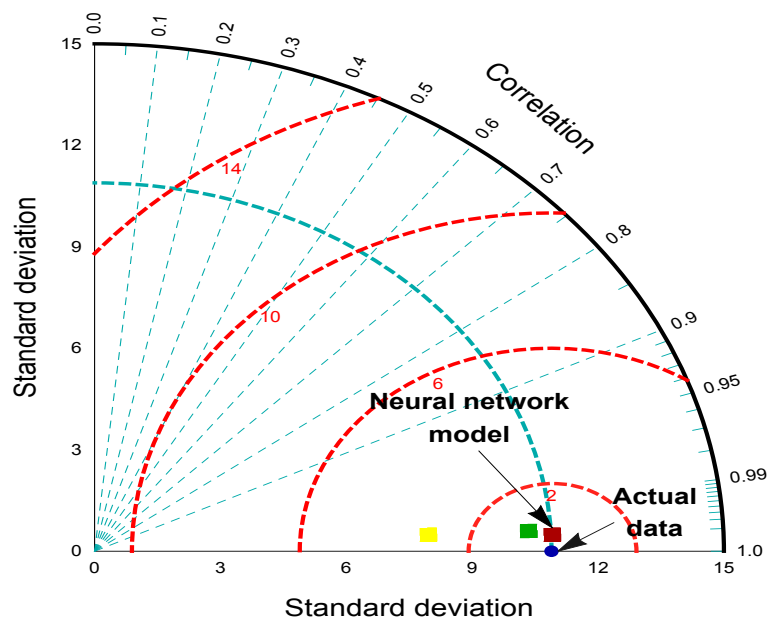
شکل ۷.۳: رفتار مسیر A از شبکه عصبی (۱۰.۳) و (۱۱.۳) در مثال ۲.۴.۳.



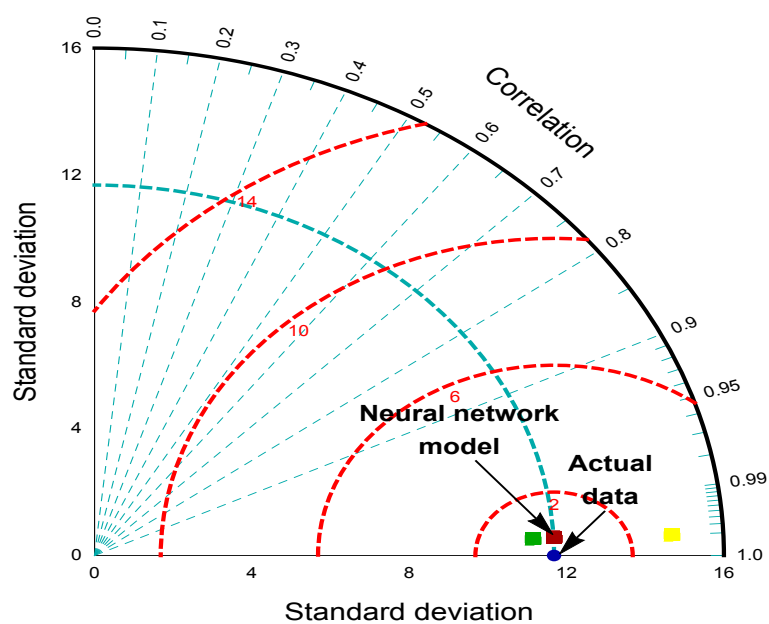
شکل ۸.۳: رفتار مسیرهای A_1 با ۲۰ نقطه شروع اولیه از شبکه عصبی (۱۰.۳) و (۱۱.۳) در مثال ۲.۴.۳.



شکل ۹.۳: نمودار تیلور برای مراکز اعداد فازی در مثال ۲.۴.۳. مدل [۸۶] با مربع سبز و مدل [۸۷] با مربع زرد نشان داده شده است.



(آ) نمودار تیلور برای مقادیر سمت چپ اعداد فازی.



(ب) نمودار تیلور برای مقادیر سمت راست اعداد فازی.

شکل ۱۰.۳: نمودار تیلور در مثال ۲.۴.۳. مدل [۸۶] با مربع سبز و مدل [۸۷] با مربع زرد نشان داده شده است.

۵.۳ مزایای روش پیشنهادی

علاوه بر مزایای شماره ۱ تا ۹ بیان شده در ۴.۲.۲، برخی دیگر از مزایای روش پیشنهادی را به شرح زیر خلاصه می‌کنیم:

- در روش‌های ارائه شده در [۲۷-۲۹] نویسندگان مسئله را برای تعداد محدود h برش مدل‌سازی می‌کنند. با این حال، در مدل ارائه شده برای مدل‌سازی مسئله از فاصله‌ای استفاده می‌کنیم که به h وابسته نیست.
- به منظور بهره‌وری بیشتر، اعداد فازی در این بخش به شکل LR در نظر گرفته می‌شوند. برای روشن‌تر شدن این موضوع، در انتهای این بخش یک مثال با اعداد فازی مثلثی و یک مثال دیگر با اعداد فازی نرمال ارائه شده است.
- برای نشان دادن برخی از مزایای مدل پیشنهادی نسبت به دیگر مدل‌ها، از نمودار آماری شناخته شده‌ای به نام «نمودار تیلور» استفاده می‌کنیم.
- برای به تصویر کشیدن عملکرد روش پیشنهادی در حضور هم‌خطی، شبیه‌سازی مونت کارلو ارائه شده است. برای تأکید بیشتر، یک مثال از وضعیتی که در آن یک متغیر ترکیب خطی کامل از متغیر دیگر است نیز بررسی می‌شود.

فصل ۴

حل عددی رگرسیون بریج فازی برای مقادیر مختلف γ

به عنوان تعمیمی از روش رگرسیون ریج برای غلبه بر مشکل هم‌خطی نویسندگان در [۱۲] رگرسیون بریج را معرفی کرده‌اند. رگرسیون بریج با موفقیت برای حل دامنه وسیعی از مسائل رگرسیون با پیچیدگی‌های متفاوت به کار برده شده است [۹۱-۹۸]. با قرار دادن $\gamma = 2$ در مدل بریج، مسئله به یک مدل رگرسیون ریج تبدیل می‌شود. همچنین به عنوان یک مورد خاص از رگرسیون بریج (با $\gamma = 1$) نویسندگان در [۱۳] مدل رگرسیون لاسو را معرفی کرده‌اند. برای اطلاعات بیشتر در مورد این مدل به [۹۹-۱۰۱] مراجعه شود.

۱.۴ مدل رگرسیون بریج فازی پیشنهادی

همانند مدل‌های رگرسیون معمولی، هم‌خطی در مدل‌های رگرسیون فازی نیز مسئله مهمی است. بر این اساس، مدل رگرسیون فازی بریج که ترکیبی از رگرسیون بریج با مدل رگرسیون فازی است به منظور کاهش اثر هم‌خطی در مدل فازی برای اولین بار در این رساله در نظر گرفته می‌شود.

با توجه به توضیحات ارائه شده در ۱.۲.۱ به ازای هر $\gamma \geq 0$ مدل رگرسیون بریج زیر را داریم:

$$\text{minimize } \sum_{i=1}^n (Y_i - \sum_{j=0}^p A_j x_{ij})^2 \quad (1.4)$$

subject to

$$\sum_{j=0}^p |A_j|^\gamma \leq t \quad (2.4)$$

که در آن پارامتر مثبت t نشان دهنده ثابت تنظیم است. همچنین، انتخاب بهینه γ باعث افزایش کارایی مدل میشود. نشان داده شده است که برای $\gamma \geq 1$ مدل رگرسیون بریج یک مسئله محدب است (برای توضیحات بیشتر به [۹۲] مراجعه شود). علاوه بر این، با فرض $\gamma = 0$ ، $\gamma = 1$ و $\gamma = 2$ در مسئله بهینه سازی (۱.۴) و (۲.۴)، به ترتیب، مدل رگرسیون کمترین توان‌های دوم، مدل رگرسیون لاسو و مدل رگرسیون ریج به دست می‌آید.

در این بخش، به منظور توسعه مدل رگرسیون بریج برای اعداد فازی، مجموعه داده‌های شماره (۱) ذکر شده در ۱.۲.۱ در نظر گرفته شده است. سپس، یک روش پیشنهادی برای حل مشکل هم‌خطی با استفاده از مدل رگرسیون بریج فازی با اعداد فازی LR بحث شده است. تحت مفروضات مدل (۷.۱)، فرض کنید تابع مشخصه I_{ij} تنها زمانی برابر با یک است که $x_{ij} \geq 0$. بنابراین عدد فازی LR ، $\hat{Y}_i$ به صورت زیر تعریف می‌شود:

$$\hat{Y}_i = \left(\sum_{j=0}^p a_{jm} x_{ij}, \sum_{j=0}^p (I_{ij} a_{jl} x_{ij} - (1 - I_{ij}) a_{ju} x_{ij}), \sum_{j=0}^p (I_{ij} a_{ju} x_{ij} - (1 - I_{ij}) a_{jl} x_{ij}) \right)_{LR}$$

برای $i = 1, \dots, n$ و $j = 0, 1, \dots, p$.

اکنون، مدل چند هدفه پیشنهادی برای حل مسئله رگرسیون بریج فازی، با استفاده از روش کمترین توان‌های دوم برای مرکز و پهنای چپ و راست به صورت زیر ارائه می‌شود:

$$\text{minimize } \sum_{i=1}^n (y_{im} - \sum_{j=0}^p a_{jm} x_{ij})^2 \quad (3.4)$$

$$\text{minimize } \sum_{i=1}^n \left(y_{il} - \sum_{j=0}^p (I_{ij} a_{jl} x_{ij} - (1 - I_{ij}) a_{ju} x_{ij}) \right)^2 \quad (4.4)$$

$$\text{minimize } \sum_{i=1}^n \left(y_{iu} - \sum_{j=0}^p (I_{ij} a_{ju} x_{ij} - (1 - I_{ij}) a_{jl} x_{ij}) \right)^2 \quad (5.4)$$

subject to

$$\sum_{j=0}^p |a_{jm}|^\gamma \leq t_1 \quad (6.4)$$

$$\sum_{j=0}^p (a_{jl})^\gamma \leq t_2 \quad a_{jl} \geq 0, \quad (7.4)$$

$$\sum_{j=0}^p (a_{ju})^\gamma \leq t_3 \quad a_{ju} \geq 0, \quad (8.4)$$

که در آن، t_1, t_2 و t_3 نشان دهنده ثابت‌های تنظیم هستند. لازم به ذکر است که استفاده از t_s های مختلف برای $s = 1, 2, 3$ نتایج بهتری را نسبت به زمانی که تنها یک t را در کل محاسبات استفاده می‌کنیم، در پی دارد (برای اطلاعات بیشتر به [۸۹] مراجعه شود).

۲.۴ مدل بهینه سازی

به منظور به دست آوردن ضرایب $A_0, A_1, \dots, A_p$ از مدل رگرسیون بریج فازی، به دنبال کمینه کردن هم‌زمان توابع هدف (۳.۴) – (۵.۴) تحت قیود (۶.۴) – (۸.۴) هستیم. برای این منظور با انتخاب روش مجموع وزن دار^۱ از میان روش های موجود تک هدف سازی^۲ به مسئله زیر می‌رسیم (برای اطلاعات بیشتر به [۱۰۲، ۱۰۳] مراجعه شود):

$$\begin{aligned} \text{minimize } C(\mathbf{A}) = & w_1 \sum_{i=1}^n \left(y_{im} - \sum_{j=0}^p a_{jm} x_{ij} \right)^2 \\ & + w_2 \sum_{i=1}^n \left(y_{il} - \sum_{j=0}^p (I_{ij} a_{jl} x_{ij} - (1 - I_{ij}) a_{ju} x_{ij}) \right)^2 \\ & + w_3 \sum_{i=1}^n \left(y_{iu} - \sum_{j=0}^p (I_{ij} a_{ju} x_{ij} - (1 - I_{ij}) a_{jl} x_{ij}) \right)^2 \end{aligned} \quad (9.4)$$

subject to

$$\sum_{j=0}^p |a_{jm}|^\gamma \leq t_1 \quad (10.4)$$

$$\sum_{j=0}^p (a_{jl})^\gamma \leq t_2 \quad a_{jl} \geq 0, \quad (11.4)$$

$$\sum_{j=0}^p (a_{ju})^\gamma \leq t_3 \quad a_{ju} \geq 0, \quad (12.4)$$

که در آن

$$\mathbf{A} = (a_{0m}, \dots, a_{pm}, a_{0l}, \dots, a_{pl}, a_{0u}, \dots, a_{pu})^\top \quad (13.4)$$

بردار ستونی $(3p+3)$ بعدی است. همچنین $w_s \geq 0$ برای $s = 1, 2, 3$ و $\sum_{s=1}^3 w_s = 1$. در ادامه برای حل مدل (۹.۴) – (۱۲.۴) از گزاره زیر استفاده می‌کنیم:

گزاره ۱.۲.۴. [۱۰۴] فرض کنید $u, v \in \mathbb{R}^n$ متغیرهای کمکی باشند در این صورت:

$$|x| = u - v \quad u \geq 0, v \geq 0$$

^۱Sum weighted

^۲Scalarization

بنابراین، مقدار قدر مطلق موجود در مدل بهینه سازی به $a'_{jm} - a''_{jm}$ با $a'_{jm}, a''_{jm} \geq 0$ تبدیل می شود. در نتیجه مسئله بهینه سازی درجه دوم زیر برای $\gamma \geq 1$ به دست می آید:

$$\text{minimize } C_{\text{Bridge}}(\bar{\mathbf{A}}) = \bar{\mathbf{A}}^\top \bar{\mathbf{Q}} \bar{\mathbf{A}} + \bar{\mathbf{B}}^\top \bar{\mathbf{A}} + M \quad (14.4)$$

subject to

$$\bar{g}_q(\bar{\mathbf{A}}) \leq 0, \quad q = 1, 2, 3 \quad (15.4)$$

$$\bar{\mathbf{G}} \bar{\mathbf{A}} \leq 0 \quad (16.4)$$

که در آن

$$\bar{\mathbf{A}} = (a'_{om}, \dots, a'_{pm}, a''_{om}, \dots, a''_{pm}, a_{ol}, \dots, a_{pl}, a_{ou}, \dots, a_{pu})^\top, \quad (17.4)$$

یک بردار $(4p+4)$ بعدی، $w_s \geq 0$ و $\sum_{s=1}^3 w_s = 1$ برای $s = 1, 2, 3$. در اینجا

$$\bar{\mathbf{Q}} = \begin{bmatrix} w_1 \mathbf{P}_1 & -2w_1 \mathbf{P}_1 & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & w_1 \mathbf{P}_1 & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & (w_2 \mathbf{P}_2 + w_3 \mathbf{P}_2) & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & (w_2 \mathbf{P}_2 + w_3 \mathbf{P}_2) \end{bmatrix} \quad (18.4)$$

و

$$\mathbf{P}_1 = \begin{bmatrix} \sum_{i=1}^n x_{i0} & \sum_{i=1}^n x_{i1} & \dots & \sum_{i=1}^n x_{ip} \\ \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1}^2 & \dots & \sum_{i=1}^n x_{i1} x_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n x_{ip} & \sum_{i=1}^n x_{ip} x_{i1} & \dots & \sum_{i=1}^n x_{ip}^2 \end{bmatrix} \quad (19.4)$$

$$\mathbf{P}_2 = \begin{bmatrix} \sum_{i=1}^n I_{i0} x_{i0} & \sum_{i=1}^n I_{i1} x_{i1} & \dots & \sum_{i=1}^n I_{ip} x_{ip} \\ \sum_{i=1}^n I_{i1} x_{i1} & \sum_{i=1}^n I_{i1} x_{i1}^2 & \dots & \sum_{i=1}^n I_{i1} x_{i1} I_{ip} x_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n I_{ip} x_{ip} & \sum_{i=1}^n I_{ip} x_{ip} I_{i1} x_{i1} & \dots & \sum_{i=1}^n I_{ip} x_{ip}^2 \end{bmatrix} \quad (20.4)$$

$$\mathbf{P}_3 = \begin{bmatrix} \sum_{i=1}^n (1 - I_{i0}) x_{i0} & \sum_{i=1}^n (1 - I_{i1}) x_{i1} & \dots & \sum_{i=1}^n (1 - I_{ip}) x_{ip} \\ \sum_{i=1}^n (1 - I_{i1}) x_{i1} & \sum_{i=1}^n (1 - I_{i1}) x_{i1}^2 & \dots & \sum_{i=1}^n (1 - I_{i1}) x_{i1} (1 - I_{ip}) x_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n (1 - I_{ip}) x_{ip} & \sum_{i=1}^n (1 - I_{ip}) x_{ip} (1 - I_{i1}) x_{i1} & \dots & \sum_{i=1}^n (1 - I_{ip}) x_{ip}^2 \end{bmatrix} \quad (21.4)$$

همچنین O یک ماتریس $(p+1) \times (p+1)$ صفر، $\bar{G}$ یک ماتریس قطری $(4p+4) \times (4p+4)$ با درایه های قطری (-1) و $\bar{B}$ یک بردار قطری به شکل زیر است:

$$\bar{B} = (b'_{10}, b'_{11}, \dots, b'_{1p}, b''_{10}, \dots, b''_{1p}, b_{20}, \dots, b_{2p}, b_{30}, \dots, b_{3p})^T \quad (22.4)$$

که در آن

$$b'_{1j} = -2w_1 \sum_{i=1}^n x_{ij}y_{im}$$

$$b''_{1j} = 2w_1 \sum_{i=1}^n x_{ij}y_{im}$$

$$b_{2j} = -2w_2 \sum_{i=1}^n I_{ij}x_{ij}y_{il} + 2w_3 \sum_{i=1}^n (1 - I_{ij})x_{ij}y_{iu}$$

$$b_{3j} = -2w_3 \sum_{i=1}^n I_{ij}x_{ij}y_{iu} + 2w_2 \sum_{i=1}^n (1 - I_{ij})x_{ij}y_{il}$$

و

$$\bar{g}_1 = \sum_{j=0}^p (a'_{jm} - a''_{jm})^\gamma - t_1 \quad (23.4)$$

$$\bar{g}_2 = \sum_{j=0}^p (a_{jl})^\gamma - t_2 \quad (24.4)$$

$$\bar{g}_3 = \sum_{j=0}^p (a_{ju})^\gamma - t_3 \quad (25.4)$$

در نهایت M به صورت زیر تعریف می شود:

$$M = w_1 \sum_{i=1}^n y_{im}^2 + w_2 \sum_{i=1}^n y_{il}^2 + w_3 \sum_{i=1}^n y_{iu}^2 \quad (26.4)$$

به راحتی نتیجه می شود که ماتریس هسین $C_{Bridge}(\bar{A})$ معین مثبت است زیرا:

$$\nabla^2 C_{Bridge}(\bar{A}) = \bar{Q}$$

بنابراین مسئله بهینه سازی (۱۴.۴) - (۱۶.۴) به یک مسئله درجه دوم محدب با قيود نامساوی تبدیل می شود.

۳.۴ مدل رگرسیون بریج فازی برای $\gamma = 1$ و $\gamma = 2$

در این بخش، برای دستیابی به دو مدل شناخته شده لاسو و ریج از مدل رگرسیون بریج، قرار می دهیم $\gamma = 1$ و $\gamma = 2$. بنابراین، مسائل بهینه سازی درجه دوم زیر به ترتیب به دست می آیند:

$$\text{minimize } C_{Lasso}(\bar{A}) = \bar{A}^T \bar{Q} \bar{A} + \bar{B}^T \bar{A} + M \quad (27.4)$$

subject to

$$\bar{g}_q(\bar{\mathbf{A}}) \leq 0, \quad q = 1, 2, 3 \quad (28.4)$$

$$\bar{\mathbf{G}}\bar{\mathbf{A}} \leq 0 \quad (29.4)$$

$$\text{minimize } C_{\text{Ridge}}(\mathbf{A}) = \mathbf{A}^\top \mathbf{Q} \mathbf{A} + \mathbf{B}^\top \mathbf{A} + M \quad (30.4)$$

subject to

$$g_q(\mathbf{A}) \leq 0, \quad q = 1, 2, 3 \quad (31.4)$$

$$\mathbf{G} \mathbf{A} \leq 0 \quad (32.4)$$

که در آن $\mathbf{Q}$ ماتریس بلوکی

$$\mathbf{Q} = \begin{bmatrix} w_1 \mathbf{P}_1 & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & (w_3 \mathbf{P}_3 + w_2 \mathbf{P}_2) & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & (w_2 \mathbf{P}_3 + w_3 \mathbf{P}_2) \end{bmatrix}$$

$\mathbf{G}$ یک ماتریس $(3p+3) \times (3p+3)$ به شکل

$$\mathbf{G} = \begin{bmatrix} 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & -1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & -1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 & -1 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & -1 \end{bmatrix}$$

$\mathbf{O}$ یک ماتریس $(p+1) \times (p+1)$ صفر، و $\bar{\mathbf{G}}$ یک ماتریس قطری $(4p+4) \times (4p+4)$ با درایه‌های قطر اصلی (-1) است. از طرفی، $\mathbf{B}$ بردار ستونی به شکل زیر است:

$$\mathbf{B} = (b_{10}, b_{11}, \dots, b_{1p}, b_{20}, \dots, b_{2p}, b_{30}, \dots, b_{3p})^\top \quad (33.4)$$

که در آن

$$b_{1j} = -2w_1 \sum_{i=1}^n x_{ij} y_{im}$$

$$b_{2j} = -2w_2 \sum_{i=1}^n I_{ij} x_{ij} y_{il} + 2w_3 \sum_{i=1}^n (1 - I_{ij}) x_{ij} y_{iu}$$

$$b_{3j} = -2w_3 \sum_{i=1}^n I_{ij} x_{ij} y_{iu} + 2w_2 \sum_{i=1}^n (1 - I_{ij}) x_{ij} y_{il}$$

$$\bar{g}_1 = \sum_{j=0}^p (a'_{jm} - a''_{jm}) - t_1 \quad (34.4)$$

$$\bar{g}_2 = \sum_{j=0}^p (a_{jl}) - t_2 \quad (35.4)$$

$$\bar{g}_3 = \sum_{j=0}^p (a_{ju}) - t_3 \quad (36.4)$$

$$g_1 = \sum_{j=0}^p (a_{jm})^2 - t_1 \quad (37.4)$$

$$g_2 = \sum_{j=0}^p (a_{jl})^2 - t_2 \quad (38.4)$$

$$g_3 = \sum_{j=0}^p (a_{ju})^2 - t_3 \quad (39.4)$$

در مدل بالا $\bar{A}$ ، A ، $\bar{Q}$ ، P_1 ، P_2 ، P_3 ، $\bar{B}$ و M به ترتیب در (۱۳.۴)، (۱۷.۴)، (۱۸.۴)، (۱۹.۴)، (۲۰.۴)، (۲۱.۴)، (۲۲.۴) و (۲۶.۴) تعریف شده اند.

هسین ماتریس‌های $C_{Lasso}(\bar{A})$ و $C_{Ridge}(A)$ معین مثبت هستند، زیرا

$$\nabla^2 C_{Lasso}(\bar{A}) = \bar{Q}$$

$$\nabla^2 C_{Ridge}(A) = Q$$

بنابراین مدل‌های (۲۷.۴)–(۲۹.۴) و (۳۰.۴)–(۳۲.۴) به دو مسئله بهینه سازی درجه دوم محدب با قیود نامساوی تبدیل می‌شوند.

چندین روش بهینه سازی برای حل مسائل درجه دوم محدب با قیود نامساوی وجود دارد که هر یک از این روش‌ها نقاط ضعف و قوت خود را دارند. در بخش بعد با تکیه بر مطالب ارائه شده در ۴.۳.۱ به معرفی یک شبکه عصبی مناسب برای حل این مسائل می‌پردازیم.

۴.۴ رویکرد شبکه عصبی بازگشتی

در این قسمت به دنبال معرفی یک شبکه عصبی مناسب برای حل مسئله بهینه سازی درجه دوم محدب (۱۴.۴)–(۱۶.۴) هستیم. همانند روش اجرا شده در بخش ۴.۳.۱ در این بخش نیز در تلاش می‌کنیم که با استفاده از تکنیک‌های بهینه سازی، مسئله برنامه‌ریزی (۱۴.۴)–(۱۶.۴) را به یک سیستم غیر خطی تبدیل کنیم. برای تمرکز بیشتر بر روی مباحث ارائه شده تمام اثبات‌های این قسمت به پیوست ب انتقال داده شده است.

می‌دانیم که $\bar{A}^* \in \mathbb{R}^{p+4}$ یک جواب بهینه از (۱۴.۴)–(۱۶.۴) است، اگر و تنها اگر وجود داشته باشد $\bar{U}^* \in \mathbb{R}^{p+4}$ و $\bar{U}_q^* \in \mathbb{R}^3$ به طوری که $((\bar{A}^*)^\top, (\bar{U}^*)^\top, (\bar{U}_q^*)^\top)^\top$ در شرایط

کروش-کان-تاکر زیر صدق کند:

$$\begin{cases} \bar{U}^* \geq 0, & \bar{U}_q^* \geq 0, & q = 1, 2, 3 \\ \bar{G}\bar{A}^* \leq 0, & \bar{g}_q(\bar{A}^*) \leq 0, & q = 1, 2, 3 \\ (\bar{U}^*)^\top \bar{G}\bar{A}^* = 0, & (\bar{U}_q^*)^\top \bar{g}_q(\bar{A}^*) = 0, & q = 1, 2, 3 \\ 2\bar{Q}\bar{A}^* + \bar{B} + \bar{G}^\top \bar{U}^* + \bar{g}_q^\top(\bar{A}^*)\bar{U}_q^* = 0, & q = 1, 2, 3 \end{cases} \quad (40.4)$$

که در آن $\bar{A}^*$ را نقطه کروش-کان-تاکر مدل (۱۴.۴)–(۱۶.۴) و $\bar{U}^*$ و $\bar{U}_q^*$ بردارهای ضریب لاگرانژ متناظر با $\bar{A}^*$ می‌باشند. و با فرض این که $A(\cdot)$ ، $U(\cdot)$ و $U_q(\cdot)$ برای $q = 1, 2, 3$ ، متغیرهای وابسته به زمان باشند، آن‌گاه شبکه عصبی زیر را معرفی می‌کنیم:

$$\frac{dy}{dt} = \tau \phi(y) \quad (41.4)$$

$$y(t_0) = y_0 = (\bar{A}(t_0)^\top, \bar{U}(t_0)^\top, \bar{U}_q(t_0)^\top)^\top \in \mathbb{R}^{11p+11}, \tau > 0 \quad (42.4)$$

که در آن برای $q = 1, 2, 3$ داریم:

$$\phi(y) = \begin{pmatrix} - \left(2\bar{Q}\bar{A} + \bar{B} + \bar{G}^\top (\bar{U} + \bar{G}\bar{A})^+ + \nabla \bar{g}_q(\bar{A})^\top (\bar{U}_q + \bar{g}_q(\bar{A}))^+ \right) \\ (\bar{U} + \bar{G}\bar{A})^+ - \bar{U} \\ (\bar{U}_q + \bar{g}_q(\bar{A}))^+ - \bar{U}_q \end{pmatrix} \quad (43.4)$$

برای سادگی فرض می‌کنیم که $\tau = 1$.

۱.۴.۴ بررسی همگرایی و پایداری

در این قسمت به بررسی گزاره‌های مربوط به همگرایی و پایداری مدل (۱۴.۴)–(۱۶.۴) می‌پردازیم.

قضیه ۱.۴.۴. فرض کنید $y^* = (\bar{A}^{*\top}, \bar{U}^{*\top}, \bar{U}_q^{*\top})^\top$ نقطه تعادل مدل شبکه عصبی (۴۱.۴) و (۴۲.۴) باشد. بنابراین A^* یک نقطه کروش-کان-تاکر برای مسئله (۱۴.۴)–(۱۶.۴) است. از سمت دیگر، اگر $A^* \in \mathbb{R}^{p+4}$ یک جواب مسئله (۱۴.۴)–(۱۶.۴) باشد، بنابراین $U^* \in \mathbb{R}^{p+4}$ و $U_q^* \in \mathbb{R}^3$ وجود دارند به طوری که $y^* = (A^{*\top}, U^{*\top}, U_q^{*\top})^\top$ نقطه تعادل مدل شبکه عصبی (۴۱.۴) و (۴۲.۴) است.

در ادامه اثبات می‌کنیم که شبکه عصبی (۴۱.۴) و (۴۲.۴) پایدار لیاپانوف است.

لم ۱.۴.۴. ژاکوبین ماتریس ϕ که در (۴۳.۴) تعریف شده است، یک ماتریس نیمه معین منفی است.

قضیه ۲.۴.۴. تحت مفروضات لم ۱.۴.۴ شبکه عصبی (۴۱.۴) و (۴۲.۴) پایدار لیاپانوف است.

برای نشان دادن همگرایی سراسری مدل، لم و قضیه زیر را داریم:

لم ۲.۴.۴. (i) برای هر نقطه ابتدایی $y(t_0) = (A_0(t_0)^T, U(t_0)^T, U_q(t_0)^T)^T$ یک جواب پیوسته یکتا $y(t) = (A(t)^T, U(t)^T, U_q(t)^T)^T$ برای سیستم (۴۱.۴) و (۴۲.۴) وجود دارد. (ii) فرض کنید $y(t) = (A(t)^T, U(t)^T, U_q(t)^T)^T$ یک مسیر از شبکه عصبی (۴۱.۴) و (۴۲.۴) با نقطه ابتدایی $y(t_0) = (A_0(t_0)^T, U(t_0)^T, U_q(t_0)^T)^T$ باشد. در این صورت اگر، $U(t_0) \geq 0$ و $U_q(t_0) \geq 0$ بنابراین $U(t) \geq 0$ و $U_q(t) \geq 0$

قضیه ۳.۴.۴. مسیرهای شبکه عصبی (۴۱.۴) و (۴۲.۴) برای هر نقطه ابتدایی

$$y_0 = (A(t_0)^T, U(t_0)^T, U_q(t_0)^T)^T \in \mathbb{R}^{11p+11}$$

به نقطه تعادل شبکه عصبی همگرا می‌شود. به طور خاص، شبکه عصبی (۴۱.۴) و (۴۲.۴) با هر نقطه ابتدایی $y_0 = (A(t_0)^T, U(t_0)^T, U_q(t_0)^T)^T \in \mathbb{R}^{11p+11}$ پایدار مجانبی سراسری است و D^* یک نقطه تعادل یکتا دارد. که در آن D^* نشان دهنده مجموعه نقاط بهینه مسئله اصلی و دوگان است.

قضیه ۴.۴.۴. سرعت همگرایی شبکه عصبی (۴۱.۴) و (۴۲.۴) با افزایش τ افزایش می‌یابد.

۵.۴ مثال‌های عددی

برای نشان دادن عملکرد روش پیشنهادی، در این بخش چند مثال کاربردی را توسط شبکه عصبی (۴۱.۴) و (۴۲.۴) آزمایش می‌کنیم. در این قسمت برای نشان دادن برتری روش پیشنهادی به حل یک مثال شبیه سازی شده و یک مثال شناخته شده می‌پردازیم. در این مثال‌ها، درستی روش پیشنهادی را با چندین نقطه شروع مختلف و فرض $w_1 = w_2 = w_3$ ، آزمایش می‌کنیم. همچنین برای تأکید بیشتر بر مدل پیشنهادی، مقایسه‌هایی با مدل‌های موجود در [۴۱، ۸۹، ۱۰۵، ۱۰۶] به وسیله فاصله‌های فازی مختلف انجام شده است.

مثال ۱.۵.۴. (مثال شبیه‌سازی شده) برای نشان دادن عملکرد روش پیشنهادی، مجموعه 3° داده شبیه سازی شده با اندازه $n = 2^\circ$ با خروجی فازی و ورودی غیر فازی را تولید می‌کنیم. مدل رگرسیون خطی زیر را در نظر بگیرید:

$$Y_i = A_0 + A_1 x_{i1} + \dots + A_{25} x_{i25} + \epsilon_i \quad (44.4)$$

که در آن برای $i = 1, 2, \dots, 2^\circ$

$$x_{ij} = \begin{cases} z_1 + e_1 & \text{برای } j = 1, 2, \dots, 5 \\ z_2 + e_2 & \text{برای } j = 6, 7, \dots, 10 \\ z_3 + e_3 & \text{برای } j = 11, 12, \dots, 15 \\ z_4 & \text{برای } j = 16, 17, \dots, 25 \end{cases}$$

$$A_j = \begin{cases} (°, °, °)_T & \text{برای } j = ° \\ (۱/۵, °/۵, °/۳۵)_T & \text{برای } j = ۱, ۲, \dots, ۵ \\ (°, °/۵, °/۳۵)_T & \text{برای } j = ۶, ۷, \dots, ۱۰ \\ (°, °/۳۵, °/۵)_T & \text{برای } j = ۱۱, ۱۲, \dots, ۱۵ \\ (°, °, °/۵)_T & \text{برای } j = ۱۶, ۱۷, \dots, ۲۰ \\ (°, °, °)_T & \text{برای } j = ۲۱, ۲۲, \dots, ۲۵ \end{cases}$$

$$\epsilon_i = (\epsilon_{im}, \epsilon_{il}, \epsilon_{iu})_T, \quad \epsilon_{il} = \epsilon_{iu} = |\epsilon_{im}|$$

با

$$z_1 \sim N(°, °/۵), \quad z_2 \sim N(°, °/۸), \quad z_3 \sim N(°, ۱/۲), \quad z_4 = N_2(°, \Sigma)$$

$$e_1 \sim N(°, °/°۱), \quad e_2 \sim N(°, °/°۲), \quad e_3 \sim N(°, °/°۳)$$

$$\epsilon_{im} = N(°, ۱/۵)$$

$$[\Sigma]_{lk} = Cov(x_{jl}, x_{jk}) = °/۵^{|l-k|}$$

جدول ۱.۴ میانگین خطاها برای هر ۳۰ مجموعه داده شبیه سازی شده را با استفاده از روش پیشنهادی نشان می دهد.

جدول ۱.۴: میانگین خطای ۳۰ مجموعه داده شبیه سازی شده برای مقادیر مختلف γ در مقایسه با برخی تکنیک های متداول برای حل مدل رگرسیون فازی در مثال ۱.۵.۴.

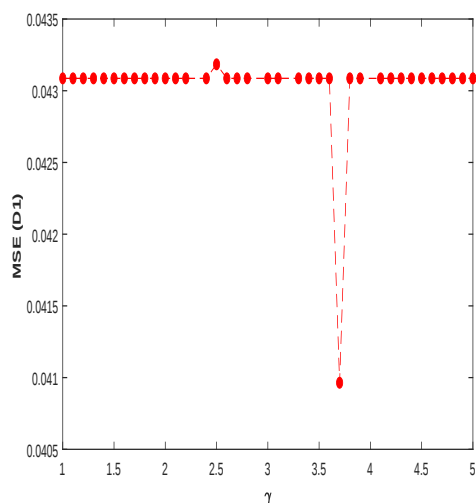
$\overline{MCV}$			$\overline{MSPE}$			$\overline{MSM}$	روش ها
D_2	D_4	D_7	D_2	D_4	D_7		
۳/۱۸۴	۳/۰۴۶	۱/۳۴۹	۰/۷۳۴	۰/۱۴۶	۰/۰۳۹	۰/۹۲۰	بریج فازی با $(\gamma = ۱)$
۳/۱۸۴	۳/۱۴۸	۱/۳۱۸	۰/۷۸۳	۰/۱۴۴	۰/۰۳۹	۰/۹۲۲	بریج فازی با $(\gamma = ۱/۵)$
۳/۱۸۹	۳/۱۳۶	۱/۳۲۹	۰/۸۴۵	۰/۱۵۶	۰/۰۳۷	۰/۹۱۱	بریج فازی با $(\gamma = ۲)$
۳/۱۸۴	۳/۱۴۸	۱/۳۱۸	۰/۷۸۲	۰/۱۵۳	۰/۰۳۹	۰/۹۲۲	بریج فازی با $(\gamma = ۳)$
۳/۹۷۶	۳/۳۴۴	۱/۵۲۲	۱/۲۲۵	۳/۰۰۹	۰/۲۰۴	۰/۸۰۵	مدل ارائه شده در [۸۹]
۵/۰۵۹	۵/۹۱۲	۳/۶۶۱	۲/۱۳۰	۵/۲۳۳	۰/۵۰۴	۰/۶۰۷	مدل ارائه شده در [۴۱]
۴/۳۷۰	۵/۷۳۴	۲/۳۲۵	۲/۳۶۳	۵/۹۹۷	۰/۵۹۲	۰/۶۵۴	مدل ارائه شده در [۱۰۵]
۴/۴۴۶	۴/۷۴۰	۲/۹۲۲	۱/۷۶۰	۴/۷۶۸	۰/۴۸۴	۰/۷۲۱	مدل ارائه شده در [۱۰۷]

اکنون روش شبکه عصبی پیشنهادی را برای مجموعه داده های شبیه سازی شده ۱۵ ام استفاده می کنیم. می دانیم که انتخاب بهینه γ باعث افزایش کارایی مدل می شود. بنابراین، برای تخمین بهترین مقدار Y با استفاده از $(۱۴.۴) - (۱۶.۴)$ ، باید پارامتر بریج بهینه γ را پیدا کنیم. جدول ۲.۴ مقادیر متناظر $MSPE$ و MSM را برای مدل رگرسیون بریج فازی پیشنهادی با مقادیر مختلف γ نشان می دهد.

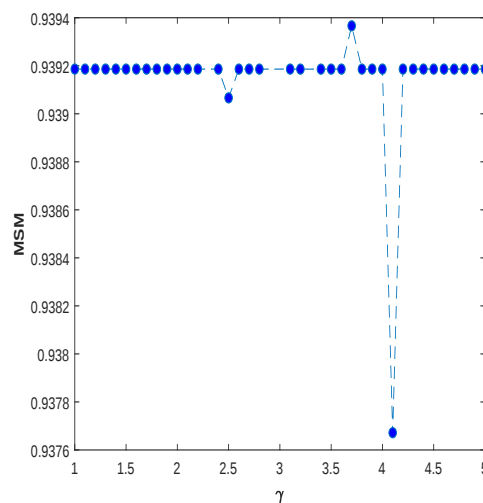
واضح است که خطای کل مدل رگرسیون بریج فازی پیشنهادی با $\gamma = ۳/۷$ ، از خطای کل محاسبه شده برای $\gamma = ۱, ۲, ۳$ و بهتر است (به شکل ۱.۴ مراجعه شود). علاوه بر این،

جدول ۲.۴: خطای ۱۵ امین مجموعه داده شبیه‌سازی شده برای مقادیر مختلف γ در مثال ۱.۵.۴.

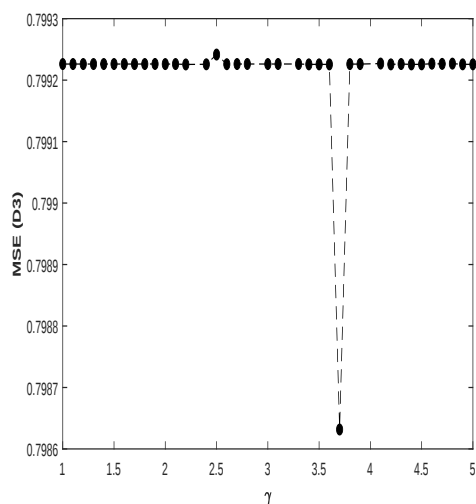
MCV			MSPE			MSM	برج فازی
D_2	D_4	D_3	D_2	D_4	D_{13}		γ
۳/۱۸۹۷۰	۳/۱۱۱۷۵	۱/۳۴۹۰۱	۰/۷۹۹۲۱	۰/۱۲۹۲۹	۰/۰۴۳۱۰	۰/۹۳۹۲۸	۱
۳/۱۸۹۷۴	۳/۱۳۷۴۸	۱/۳۲۹۱۲	۰/۷۹۹۲۱	۰/۱۲۹۵۲	۰/۰۴۳۱۸	۰/۹۳۸۲۱	۲
۳/۱۸۸۹۲	۳/۱۱۱۰۰	۱/۳۱۸۸۲	۰/۷۹۹۲۱	۰/۱۲۹۵۲	۰/۰۴۳۱۱	۰/۹۳۹۲۱	۳
۳/۱۸۸۹۲	۳/۱۱۰۸۵	۱/۳۱۸۵۵	۰/۷۹۸۶۲	۰/۱۲۶۷۹	۰/۰۴۲۱۵	۰/۹۳۹۴۰	۳/۷



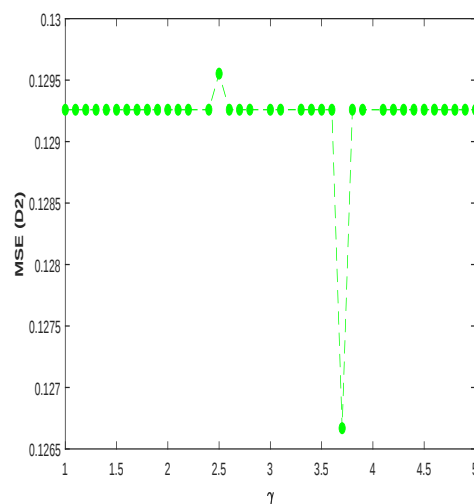
(ب) $MSPE$ با فاصله D_3



(ا) MSM



(د) $MSPE$ با فاصله D_2



(ج) $MSPE$ با فاصله D_4

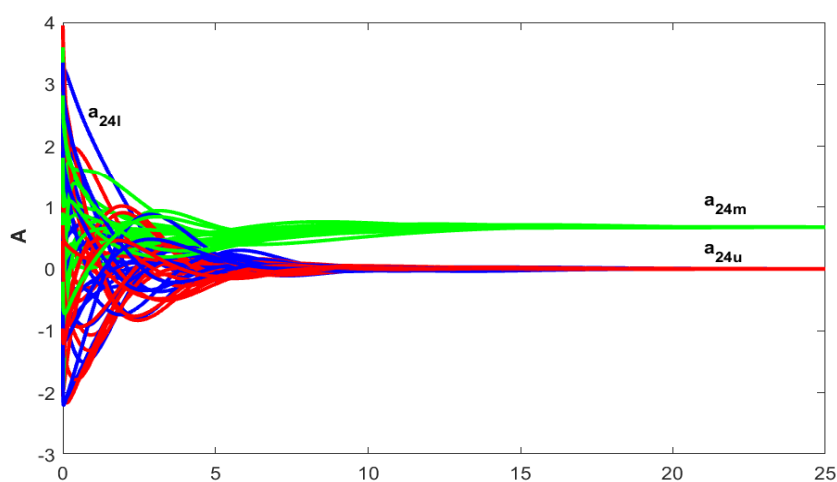
شکل ۱.۴: نیکویی برازش برای مقادیر مختلف γ با سه فاصله متفاوت D_2 ، D_4 و D_3 در مثال ۱.۵.۴.

جدول ۳.۴ ضرایب فازی برآورد شده مدل رگرسیون برج فازی را برای $\gamma = ۳/۷$ ، $t_1 = ۱۶۶۹$ ، $t_2 = ۴۱$ و $t_3 = ۳۸$ نشان می‌دهد.

شکل‌های ۲.۴-۷.۴ مسیرهای همگرایی مدل شبکه عصبی را نمایش می‌دهد. برای تجسم چگونگی تناسب مدل‌ها با داده‌های فازی مشاهده شده، از نمودارهای حبابی

جدول ۳.۴: ضرایب مدل رگرسیون بریج با $\gamma = 3/7$ ، مطابق با ۱۵ امین مجموعه داده در مثال ۱.۵.۴.

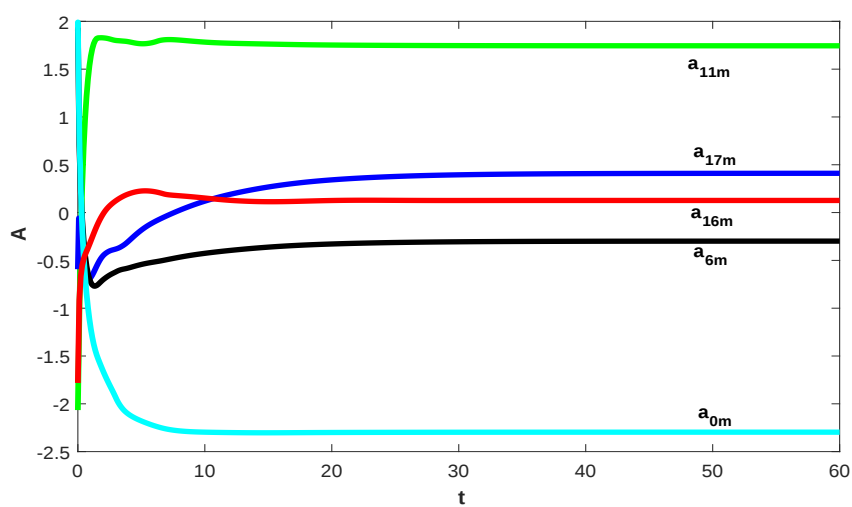
ضرایب فازی	بریج فازی با $(\gamma = 3/7)$	ضرایب فازی	بریج فازی با $(\gamma = 3/7)$
A_0	$(-2/2756, 0, 0)_T$	A_{13}	$(-2/2556, 0/6890, 1/0281)_T$
A_1	$(-2/8291, 0/6433, 0/9030)_T$	A_{14}	$(1/9892, 0, 0/1928)_T$
A_2	$(3/5032, 0, 0)_T$	A_{15}	$(4/2537, 0, 0)_T$
A_3	$(2/2692, 0, 0)_T$	A_{16}	$(0/2557, 0/0792, 0/4831)_T$
A_4	$(-1/1800, 0, 0)_T$	A_{17}	$(0/4384, 0, 0)_T$
A_5	$(0/2767, 0, 0)_T$	A_{18}	$(0/7506, 0, 0)_T$
A_6	$(-0/3055, 0/3847, 0/2925)_T$	A_{19}	$(0/8426, 0, 0/4403)_T$
A_7	$(0/1711, 0/2676, 0/2585)_T$	A_{20}	$(2/2863, 0, 0/4810)_T$
A_8	$(-1/4283, 0/9777, 0/7293)_T$	A_{21}	$(0/0551, 0/0341, 0/1301)_T$
A_9	$(0/1526, 0/6975, 0/3733)_T$	A_{22}	$(-0/4957, 0/8162, 0/7252)_T$
A_{10}	$(0/2140, 2/4168, 2/6125)_T$	A_{23}	$(-2/9226, 0, 0)_T$
A_{11}	$(1/4807, 0/3650, 0/3956)_T$	A_{24}	$(0/6620, 0, 0)_T$
A_{12}	$(-0/5725, 0/8036, 0/9055)_T$	A_{25}	$(0/0260, 0/0953, 0/5637)_T$



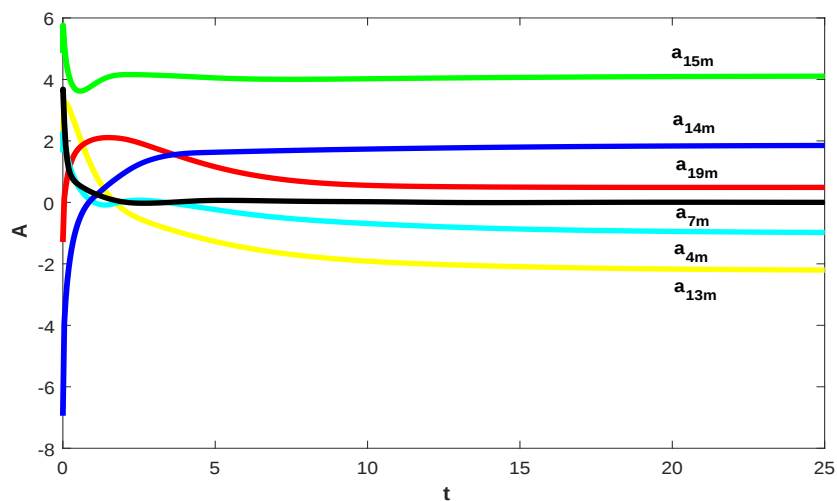
شکل ۲.۴: رفتار مسیر A_{24} در مدل رگرسیون بریج با $(\gamma = 3/7)$ برای شبکه عصبی (۴۱.۴) و (۴۲.۴) با ۲۰ نقطه شروع متفاوت در مثال ۱.۵.۴.

فازی استفاده می شود (شکل ۸.۴ مشاهده شود). نمودارهای حبابی فازی نشان می دهد که مشاهدات به دلیل نزدیک بودن دایره ها به نیمساز صفحه و شعاع کوتاه هر دایره، به خوبی با مدل های پیشنهادی تناسب دارد.

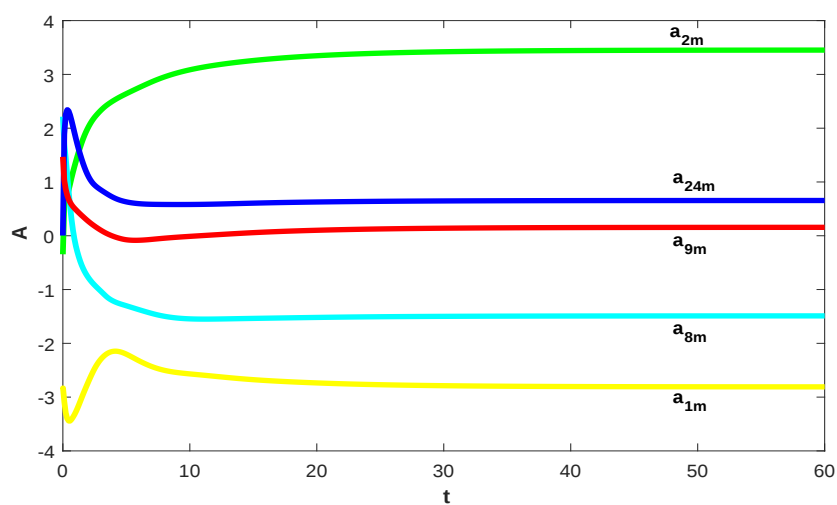
عملکرد مدل در حالت گرافیکی در نمودارهای ۹.۴ – ۱۱.۴، از طریق نمودارهای تیلور نشان داده شده است. شکل ۹.۴ شبکه عصبی پیشنهادی (۱۹.۲) و (۲۰.۲) را برای مدل رگرسیون بریج مراکز با $\gamma = 1$ ، $\gamma = 2$ ، $\gamma = 3$ ، و $\gamma = 3/7$ به ترتیب با دایره قهوه ای، سبز، سیاه و سرخابی نشان داده شده است. واضح است که مدل پیشنهادی با $\gamma = 3/7$ بسیار نزدیک به مجموعه داده واقعی است. این نتیجه برای پهنای چپ و راست نیز برقرار است. بنابراین، نتیجه گرفته می شود که مدل پیشنهادی مورد استفاده در این مقاله در عمل دارای شایستگی های خاصی است.



(ا)

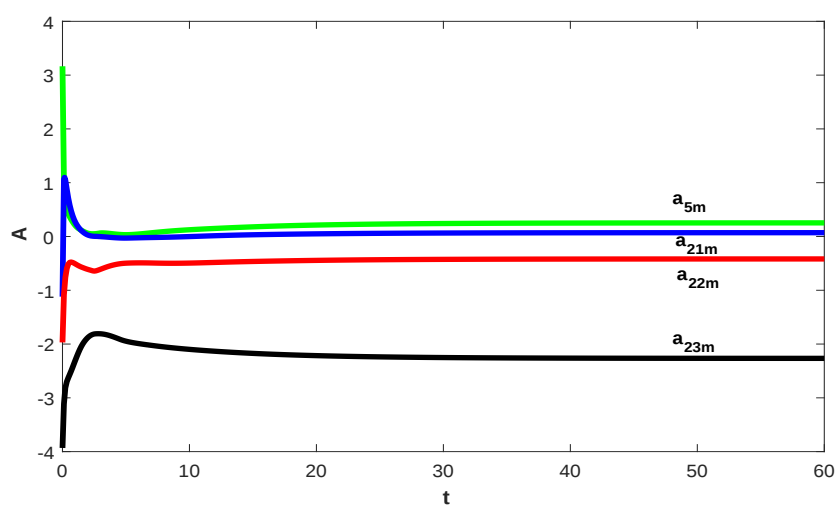


(ب)

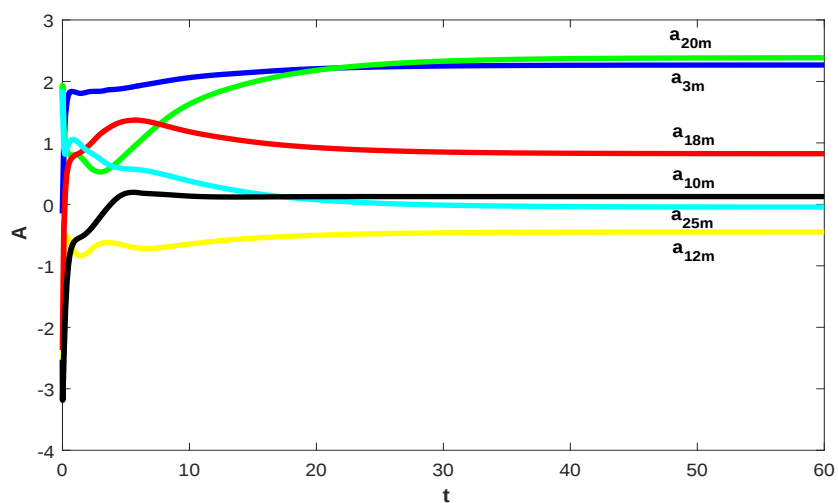


(ج)

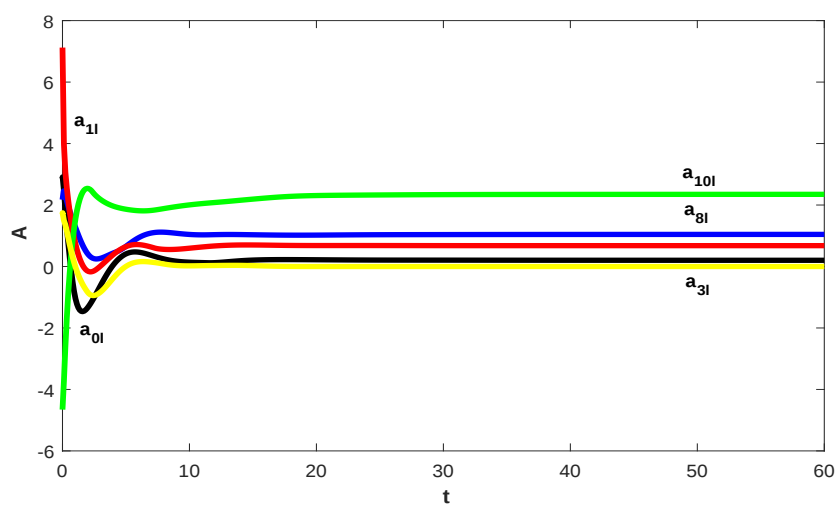
شکل ۳.۴: ادامه رفتار مسیرهای شبکه عصبی (۴۱.۴) و (۴۲.۴) در مثال ۱.۵.۴.



(ا)

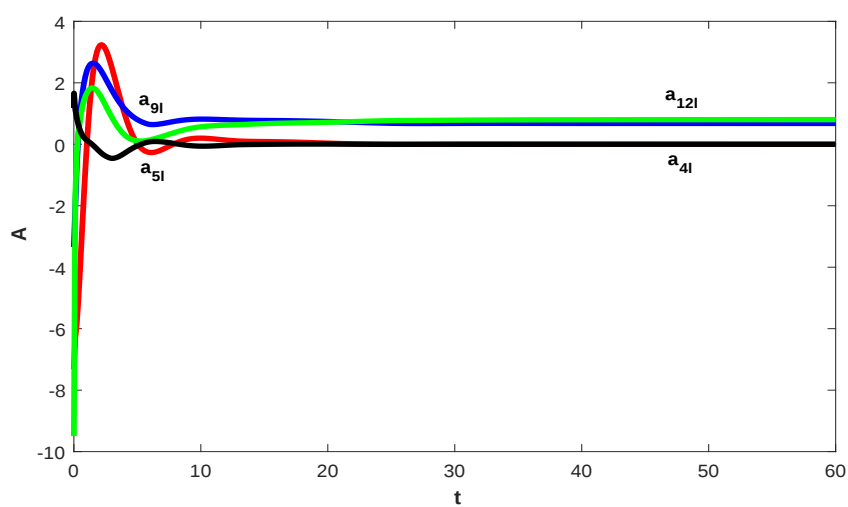


(ب)

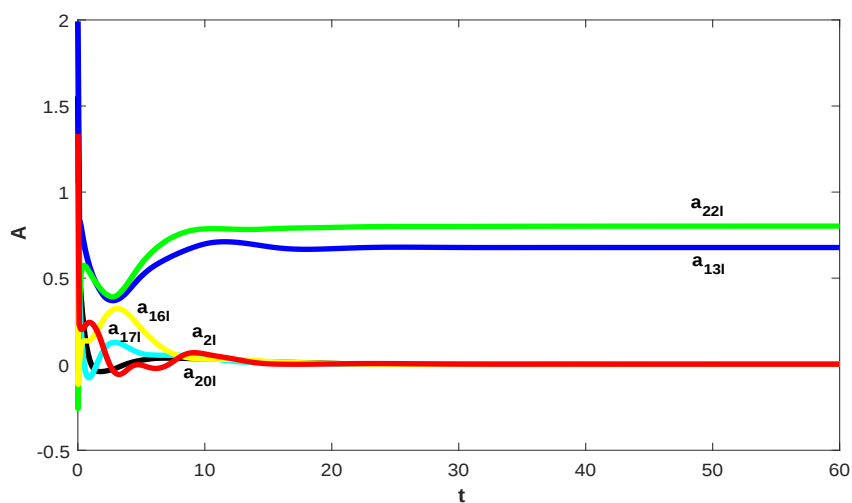


(ج)

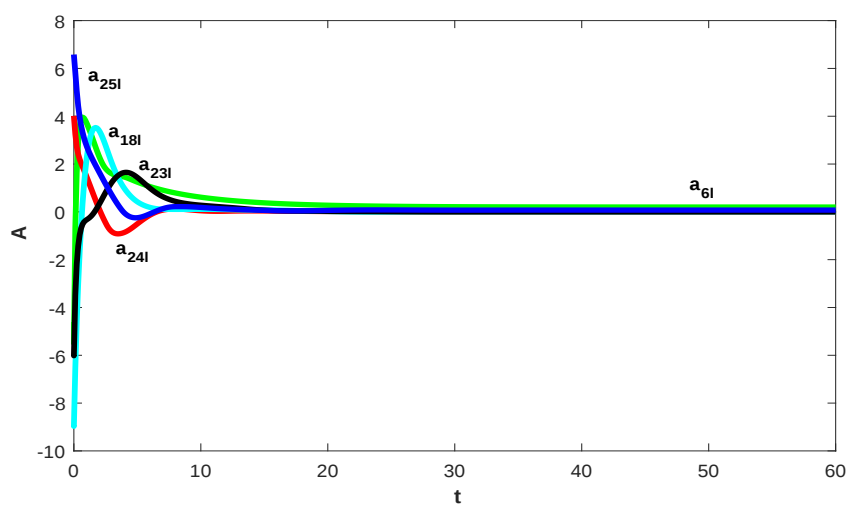
شکل ۴.۴: ادامه رفتار مسیرهای شبکه عصبی (۴۱.۴) و (۴۲.۴) در مثال ۱.۵.۴.



(ا)

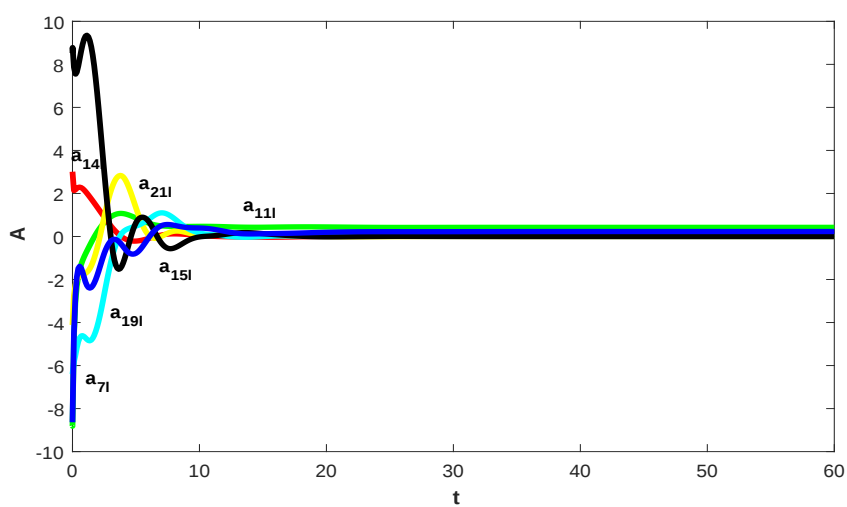


(ب)

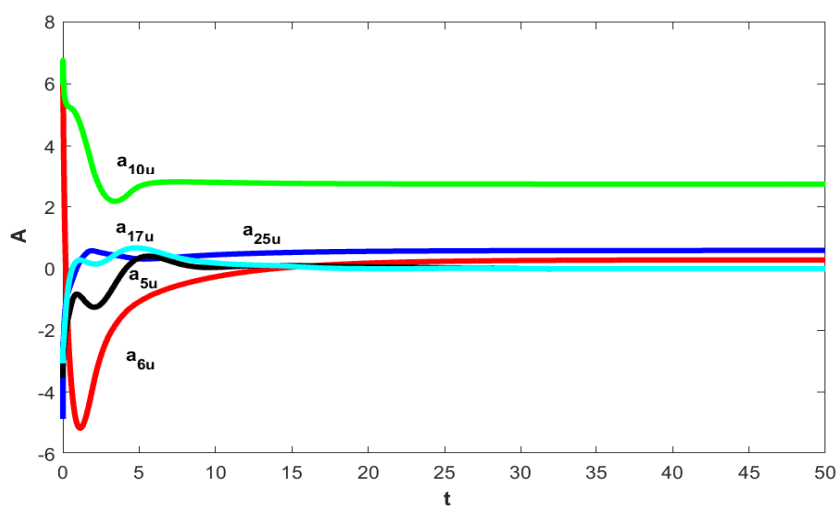


(ج)

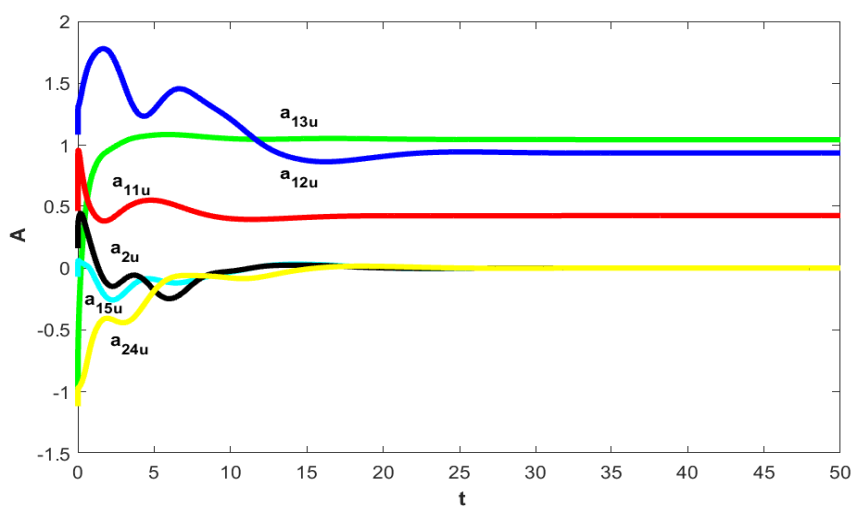
شکل ۵.۴: ادامه رفتار مسیرهای شبکه عصبی (۴۱.۴) و (۴۲.۴) در مثال ۱.۵.۴.



(I)

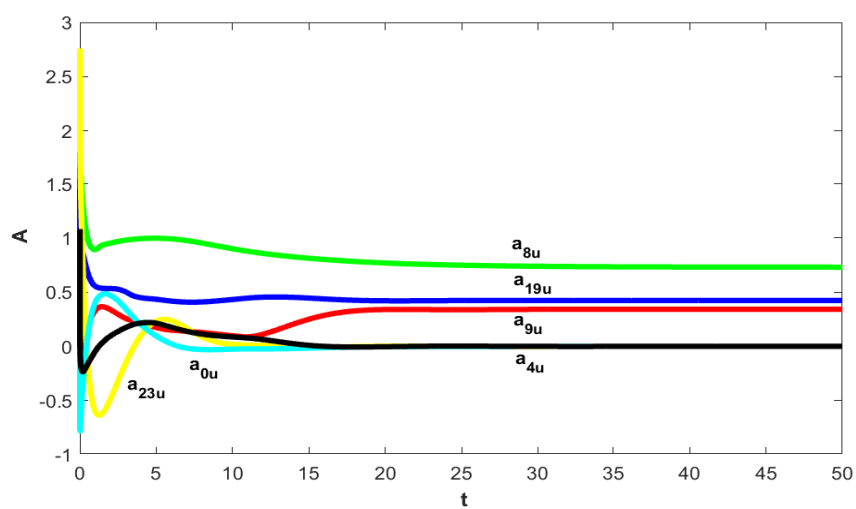


(ب)

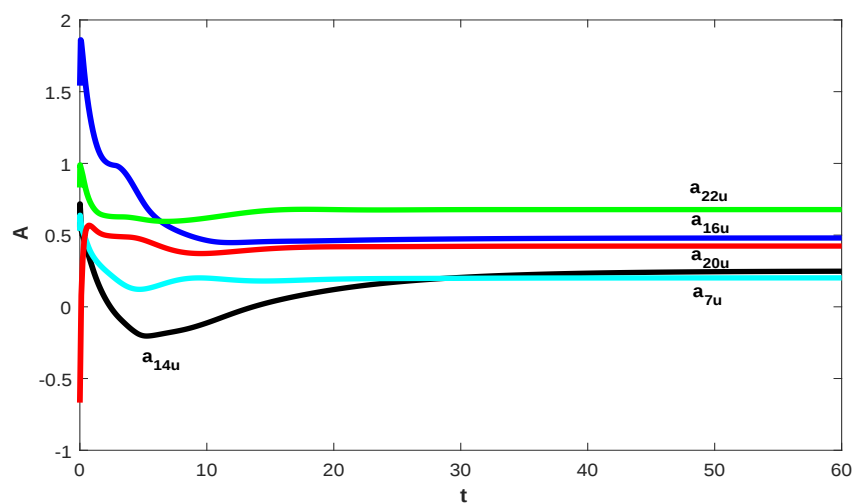


(ج)

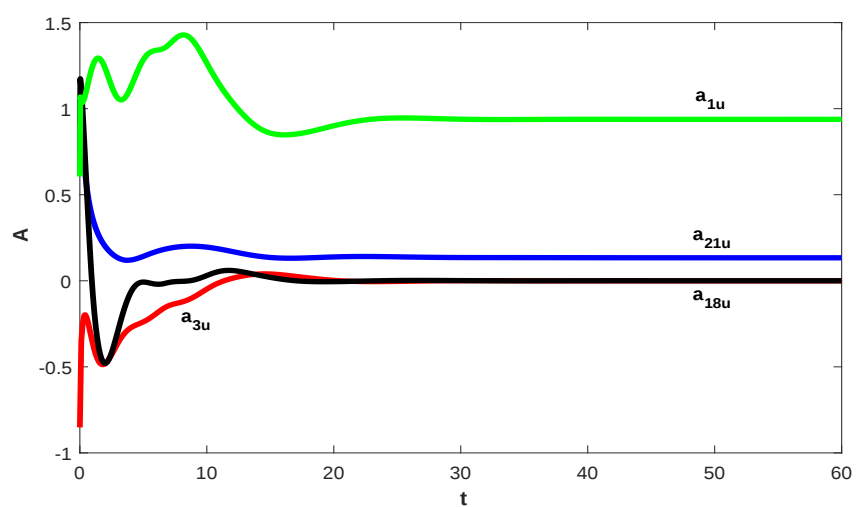
شکل ۶.۴: ادامه رفتار مسیرهای شبکه عصبی (۴۱.۴) و (۴۲.۴) در مثال ۱.۵.۴.



(ا)

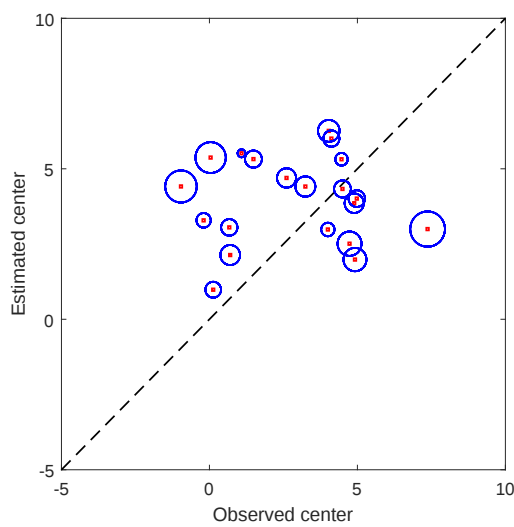


(ب)

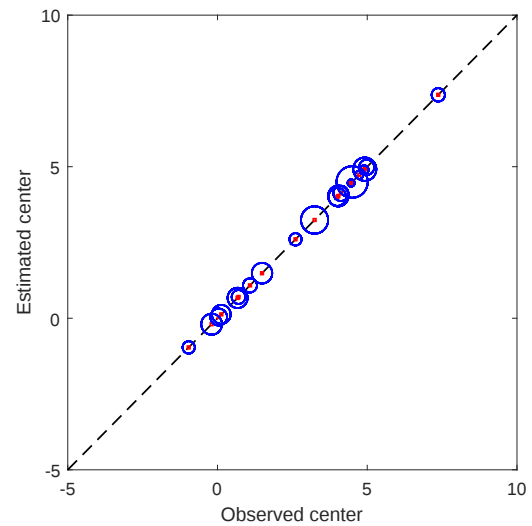


(ج)

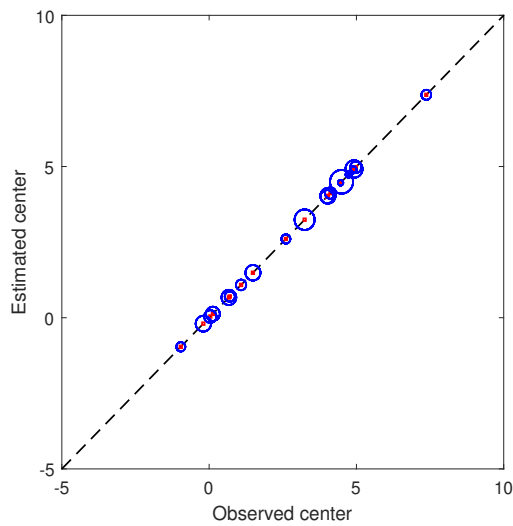
شکل ۷.۴: ادامه رفتار مسیرهای شبکه عصبی (۴۱.۴) و (۴۲.۴) در مثال ۱.۵.۴.



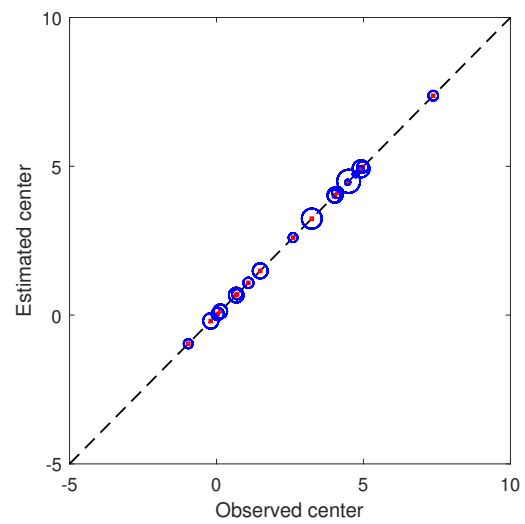
(ب) مدل رگرسیون بریج فازی پیشنهادی با $\gamma = 2$



(ا) مدل رگرسیون بریج فازی پیشنهادی با $\gamma = 1$

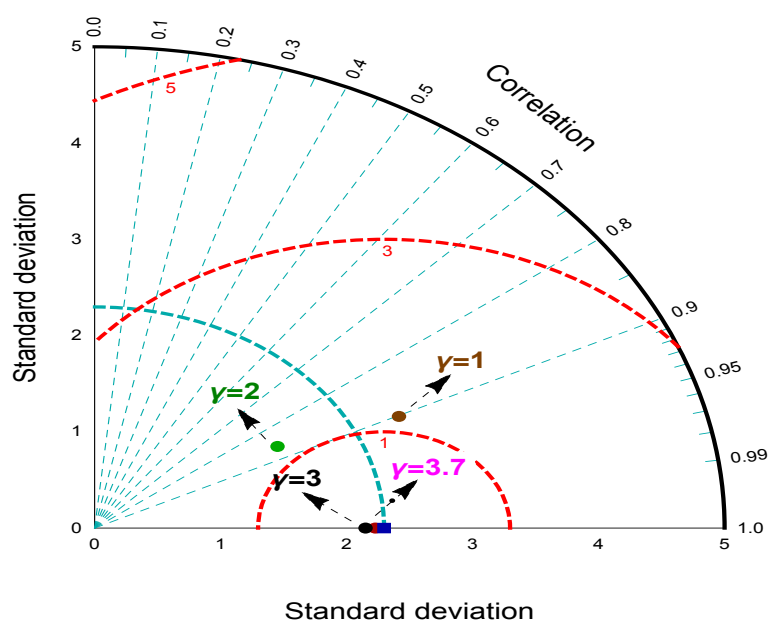


(د) مدل رگرسیون بریج فازی پیشنهادی با $\gamma = 3/7$

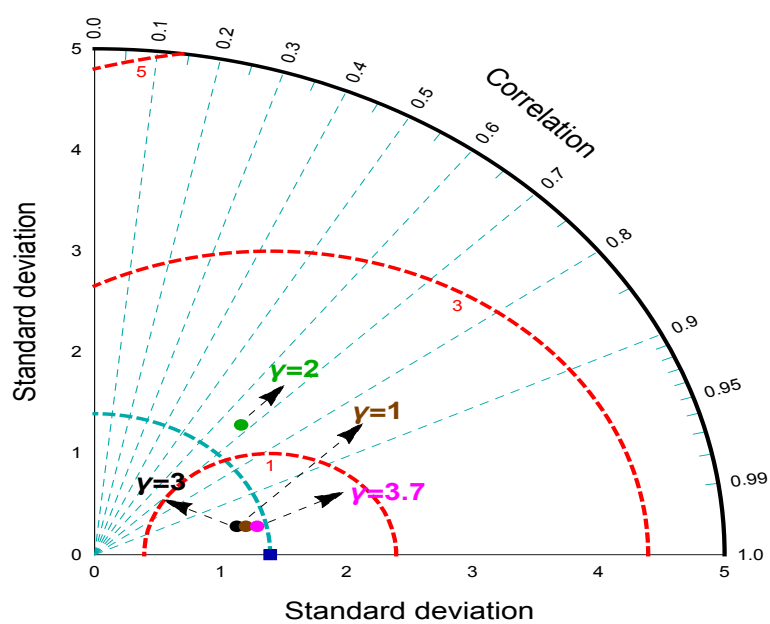


(ج) مدل رگرسیون بریج فازی پیشنهادی با $\gamma = 3$

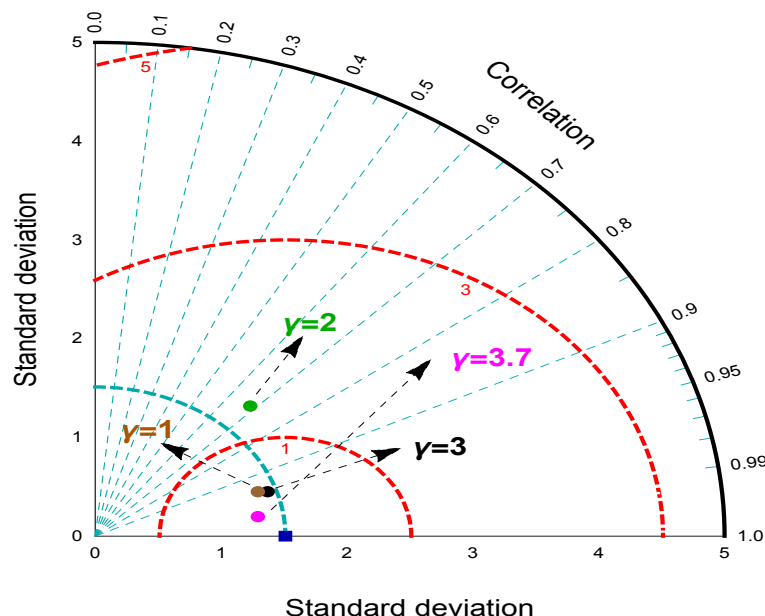
شکل ۸.۴: نمودار حبابی فازی در مثال ۱.۵.۴.



شکل ۹.۴: نمودار تیلور برای مراکز اعداد فازی در مثال ۱.۵.۴.



شکل ۱۰.۴: نمودار تیلور برای مقادیر سمت چپ اعداد فازی ۱.۵.۴.



شکل ۱۱.۴: نمودار تیلور برای مقادیر سمت راست اعداد فازی ۱.۵.۴.

مثال ۲.۵.۴. در این مثال، مجموعه داده‌های نشان داده شده در جدول ۴.۴ از [۱۰۸] گرفته شده است. با استفاده از این داده‌ها در نظر داریم ضرایب مدل رگرسیون زیر را به دست آوریم:

جدول ۴.۴: مقادیر مشاهده شده و خطای به دست آمده در مثال ۲.۵.۴.

شماره	x_1	x_2	x_3	x_4	x_5	Y
۱	۱	۵	۳۶/۴۳	۳۸/۰۹	۱	$(۶۰۶۰, ۵۵۰, ۵۵۰)_T$
۲	۱	۶	۲۶/۵۰	۶۲/۱۰	۱	$(۷۱۰۰, ۵۰, ۵۰)_T$
۳	۱	۷	۴۴/۷۱	۶۳/۷۰	۱	$(۸۰۸۰, ۴۰۰, ۴۰۰)_T$
۴	۱	۸	۳۸/۰۹	۷۴/۵۲	۱	$(۸۲۶۰, ۱۵۰, ۱۵۰)_T$
۵	۱	۷	۴۱/۴۰	۷۵/۳۸	۲	$(۸۶۵۰, ۷۵۰, ۷۵۰)_T$
۶	۲	۴	۲۶/۴۹	۵۲/۹۹	۲	$(۸۵۲۰, ۴۵۰, ۴۵۰)_T$
۷	۲	۵	۲۶/۴۹	۶۲/۹۳	۲	$(۹۱۷۰, ۷۰۰, ۷۰۰)_T$
۸	۲	۶	۳۳/۱۲	۷۲/۰۴	۳	$(۱۰۳۱۰, ۲۰۰, ۲۰۰)_T$
۹	۲	۶	۴۳/۰۶	۷۶/۱۲	۲	$(۱۰۹۲۰, ۶۰۰, ۶۰۰)_T$
۱۰	۲	۷	۴۲/۶۴	۹۰/۲۶	۲	$(۱۲۰۳۰, ۱۰۰, ۱۰۰)_T$
۱۱	۳	۷	۳۱/۳۳	۸۵/۷۰	۳	$(۱۳۹۴۰, ۳۵۰, ۳۵۰)_T$
۱۲	۳	۶	۲۷/۶۴	۹۵/۲۷	۳	$(۱۴۲۰۰, ۲۵۰, ۲۵۰)_T$
۱۳	۳	۶	۲۷/۶۴	۱۰۵/۹۸	۳	$(۱۶۰۱۰, ۳۰۰, ۳۰۰)_T$
۱۴	۳	۶	۶۶/۸۱	۷۹/۲۵	۳	$(۱۶۳۲۰, ۵۰۰, ۵۰۰)_T$
۱۵	۳	۶	۳۲/۲۵	۱۲۰/۵	۳	$(۱۶۹۹۰, ۶۵۰, ۶۵۰)_T$

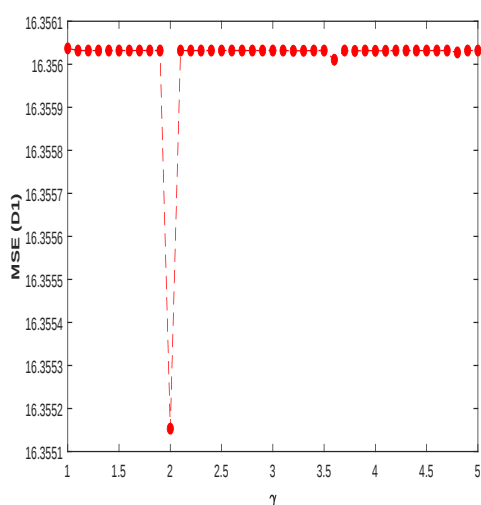
$$Y = A_0 + A_1x_1 + A_2x_2 + A_3x_3 + A_4x_4 + A_5x_5 \quad (۴۵.۴)$$

برای برآورد مقدار Y با استفاده از (۱۴.۴) – (۱۶.۴) لازم است که مقدار بهینه پارامتر بریج γ را پیدا کنیم.

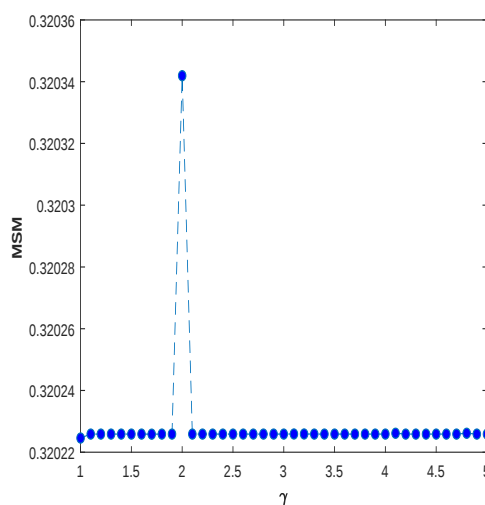
شکل ۱۲.۴، با در نظر گرفتن دو مدل نیکویی برازش و با فواصل D_3 ، D_4 و D_5 افزایش چشمگیر شاخص MSM را در $\gamma = 2$ نشان می دهد. از طرف دیگر، شاخص $MSPE$ با سه

فاصله متفاوت D_3 ، D_4 و D_2 روند متفاوتی را نشان می‌دهد. مقادیر به میزان قابل توجهی در $\gamma = 2$ کاهش می‌یابد. بنابراین مسئله رگرسیون بریج فازی برای $\gamma = 2$ ، $r_1 = 10^8 \times 1139$ و $r_2 = 10^6 \times 3893$ به صورت زیر مدل‌بندی می‌شود:

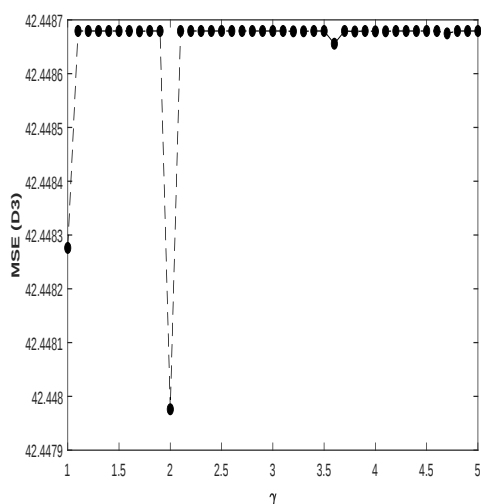
$$\begin{aligned} \hat{Y} = & (-1831/984, 201/7570)_T + (2624/075, 0/7776 \times 10^{-5})_{Tx_1} \\ & + (82/16977, 0/4976 \times 10^{-7})_{Tx_2} + (72/51824, 3/6588)_{Tx_3} \\ & + (-45/344, 0)_{Tx_4} + (-39/2779, 70/6566)_{Tx_5} \end{aligned} \quad (46.4)$$



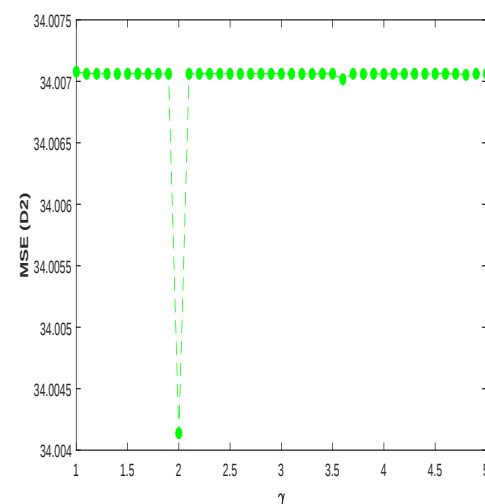
(ب) MSPE با فاصله D_1



(آ) MSM



(د) MSPE با فاصله D_3



(ج) MSPE با فاصله D_2

شکل ۱۲.۴: نیکویی برازش برای مقادیر مختلف γ با سه فاصله متفاوت D_1 ، D_2 و D_3 در مثال ۲.۵.۴.

شکل ۱۳.۴ نشان می‌دهد که مسیرهای مدل (۴۱.۴) و (۴۲.۴) برای حل مدل رگرسیون بریج فازی با $\gamma = 2$ همگرا به جواب بهینه مسئله است. همچنین شکل ۱۶.۴ رفتار مسیر A_2

از مدل رگرسیون بریج فازی را با ۲۰ نقطه اولیه متفاوت نشان می‌دهد. به علاوه، مقایسه میان مدل پیشنهادی و چندین مدل رگرسیون فازی در جدول ۶.۲ ارائه شده است.

جدول ۵.۴: مقایسه خطای به دست آمده از مدل رگرسیون بریج برای $\gamma = 2$ با برخی تکنیک‌های متداول برای حل مدل رگرسیون فازی در مثال ۲.۵.۴.

MCV			MSPE			روش‌ها	MSM
D_2	D_4	D_3	D_2	D_4	D_3		
۵۹/۸۴۱۴	۵۸/۰۴۸۸	۲۷/۱۷۸۸	۴۲/۴۴۷۹	۳۴/۰۰۴۲	۱۶/۳۵۵۱	بریج فازی با $(\gamma = 2)$	۰/۳۲۰۳
۷۴/۰۸۲۷	۱۰۲/۲۰۷۷	۴۶/۶۴۵۲	۱۰۸/۸۶۱۷	۱۴۳/۰۷۹۹	۶۷/۹۴۳۴	مدل ارائه شده در [۸۹]	۰/۱۴۳۹
۷۵/۷۵۱۶	۷۲/۵۷۴۱	۲۹/۴۹۹۴	۷۴/۲۶۰۶	۷۱/۱۳۴۳	۳۳/۷۷۹۱	مدل ارائه شده در [۱۰۵]	۰/۲۳۵۵
۷۵/۵۳۲۵	۱۱۵/۰۰۹۱	۵۳/۲۸۴۲	۱۱۸/۹۴۳۹	۱۰۰/۷۸۷۲	۴۴/۲۷۹۷	مدل ارائه شده در [۴۱]	۰/۱۳۴۴
۷۶/۱۹۱۰	۸۶/۸۲۹۹	۳۸/۲۹۱۱	۶۵/۲۱۲۶	۶۹/۹۹۰۹	۳۳/۰۸۲۰	مدل ارائه شده در [۱۰۹]	۰/۲۰۸۳
۶۴/۳۰۵۵	۸۷/۶۶۳۴	۳۹/۹۲۹۹	۸۶/۶۲۷۶	۹۶/۴۸۷۹	۴۴/۳۱۴۹	مدل ارائه شده در [۱۰۷]	۰/۱۹۳۱
۱۱۳/۲۶۶۹	۱۴۴/۳۰۰۳	۵۹/۳۰۵۴	۸۲/۴۸۰۳	۹۳/۵۲۵۹	۴۲/۸۰۳۹	مدل ارائه شده در [۱۰۶]	۰/۱۱۳۲

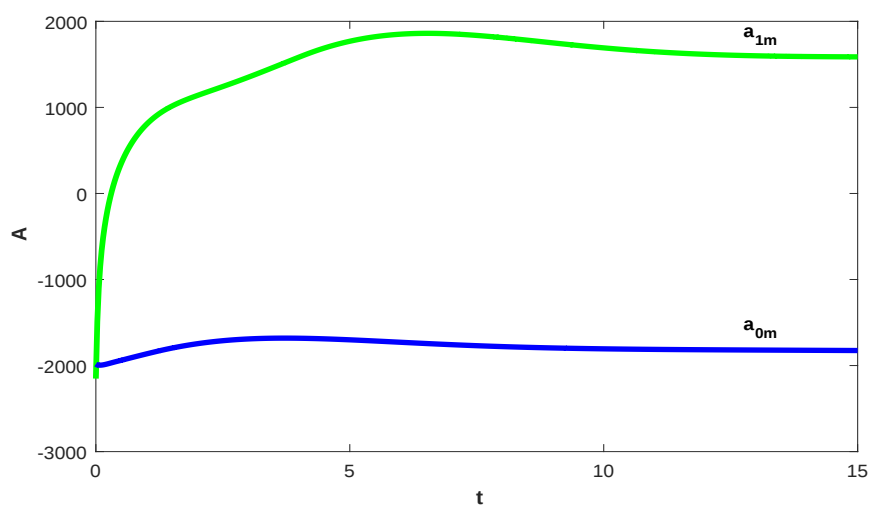
با مقایسه معیار نیکویی برازش با سه فاصله متفاوت D_2 ، D_4 و D_3 ، می‌توان نتیجه گرفت که روش پیشنهادی عملکرد بهتری در مقایسه با سایر روش‌ها دارد. همانطور که از جداول ۵.۴ مشاهده می‌شود مقادیر MCV و $MSPE$ تقریباً برابر است. که نشان دهنده عملکرد مناسب مدل برای این مثال است.

نمودارهای حبابی فازی نشان می‌دهد که مشاهدات به دلیل نزدیک بودن دایره‌ها به نیمساز صفحه و شعاع کوتاه هر دایره، به خوبی با مدل انتخاب شده متناسب است. بنابراین، نتیجه گرفته می‌شود که مدل پیشنهادی مورد استفاده در این مقاله از قدرت برازش خوبی برخوردار است. علاوه بر این در شکل‌های ۱۷.۴ و ۱۸.۴ نمودارهای تیلور برای مقایسه کارایی مدل رگرسیون بریج فازی پیشنهادی با مدل‌های موجود در [۴۱، ۸۹، ۱۰۵-۱۰۷، ۱۰۹] ترسیم شده است. موقعیت هر دایره روی نمودار نشان دهنده یک مدل متفاوت است که توسط مقادیر ضریب همبستگی و انحراف معیار تعیین می‌شود. واضح است که مدل پیشنهادی در مقایسه با سایر مدل‌های موجود از قابلیت بهتری برخوردار است.

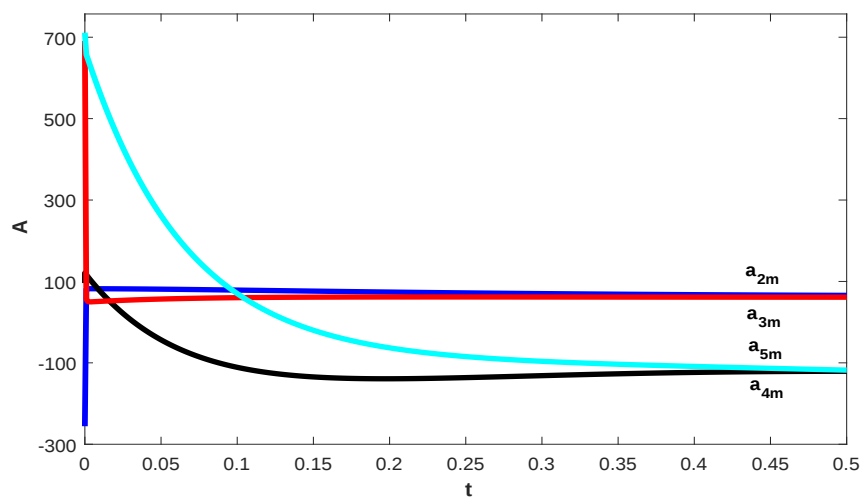
۶.۴ مزایای روش پیشنهادی

علاوه بر مزایای شماره ۱ تا ۹ بیان شده در ۴.۲.۲، برخی دیگر از مزایای روش پیشنهادی را به شرح زیر خلاصه می‌کنیم:

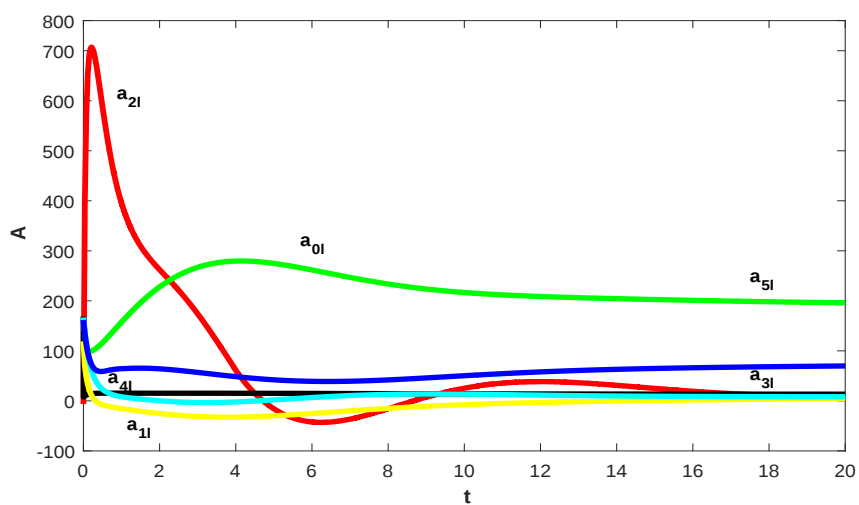
- به منظور بهره‌وری بیشتر، اعداد فازی در این بخش به شکل LR در نظر گرفته می‌شوند.
- برای نشان دادن برخی از مزایای مدل پیشنهادی نسبت به دیگر مدل‌ها، از نمودارهای آماری شناخته شده‌ای به نام‌های «نمودار تیلور» و «نمودار حبابی فازی» استفاده می‌کنیم.
- برای به تصویر کشیدن عملکرد روش پیشنهادی یک مثال شبیه سازی شده و یک مثال کاربردی برای نشان دادن کارایی مدل در حل مسائل واقعی ارائه شده است.



(ا)

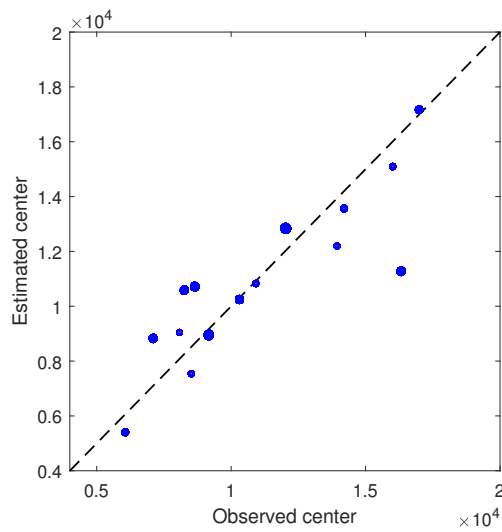


(ب)

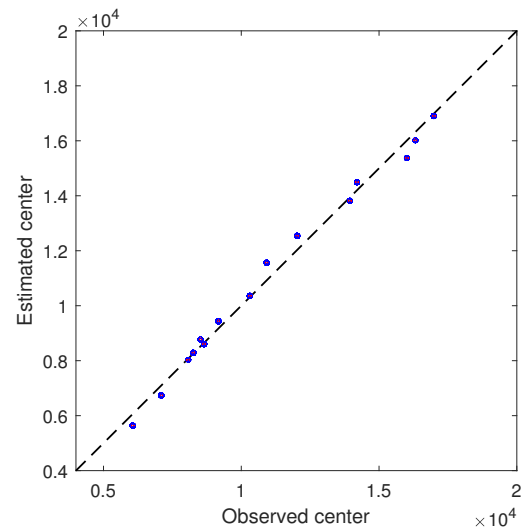


(ج)

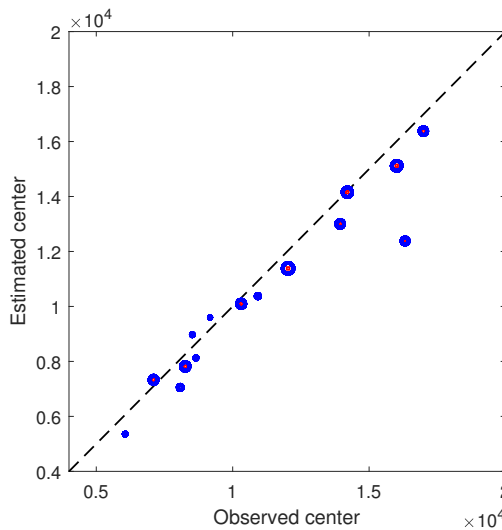
شکل ۱۳.۴: رفتار مسیرهای شبکه عصبی (۴۱.۴) و (۴۲.۴) در مثال ۲.۵.۴.



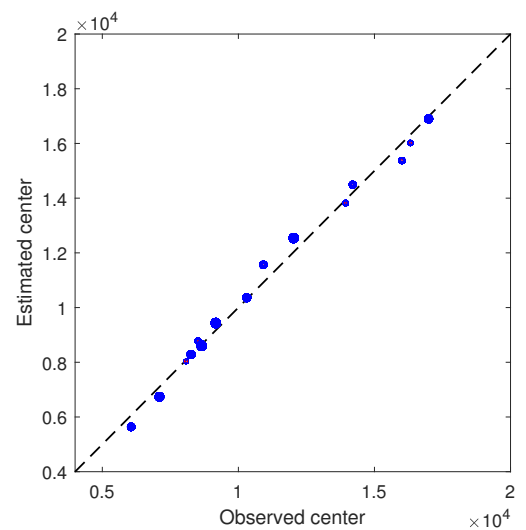
(ب) مدل ارائه شده در [۸۹]



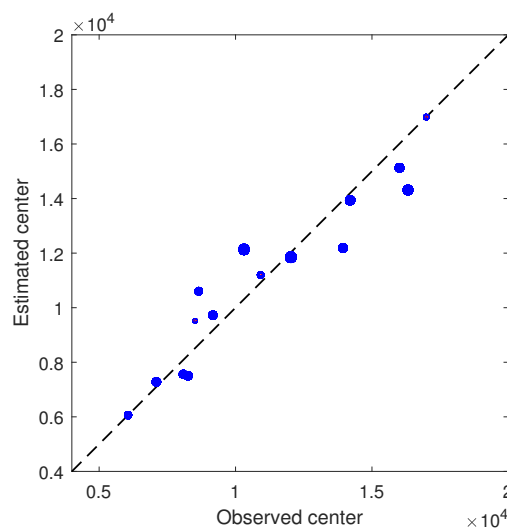
(ا) مدل رگرسیون بریج پیشنهادی با $\gamma = 2$



(د) مدل ارائه شده در [۴۱]

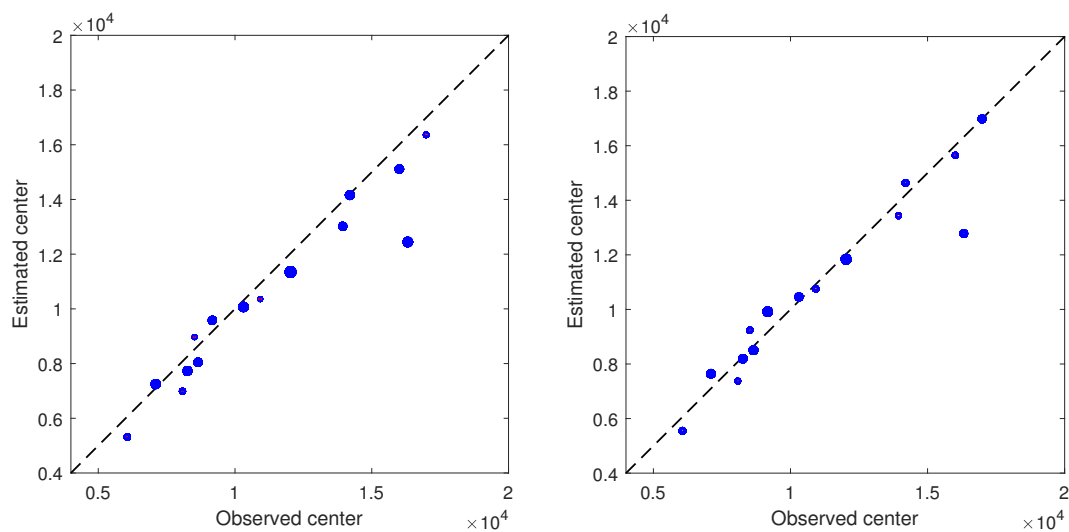


(ج) مدل ارائه شده در [۱۰۵]

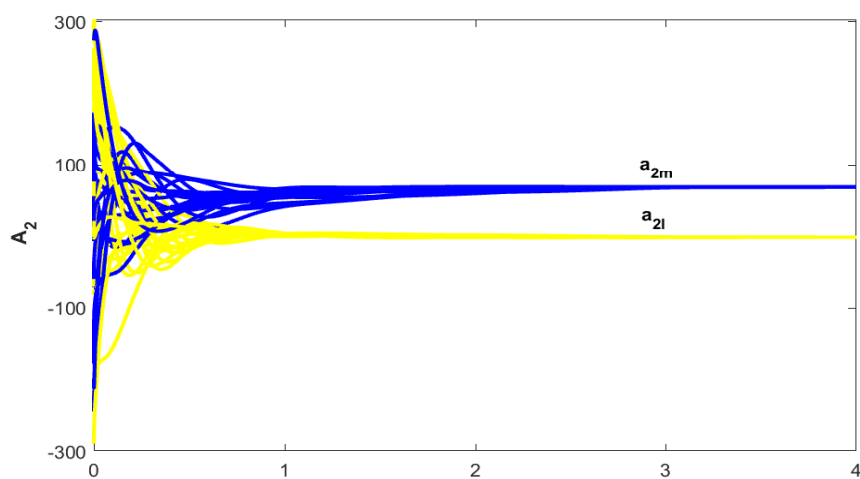


(ه) مدل ارائه شده در [۱۰۹]

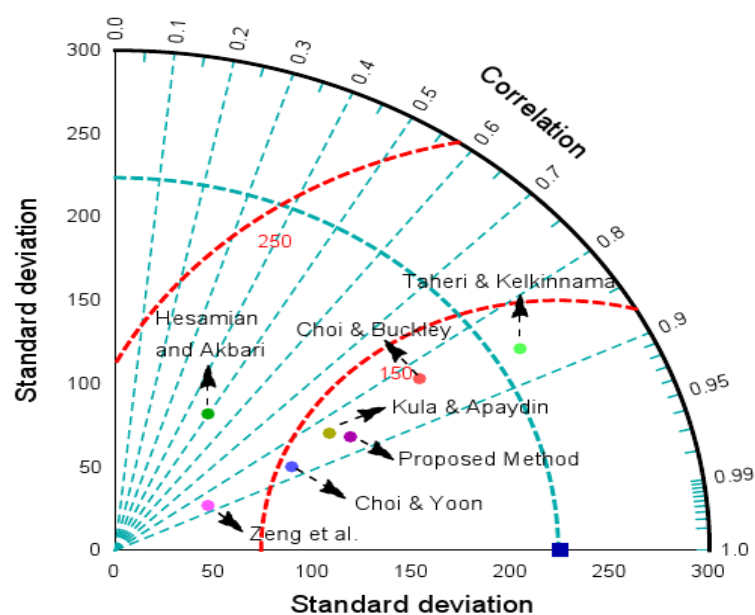
شکل ۱۴.۴: نمودار حبابی فازی در مثال ۱.۵.۴.



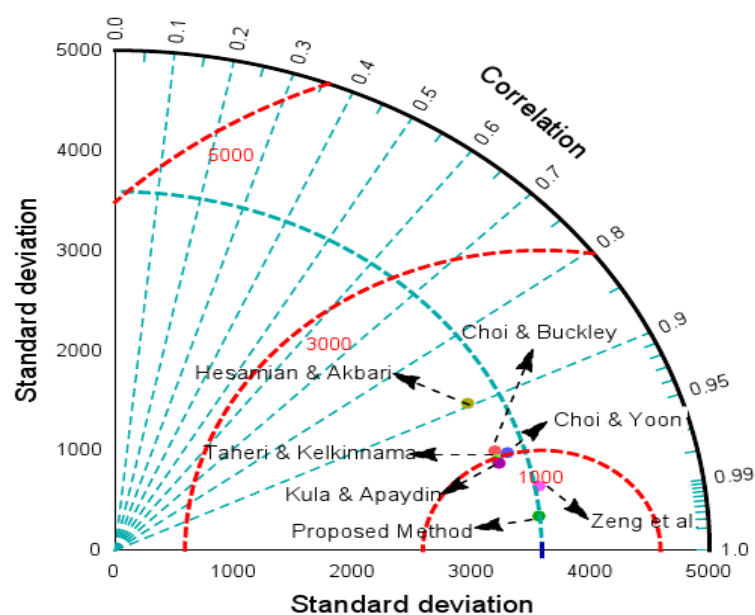
(ب) مدل ارائه شده در [۱۰۶]
 (آ) مدل ارائه شده در [۱۰۷]
 شکل ۱۵.۴: ادامه نمودار حبابی فازی در مثال ۱.۵.۴.



شکل ۱۶.۴: رفتار مسیر A_2 برای شبکه عصبی (41.4) و (42.4) با 20 نقطه شروع متفاوت در مثال ۲.۵.۴.



شکل ۱۷.۴: نمودار تیلور برای مراکز اعداد فازی در مثال ۱.۵.۴.



شکل ۱۸.۴: نمودار تیلور برای مقادیر سمت چپ و راست اعداد فازی در مثال ۱.۵.۴.

مراجع

- [۱] نیرومند ح، (۱۳۸۷)، ”تحلیل رگرسیون خطی: ابزاری برای تحقیق” چاپ اول، انتشارات دانشگاه فردوسی مشهد
- [۲] سلیمانپورباکفایت ا، (۱۹۹۳)، ”مفاهیم و کاربردهای شبکه‌های عصبی مصنوعی در علوم پایه قابل استفاه برای دانشجویان و محققین علوم پایه و فنی و مهندسی” چاپ اول، انتشارات ارومیه
- [۳] منهای م، (۱۳۹۴)، ”محاسبات فازی: هوش محاسباتی” چاپ اول، انتشارات دانش نگار
- [۴] طاهری س.م، (۱۳۷۸)، ”آشنایی با مجموعه‌های فازی” چاپ سوم، انتشارات جهاد دانشگاهی مشهد
- [5] Galton, F. (1886), ” Regression towards mediocrity in hereditary stature” **The Journal of the Anthropological Institute of Great Britain and Ireland**, 15, pp 246–263
- [6] Gunst, R.F. and Webster, J.T. (1975), ” Regression analysis and problems of multicollinearity” **Communications in Statistics-Theory and Methods**, 4,3, pp 277–292
- [7] Montgomery, D.C, Peck, E.A. and Vining, G.G. (2012), ” **Introduction to linear regression analysis**”, vol. 821, John Wiley & Sons
- [8] Marquardt, D.W. (1970), ” Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation” **Technometrics**, 12,3, pp 591–612
- [9] Akdeniz, F. and Kaçiranlar, S. (1995), ” On the almost unbiased generalized liu estimator and unbiased estimation of the bias and mse” **Communications in Statistics-Theory and Methods**, 24,7, pp 1789–1797

-
- [10] Liu, K. (2003), " Using liu-type estimator to combat collinearity" **Communications in Statistics-Theory and Methods**, 32,5, pp 1009–1020
 - [11] Swindel, B.F. (1976), " Good ridge estimators based on prior information" **Communications in Statistics-Theory and Methods**, 5,11, pp 1065–1075
 - [12] Frank, L.E. and Friedman, J.H. (1993), " A statistical view of some chemometrics regression tools" **Technometrics**, 35,2, pp 109–135
 - [13] Tibshirani, R. (1996), " Regression shrinkage and selection via the lasso" **Journal of the Royal Statistical Society: Series B (Methodological)**, 58,1, pp 267–288
 - [14] Hoerl, A.E. and Kennard, R.W. (1970), " Ridge regression: Biased estimation for nonorthogonal problems" **Technometrics**, 12,1, pp 55–67
 - [15] Diamond, P. and Kloeden, P. (1990), " Metric spaces of fuzzy sets" **Fuzzy Sets and Systems**, 35,2, pp 241–249
 - [16] Diamond, P. and Kloeden, P.E. (1994), " **Metric spaces of fuzzy sets: theory and applications**" World scientific
 - [17] Diamond, P. and Tanaka, H. (1998), " Fuzzy regression analysis" in **Fuzzy sets in decision analysis, operations research and statistics** Springer, pp. 349–387
 - [18] Poleshchuk, O. and Komarov, E. (2010), " Hybrid fuzzy least-squares regression model for qualitative characteristics" in **Integrated Uncertainty Management and Applications**, Springer, pp. 187–196
 - [19] Shen, S.L., Mei, C.L. and Cui, J.L. (2010), " A fuzzy varying coefficient model and its estimation" **Computers & Mathematics with Applications**, 60,6, pp 1696–1705
 - [20] DUrso, P. and Massari, R. (2013), " Weighted least squares and least median squares estimation for the fuzzy linear regression analysis" **Metron**, 71,3, pp 279–306
 - [21] Chan, K.Y., Lam, H.K., Dillon, T.S. and Ling, S.H. (2014), " A stepwise-based fuzzy regression procedure for developing customer preference models in new product development" **IEEE Transactions on Fuzzy Systems**, 23,5, pp 1728–1745
 - [22] Mashinchi, M.H., Orgun, M.A., Mashinchi, M. and Pedrycz W. (2011), " A Tabu–Harmony Search-Based Approach to Fuzzy Linear Regression" **IEEE Transactions on Fuzzy Systems**, 19,3, pp 432–448

- [23] Drezner, Z., Marcoulides, G.A. and Salhi, S. (1999), "Tabu search model selection in multiple regression analysis" **Communications in Statistics-Simulation and Computation**, 28,2, pp 349–367
- [24] Mendel, J.M. (2014), "On a novel way of processing data that uses fuzzy sets for later use in rule-based regression and pattern classification" **International Journal of Fuzzy Logic and Intelligent Systems**, 14,1, pp 1–7
- [25] Chakraborty, C. and Chakraborty, D. (2008), "Fuzzy linear and polynomial regression modelling of 'if-then'fuzzy rulebase" **International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems**, 16,2, pp 219–232
- [26] Rodríguez-Fdez, I., Mucientes, M. and Bugarín, A. (2016), "FRULER: Fuzzy rule learning through evolution for regression" **Information Sciences**, 354, pp1–18
- [27] Abbasbandy, S. and Otadi, M. (2006), "Numerical solution of fuzzy polynomials by fuzzy neural network" **Applied Mathematics and Computation**, 181,2, pp 1084–1089
- [28] Abbasbandy, S., Otadi, M. and Mosleh, M. (2008), "Numerical solution of a system of fuzzy polynomials by fuzzy neural network" **Information Sciences**, 178,8, pp 1948–1960
- [29] Mosleh, M., Otadi, M. and Abbasbandy, S. (2010), "Evaluation of fuzzy regression models by fuzzy neural network" **Journal of Computational and Applied Mathematics**, 234,3, pp 825–834
- [30] Buckley, J.J. and Eslami, E. (1997), "Neural net solutions to fuzzy problems: The quadratic equation" **Fuzzy Sets and Systems**, 86,3, 289–298
- [31] Hayashi, Y., Buckley, J.J. and Czogala, E. (1993), "Fuzzy neural network with fuzzy signals and weights" **Hayashi, Yoichi and Buckley, James J and Czogala, Ernest**, 8,4, pp 527–537
- [32] Ishibuchi, H., Kwon, K. and Tanaka, H. (1995), "A learning algorithm of fuzzy neural networks with triangular fuzzy weights" **Fuzzy sets and systems**, 71,3, pp 277–293
- [33] Diamond, P. (1988), "Fuzzy least squares" **Information Sciences**, 46,3, pp 141–157

- [34] Celmiņš, A. (1987), "Least squares model fitting to fuzzy vector data" **Fuzzy sets and systems**, 22,3, pp 245–269
- [35] Diamond, P. and Körner, R. (1997), "Extended fuzzy linear models and least squares estimates" **Computers & Mathematics with Applications**, 33,9, pp 15–32
- [36] Xu, R. and Li, C. (2001), "Multidimensional least-squares fitting with a fuzzy model" **Fuzzy sets and systems**, 119,2, pp 215–223
- [37] Coppi, R., DUrso, P., Giordani, P. and Santoro, A. (2006), "Least squares estimation of a linear regression model with LR fuzzy response" **Computational Statistics & Data Analysis**, 51,1, pp 267–286
- [38] Bishnu, S. A. (2003), "Speech synthesis based on linear prediction" **Academic Press**, pp 645–655
- [39] Wasserman, L. (2006), "All of nonparametric statistics" Springer Science & Business Media
- [40] Lu, J. and Wang, R. (2009), "An enhanced fuzzy linear regression model with more flexible spreads" **Fuzzy sets and systems**, 160,17, pp 2505–2523
- [41] Taheri, S.M. and Kelkinnama, M. (2012), "Fuzzy linear regression based on least absolute deviations" **Iranian Journal of Fuzzy Systems**, 9,1, pp 121–140
- [42] Taylor, K.E. (2001), "Summarizing multiple aspects of model performance in a single diagram" **Journal of Geophysical Research: Atmospheres**, 106,D7, pp 7183–7192
- [43] Hopfield, J.J. (1982) "Neural networks and physical systems with emergent collective computational abilities" **Proceedings of the National Academy of Sciences**, 79,8, pp 2554–2558
- [44] Hopfield, J.J. and Tank, D.W. (1985), "“neural” computation of decisions in optimization problems" **Biological Cybernetics**, 52,3, pp 141–152
- [45] Tank, D.f. and Hopfield, J. (1986), "Simple “neural” optimization networks: An A/D converter, signal decision circuit, and a linear programming circuit" **IEEE Transactions on Circuits and Systems**, 33,5, pp 533–541

- [46] Kennedy, M.P. and Chua, L.O. (1988), " Neural networks for nonlinear programming" **IEEE Transactions on Circuits and Systems**, 35,5, pp 554–562
- [47] Rodriguez-Vazquez, A., Dominguez-Castro, R., Rueda, A., Huertas, J.L. and Sanchez-Sinencio, E. (1990), " Nonlinear switched capacitor "neural" networks for optimization problems" **IEEE Transactions on Circuits and Systems**, 37,3, pp 384–398
- [48] Maa, C.Y. and Schanblatt, M.A. (1992), " A two-phase optimization neural network" **IEEE Transactions on Neural Networks**, 3,6, pp 1003–1009
- [49] Zhang, S. and Constantinides, A.G. (1992), " Lagrange programming neural networks" **IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing**, 39,7, pp 441–452
- [50] Bouzerdoun, A. and Pattison, T.R. (1993), " Neural network for quadratic optimization with bound constraints" **IEEE Transactions on Neural Networks**, 4,2, pp 293–304
- [51] Effati, S. and Baymani, M. (2005), " A new nonlinear neural network for solving convex nonlinear programming problems" **Applied Mathematics and Computation**, 168,2, pp 1370–1379
- [52] Effati, S., Ghomashi, A. and Abbasi, M. (2011), " A novel recurrent neural network for solving mlcps and its application to linear and quadratic programming problems" **Asia-Pacific Journal of Operational Research**, 28,04, pp 523–541
- [53] Effati, S., Ghomashi, A. and Nazemi, A.R. (2007), " Application of projection neural network in solving convex programming problems" **Applied Mathematics and Computation**, 188,2, pp 1103–1114
- [54] Sabouri, J., Effati, S. and Pakdaman, M. (2017), " A neural network approach for solving a class of fractional optimal control problems" **Neural Processing Letters**, 45,1, pp 59–74
- [55] Avriel, M. (1976), " **Nonlinear programming: Analysis and methods**" Prentice-Hall, New Jersey

-
- [56] Nazemi, A. (2013), " Solving general convex nonlinear optimization problems by an efficient neurodynamic model" **Engineering Applications of Artificial Intelligence**, 26,2, pp 685–696
- [57] Ortega, J.M. and Rheinboldt, W.C. (1970), " **Iterative solution of nonlinear equations in several variables**" Academic Press, New York
- [58] Zabczyk, J. (1997), " Mathematical control theory: An introduction " **Applications of Mathematics**, 42,1, pp 1–80
- [59] Hale, J.K. (1969), " **Ordinary differential equations**" Wiley-Interscience, New York
- [60] Bazaraa, M.S., Sherali, H.D. and Shetty, C.M. (2013), " **Nonlinear programming: theory and algorithms**" John Wiley & Sons
- [61] Quarteroni, A., Sacco, R. and Saleri, F. (2010), " **Numerical mathematics**, Vol. 37, Springer Science & Business Media
- [62] Fukushima, M. (1992), " Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems" **Mathematical Programming**, 53, pp 99–110
- [63] Miller, R.K. and michel A.N. (1982), " **Ordinary differential equations**" Academic Press, New York
- [64] Ishibuchi, H. and Tanaka, H. (1992), " Fuzzy regression analysis using neural networks" **Fuzzy Sets and Systems**, 50,3, pp 257–265
- [65] Mosleh, M., Otadi, M. and Abbasbandy, S. (2011), " Fuzzy polynomial regression with fuzzy neural networks" **Applied mathematical modelling**, 35,11, pp 5400–5412
- [66] Mosleh, M., Allahviranloo, T. and Otadi, M. (2012), " Evaluation of fully fuzzy regression models by fuzzy neural network" **Neural Computing and Applications**, 21,1, pp 105–112
- [67] Liu, H.T., Wang, J., He, Y.L. and Ashfaq, R.A.R. (2017), " Extreme learning machine with fuzzy input and fuzzy output for fuzzy regression" **Neural Computing and Applications**, 28,11, pp 3465–3476

- [68] Yeylaghi, S., Otadi, M. and Imankhan, N. (2017), " A new fuzzy regression model based on interval-valued fuzzy neural network and its applications to management" **Beni-Suef University Journal of Basic and Applied Sciences**, 6,2, pp 106–111
- [69] Chukhrova, N. and Johannssen, A. (2019), " Fuzzy regression analysis: systematic review and bibliography" **Applied Soft Computing**, 84, pp 105708
- [70] Fazlollahtabar, H. and Gholizadeh, H. (2020), " Fuzzy possibility regression integrated with fuzzy adaptive neural network for predicting and optimizing electrical discharge machining parameters" **Computers & Industrial Engineering**, 140, pp 106225
- [71] Fullér, R. (1995), " Neural fuzzy systems" **Abo Akademi University**
- [72] Avriel, M. (2003), " **Nonlinear programming: analysis and methods**" Courier Corporation
- [73] Nazemi, A.R. (2012), " A dynamic system model for solving convex nonlinear optimization problems" **Communications in Nonlinear Science and Numerical Simulation**, 17,4, pp 1696–1705
- [74] Otadi, M. (2014), " Fully fuzzy polynomial regression with fuzzy neural networks" **Neurocomputing**, 142, pp 486–493
- [75] Zhou, J., Zhang, H., Gu, Y. and Pantelous, A.A. (2018), " Affordable levels of house prices using fuzzy linear regression analysis: the case of shanghai" **Soft Computing**, 22, pp 1–12
- [76] Kao, C. and Chyu, C.L. (2003), " Least-squares estimates in fuzzy regression analysis" **European Journal of Operational Research**, 148,2, pp 426–435
- [77] Tanaka, H., Hayashi, I. and Watada, J. (1989), " Possibilistic linear regression analysis for fuzzy data." **European Journal of Operational Research**, 40,3, pp 389–396
- [78] Sakawa, M. and Yano, H. (1992), " Multiobjective fuzzy linear regression analysis for fuzzy input-output data" **Fuzzy Sets and Systems**, 47,2, pp 173–181
- [79] DUrso, P., Massari, R. and Santoro, A. (2011), " Robust fuzzy regression analysis" **Information Sciences**, 181,19, pp 4154–4174

-
- [80] Gong, Y., Yang, S., Ma, H. and Ge, M. (2018), " Fuzzy regression model based on incentre distance and application to employee performance evaluation" **International Journal of Fuzzy Systems**, 20,8, pp 2632–2639
- [81] Zadeh, L.A. (1975), " The concept of a linguistic variable and its application to approximate reasoning" **Information sciences**, 8,3, pp 199–249
- [82] Hoerl, A.E. and Kennard, R.W. (1970), " Ridge regression: Biased estimation for nonorthogonal problems" **Technometrics**, 12,1, pp 55–67
- [83] Lawson, C.L. and Hanson, R.J. (1974), " **Solving least squares problems** prentice-hall, Englewood Gifts
- [84] Roozbeh, M. and Arashi, M. (2016), " Shrinkage ridge regression in partial linear models" **Communications in Statistics-Theory and Methods**, 45,20, pp 6022–6044
- [85] Nasrabadi, E. and Hashemi, S.M. (2008), " Robust fuzzy regression analysis using neural networks" **International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems**, 16,4, pp 579–598
- [86] Choi, S. H., Jung, H.Y. and Kim, H. (2019), " Ridge fuzzy regression model" **International Journal of Fuzzy Systems**, 21,7, pp 2077–2090
- [87] Hong, D.H., Hwang, C. and Ahn, C. (2004), " Ridge estimation for regression models with crisp inputs and Gaussian fuzzy output" **Fuzzy Sets and Systems**, 142,2, pp 307–319
- [88] McDonald, G.C. and Galarneau, D.I. (1975), " A monte carlo evaluation of some ridge-type estimators" **Journal of the American Statistical Association**, 70,350, pp 407–416
- [89] Hesamian, G. and Akbari, M.G. (2019), " Fuzzy lasso regression model with exact explanatory variables and fuzzy responses" **International Journal of Approximate Reasoning**, 115, pp 290–300
- [90] Marquardt, D.W. (1970), " Generalized inverses, ridge regression, biased linear estimation and nonlinear estimation" **Technometrics**, 12,3, pp 591–612
- [91] Zou, H. and Yuan, M. (2008), " The F_{∞} - norm support vector machine" **Statistica Sinica**, 18,1, pp 379–398

- [92] Knight, K. and Fu, W. (2000), " Asymptotics for lasso-type estimators" **Annals of statistics**, 28,5, pp 1356–1378
- [93] Huang, J., Horowitz, J.L, Ma, S. (2028), " Asymptotic properties of bridge estimators in sparse high-dimensional regression models" **The Annals of Statistics**, 36,2, pp 587–613
- [94] Xu, Z., Zhang, H., Wang, Y., Chang, X.Y. and Liang, Y. (2010), " L1/2 regularization" **Science China Information Sciences**, 53,6, pp 1159–1169
- [95] Park, C. and Yoon, Y.J. (2011), " Bridge regression: adaptivity and group selection" **Journal of Statistical Planning and Inference**, 141,11, pp 3506–3519
- [96] Polson, N.G., Scott, J.G. and Windle, J. (2014), " The bayesian bridge" **Journal of the Royal Statistical Society: Series B: Statistical Methodology**, 76,4, pp 713–733
- [97] Mallick, H. and Yi, N. (2018), " Bayesian bridge regression" **Journal of Applied Statistics**, 45,6, pp 988–1008
- [98] Tian, Y. and Song, X. (2020), " Bayesian bridge-randomized penalized quantile regression" **Computational Statistics & Data Analysis**, 144, pp 106876
- [99] Tibshirani, R., Saunders, M., Rosset, S., Zhu, J. and Knight, K. (2005), " Sparsity and smoothness via the fused lasso" **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, 67,1, pp 91–108
- [100] Zou, H. (2006), " The adaptive lasso and its oracle properties" **Journal of the American statistical association**, 101,476, pp 1418–1429
- [101] Park, T. and Casella, G. (2008), " The bayesian lasso" **Journal of the American Statistical Association**, 103,482, pp 681–686
- [102] Jubril, A.M. (2012), " A nonlinear weights selection in weighted sum for convex multiobjective optimization" **Facta Universitatis**, 27,3, pp 357–372
- [103] Sawaragi, Y., Nakayama, H., and Tanino, T. (1985), " **Theory of multiobjective optimization**" Vol. 176, Academic Press, New York
- [104] Mohammadi, M., Tan, Y.H., Hofman, W. and Mousavi, S.H. (2018), " A novel one-layer recurrent neural network for the l1-regularized least square problem" **Neuro-computing**, 315, pp 135–144

- [105] Zeng, W., Feng, Q. and Li, J. (2017), " Fuzzy least absolute linear regression" **Applied Soft Computing**, 52, pp 1009–1019
- [106] Kula, K. and Apaydin, A. (2008), " Fuzzy robust regression analysis based on the ranking of fuzzy sets" **International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems**, 16,05, pp 663–681
- [107] Choi, S.H. and Buckley, J. (2008), " Fuzzy regression using least absolute deviation estimators" **Soft computing**, 12,3, pp 257–263
- [108] Henderson, H.V. and Velleman, P.F. (1981), " Building multiple regression models interactively" **Biometrics**, 37,2 pp 391–411
- [109] Choi, S.H. and Yoon, J.H. (2010), " General fuzzy regression using least squares method. **International Journal of Systems Science**, 41,5, pp 477–485
- [110] Zadeh, L.A. (1965), " Fuzzy sets" **Information and Control**, 8,3, pp 338–353
- [111] Dubois, D.J. (1980), "**Fuzzy Sets and Systems: Theory and Applications**" Vol. 144, Academic Press, New York
- [112] Zimmermann, H.J. (2011), " **Fuzzy set theory and its applications**" Springer Science & Business Media
- [113] Dubois, D. and Prade, H. (1980), " **Fuzzy sets and systems: theory and applications**" Mathematics in science and engineering
- [114] Davison, A.C. and Hinkley, D.V. (1997), " **Bootstrap methods and their application**" Vol. 1, Cambridge university press
- [115] Hassanpour, H., Maleki, H.R. and Yaghoobi, M.A. (2010), " Fuzzy linear regression model with crisp coefficients: a goal programming approach" **Iranian Journal of Fuzzy Systems**, 7,2, pp 19–39
- [116] Xu, R. (1991), " A linear regression model in fuzzy environment" **Advance in Modelling Simulation**, 27, pp 31–40

پیوست آ

نظریه مجموعه‌های فازی

نظریه مجموعه‌های فازی اولین بار در سال (۱۹۶۵) توسط دکتر لطفی عسگر زاده^۱ ابداع شد. این نظریه از زمان ارائه آن تا کنون گسترش و تعمیق زیادی یافته و کاربرهای گوناگونی در زمینه‌های مختلف پیدا کرده است. در واقع این نظریه، نظریه کار در شرایط عدم اطمینان و شرایط مبهم می‌باشد. این نظریه قادر است بسیاری از مفاهیم، متغیرها و دستگاه‌هایی که نادقیق هستند را صورت‌بندی ریاضی بخشیده و زمینه را برای استدلال، کنترل و تصمیم‌گیری در شرایط عدم اطمینان، نادقیق یا مبهم فراهم آورد. مطالب این بخش عمدتاً از [۱۱۰-۱۱۲] گرفته شده است.

۱.۱. مروری بر مفاهیم نظریه مجموعه‌های فازی

معیار عضویت هر عنصر در یک مجموعه، صفت مشخص کننده مجموعه است. هر عنصری که دارای آن صفت باشد، عضو مجموعه و در غیر این صورت خارج از مجموعه است. این معیار عضویت را تابع مشخصه می‌نامند و به صورت زیر نشان می‌دهند:

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

^۱Zade

که در آن $x \in X$ (مجموعه مرجع)، برد آن مجموعه دو عضوی $\{0, 1\}$ و دامنه آن X است. حال اگر برد تابع مشخصه از مجموعه دو عضوی $\{0, 1\}$ به بازه $[0, 1]$ توسعه یابد، تابعی خواهیم داشت که به هر از X عددی را از بازه $[0, 1]$ نسبت می‌دهد. این تابع را تابع عضویت A می‌گویند. اکنون یک مجموعه قطعی^۲ به شمار نمی‌آید. چنین مجموعه‌هایی را با یک تابع عضویت نشان می‌دهند و به آنها زیر مجموعه‌های فازی گفته می‌شود.

تعریف ۱.۱.۱. (زیر مجموعه فازی) فرض کنید X یک مجموعه قطعی باشد، زیر مجموعه فازی A از مجموعه X ، مجموعه‌ای از زوج‌های مرتب به صورت زیر است:

$$A = \{(x, \mu_A(x)) \mid x \in X\}$$

که در آن $\mu_A : X \rightarrow [0, 1]$ تابع عضویت و $\mu_A(x)$ درجه عضویت x در A نامیده می‌شود.

تعریف ۲.۱.۱. (تکیه‌گاه یک مجموعه فازی) تکیه‌گاه مجموعه فازی A در X ، یک مجموعه غیرفازی است و شامل تمام عضوهایی از X می‌باشد که دارای درجه عضویت مثبت هستند و به صورت زیر تعریف می‌شوند:

$$Supp(A) = \{x \in X \mid \mu_A(x) > 0\}$$

تعریف ۳.۱.۱. (ارتفاع یک مجموعه فازی) ارتفاع یک مجموعه فازی برابر حداکثر درجه عضویت عناصر آن مجموعه است. یعنی

$$h(A) = \sup_{x \in X} \{\mu_A(x)\}$$

تعریف ۴.۱.۱. (مجموعه فازی نرمال) مجموعه فازی A نرمال است، اگر ارتفاع آن برابر یک باشد. یعنی $h(A) = 1$.

تعریف ۵.۱.۱. (h -برش یک مجموعه فازی) h -برش یک مجموعه فازی A ، یک مجموعه غیرفازی $A[h]$ است که شامل تمام عضوهایی از X می‌باشد که مقادیر عضویتی بزرگ‌تر یا مساوی h دارند. یعنی:

$$A[h] = \{x \in X \mid \mu_A(x) \geq h\}, \quad h \in (0, 1]$$

و برای $h = 0$ ، $A[0] = \overline{\{x \mid \mu_A(x) > 0\}}$ ، که در آن منظور از $\overline{\{\cdot\}}$ ، بستار مجموعه است.

تعریف ۶.۱.۱. مجموعه فازی A محدب است اگر برای هر $x, y \in X$ و $\lambda \in [0, 1]$ داشته باشیم

$$\mu_A(\lambda x + (1 - \lambda)y) \geq \min\{\mu_A(x), \mu_A(y)\}$$

یا به عبارت دیگر یک مجموعه فازی محدب است اگر کلیه مجموعه‌های h -برش آن محدب باشند.

^۲Crisp

h – برش‌ها توصیفی از مجموعه‌های فازی با استفاده از مجموعه‌های قطعی ارائه می‌دهند و دارای ویژگی‌های زیر هستند:

۱. خانواده $\{A[h] \mid h \in (0, 1]\}$ یکنواست، یعنی:

$$0 < h \leq h' \leq 1 \implies A[h'] \subseteq A[h]$$

۲. برای هر h عضو بازه $(0, 1]$ ، $A[h] \subseteq B[h]$ اگر و فقط اگر $A \subseteq B$.

$$(A \cup B)[h] = A[h] \cup B[h] \quad (A \cap B)[h] = A[h] \cap B[h] \quad ۳.$$

h – برش‌های یک مجموعه فازی، خانواده‌ای از مجموعه‌های قطعی تشکیل می‌دهند که می‌تواند برای نمایش یک مجموعه فازی داده شده A در X ، استفاده شود. این مطلب را در قالب قضیه زیر بیان می‌کنیم.

قضیه ۱.۱.۱. [۴] فرض کنید A یک مجموعه فازی در X با تابع عضویت $\mu_A(x)$ باشد و همچنین $A[h]$ ، h – برش مجموعه فازی A و $\chi_{A[h]}(x)$ تابع مشخصه مجموعه قطعی $A[h]$ برای $h \in (0, 1]$ باشد، آن‌گاه:

$$\mu_A(x) = \sup_{h \in (0, 1]} (h \wedge \chi_{A[h]}(x)), \quad x \in X$$

نتیجه ۱.۱.۱. (اصل تجزیه^۳) هر مجموعه فازی A در X می‌تواند بر حسب h – برش‌های مربوط به خود به صورت زیر نمایش داده شود:

$$A = \bigcup_{h \in (0, 1]} (hA[h])$$

یکی از مفاهیم اساسی نظریه مجموعه‌های فازی اصل گسترش است که می‌توان برای تعمیم مفاهیم ریاضی قطعی به مجموعه‌های فازی به کار گرفته شود. این اصل به ویژه در تعمیم عملگرهای جبری و تعریف این عملگرها برای اعداد فازی مفید است. یک تابع معمولی $f: X \rightarrow Y$ را در نظر بگیرید. این تابع به هر $x \in X$ عددی مانند $y \in Y$ را نسبت می‌دهد. حال می‌خواهیم تابع f را به گونه‌ای گسترش دهیم که به جای این که بر یک نقطه از X عمل کند، بر یک زیر مجموعه فازی آن عمل کند. آن چه مهم است تعریف مقدار حاصل از عمل f بر یک زیر مجموعه فازی از X مانند A است. مسلماً انتظار داریم که $f(A)$ دیگر یک نقطه از Y نباشد بلکه یک زیر مجموعه فازی از Y مانند B باشد. اصل گسترش روش این تعمیم را بیان می‌کند.

^۳Decomposition Principle

تعریف ۷.۱.۱. (اصل گسترش) فرض کنید $f : X \rightarrow Y$ یک تابع و A زیر مجموعه‌ای فازی از X باشد، در این صورت $f(A) = B$ یک زیر مجموعه فازی از Y خواهد بود به طوری که:

$$\mu_B(y) = \begin{cases} \sup_{x, y=f(x)} \{\mu_A(x)\} & \text{if } f^{-1}(y) \neq \emptyset \\ 0 & \text{if } f^{-1}(y) = \emptyset \end{cases}$$

۱.۱.۱. اعداد فازی

مجموعه‌های فازی که روی اعداد حقیقی تعریف می‌شوند از اهمیت بالایی برخوردارند. تابع عضویت این نوع مجموعه‌های فازی به صورت زیر می‌باشد:

$$\mu : \mathbb{R} \rightarrow [0, 1]$$

این نوع از مجموعه‌های فازی مفهوم کمی داشته و تحت شرایط خاصی می‌توانند به عنوان اعداد فازی یا بازه‌های فازی مطرح شوند. اعداد و بازه‌های فازی در مدل‌سازی و تحلیل کنترل فازی، تصمیم‌گیری فازی، استدلال تقریبی، بهینه‌سازی و آمار و احتمالات تقریبی نقش بسیار مهمی ایفا می‌کنند.

تعریف ۸.۱.۱. (عدد فازی) فرض کنید A یک مجموعه فازی در $\mathbb{R}$ باشد. A یک عدد فازی نامیده می‌شود اگر دارای شرایط زیر باشد:

۱. نرمال باشد،

۲. $A(h), \forall h \in (0, 1]$ بازه‌ای بسته باشد.

۳. تکیه‌گاه A کراندار باشد.

تعریف ۹.۱.۱. (عدد فازی LR) عدد فازی A ، را یک عدد فازی LR گویند هرگاه:

$$\mu_A(x) = \begin{cases} L\left(\frac{a_m - x}{a_l}\right) & \text{if } x \leq m \\ R\left(\frac{x - a_m}{a_r}\right) & \text{if } x \geq m \end{cases}$$

که در آن L و R توابع پیوسته غیرصعودی از $\mathbb{R}^+$ به $[0, 1]$ هستند و در شرط $L(0) = R(0) = 1$ صدق می‌کنند. پارامترهای a_l و a_r نیز اعداد حقیقی مثبت‌اند. در این حالت a_m را مقدار میانه و a_l و a_r را به ترتیب پهنای راست و چپ عدد فازی می‌نامند. در این صورت A به شکل $A = (a_m, a_l, a_r)_{LR}$ نمایش داده می‌شود.

در تعریف بالا L و R توابع مرجع نامیده می‌شوند و معمولاً یکی از حالات زیر را اختیار می‌کنند:

۱. خطی:

$$L(x) = \max(0, 1 - x)$$

۲. نمایی:

$$L(x) = e^{-px}, \quad p \geq 1$$

۳. توان‌دار:

$$L(x) = \max(0, 1 - x^p), \quad p \geq 1$$

۴. کسری:

$$L(x) = \frac{1}{1 + x^p}, \quad p \geq 1$$

۵. نمایی توان‌دار:

$$L(x) = e^{-x^p}, \quad p \geq 1$$

برای $p = 2$ ، عدد فازی نرمال به وجود می‌آید و به صورت $A = (a_m, a_l, a_r)_N$ نمایش داده می‌شود.

تعریف ۱.۱.۱۰. (عدد فازی مثلثی) اگر در تعریف عدد فازی LR ، توابع L و R توابع خطی باشند، عدد فازی حاصل را عدد فازی مثلثی می‌نامند و آن را به صورت $A = (a_m, a_l, a_r)_T$ نمایش می‌دهند که تابع عضویت آن به شکل زیر است:

$$\mu_A(x) = \begin{cases} 1 - \left(\frac{a_m - x}{a_l}\right) & \text{if } a_m - a_l \leq x \leq a_m \\ 1 - \left(\frac{x - a_m}{a_r}\right) & \text{if } a_m \leq x \leq a_r + a \end{cases}$$

۲.۱.۱ حساب اعداد فازی

اگر A و B دو عدد فازی باشند و تعمیم عملگر $*$ را برای اعداد فازی با $\otimes$ نشان دهیم، با استفاده از اصل توسیع، حاصل $A \otimes B$ به صورت یک مجموعه فازی با تابع عضویت زیر تعریف می‌شود.

$$(A \otimes B)(z) = \sup_{z=x*y} \min[\mu_A(x), \mu_B(y)]$$

در حالت‌های خاص برای دو عمل اصلی جمع و تفریق، تعریف بالا به صورت‌های زیر در می‌آید

$$(A \oplus B)(z) = \sup_{z=x+y} \min[\mu_A(x), \mu_B(y)]$$

$$(A \ominus B)(z) = \sup_{z=x-y} \min[\mu_A(x), \mu_B(y)]$$

در این پایان نامه برای سادگی به جای علامت‌های $\oplus$ و $\ominus$ از $+$ و $-$ استفاده می‌کنیم.

برخی اعمال جبری برای اعداد فازی LR بسیار ساده و دارای الگوهای مشخصی است و برای برخی دیگر از اعمال جبری نیز الگوهای تقریبی وجود دارد. در این قسمت برخی روابط مهم در حساب اعداد فازی LR بیان می‌شوند.

قضیه ۲.۱.۱. اگر $A = (a_m, a_l, a_r)_{LR}$ ، $B = (b_m, b_l, b_r)_{LR}$ و $k \in \mathbb{R}$ آن‌گاه بر اساس اصل گسترش بیان شده در تعریف ۷.۱.۱ داریم:

$$(a_m, a_l, a_r)_{LR} + (b_m, b_l, b_r)_{LR} = (a_m + b_m, a_l + b_l, a_r + b_r)_{LR}$$

$$(a_m, a_l, a_r)_{LR} - (b_m, b_l, b_r)_{LR} = (a_m - b_m, a_l - b_l, a_r - b_r)_{LR}$$

$$k(a_m, a_l, a_r)_{LR} = (ka_m, ka_l, ka_r)_{LR} \quad k > 0$$

$$k(a_m, a_l, a_r)_{LR} = (ka_m, -ka_r, -ka_l)_{LR} \quad k < 0$$

درباره ی عمل ضرب، کار به آسانی جمع و تفریق نیست. در حالتی که A و B دو عدد فازی LR باشند نتیجه ضرب آنها یک عدد فازی LR نخواهد بود. اما برای ضرب روابط تقریبی متفاوتی پیشنهاد شده است (برای اطلاعات بیشتر در این زمینه به [۱۱۳] مراجعه شود).

۳.۱.۱ فاصله اعداد فازی

در بحث رگرسیون فازی علاقه‌مند به تخمین و برآورد پارامترهای مدل هستیم و لذا لازم است خطا و به عبارت دقیق‌تر فاصله بین دو مقدار مشاهده شده و برآورد شده فازی محاسبه شود، لذا در این بخش به معرفی برخی از انواع فاصله بین دو عدد فازی که در طول پایان‌نامه استفاده شده‌اند می‌پردازیم.

فاصله اقلیدسی

تعریف ۱۱.۱.۱. [۱۱۴] میانگین فاصله اقلیدسی بین دو عدد فازی A و B برای k امین h - برش به صورت زیر تعریف شده است:

$$D_{h_k}(A, B) = \frac{1}{2} \left[(\underline{A}(h_k) - \underline{B}(h_k))^2 + (\overline{A}(h_k) - \overline{B}(h_k))^2 \right]$$

که در آن $A[h_k] = [\underline{A}(h_k), \overline{A}(h_k)]$ و $B[h_k] = [\underline{B}(h_k), \overline{B}(h_k)]$ ، به ترتیب، k امین h - برش A و B هستند. با قرار دادن $1 = h_g < \dots < h_2 < h_1 = 0$ و جمع روی k به فاصله زیر می‌رسیم:

$$D_1(A, B) = \sum_{k=1}^g D_{h_k}(A, B) \quad (۱.۱)$$

فاصله حسن‌پور

حسن‌پور و همکاران [۱۱۵] فاصله بین دو عدد فازی را به صورت زیر بیان کردند.

تعریف ۱۲.۱.آ. فرض کنید A و B دو عدد فازی مثلثی باشند. فاصله D_2 را به صورت زیر تعریف می‌کنیم

$$D_2(A, B) = |a_m - b_m| + |a_l - b_l| + |a_r - b_r|$$

اثبات می‌شود که فاصله D_2 در تعریف بالا یک متر است.

فاصله وزن دار شده

ژو در [۱۱۶] فاصله وزن دار شده بین دو عدد فازی را به صورت زیر تعریف می‌کند:

تعریف ۱۳.۱.آ. فرض کنید A و B دو عدد فازی باشند، فاصله وزن دار شده بین این دو عدد فازی به صورت

$$D_3(A, B) = \left(\int_0^1 f(h) d^2(A[h], B[h]) dh \right)^{1/2}$$

است، که در آن

$$A[h] = [\underline{A}(h), \overline{A}(h)], \quad B[h] = [\underline{B}(h), \overline{B}(h)]$$

$$d^2(A[h], B[h]) = (\underline{A}(h) - \underline{B}(h))^2 + (\overline{A}(h) - \overline{B}(h))^2$$

همچنین $f(h)$ یک تابع صعودی روی بازه $[0, 1]$ است و داریم $f(0) = 0$ و $\int_0^1 f(h) dh = 1/2$. تابع $f(h)$ را می‌توان به عنوان وزن $d^2(A[h], B[h])$ تعبیر کرد. ویژگی صعودی بودن $f(h)$ منجر به این می‌شود که درجه عضویت‌های بالاتر، اهمیت بیشتری در تعیین فاصله بین دو عدد فازی A و B داشته باشند. شرط‌های $f(0) = 0$ و $\int_0^1 f(h) dh = 1/2$ نیز این اطمینان را می‌دهد که فاصله ی فوق، یک تعمیم معمولی از فاصله ی بین دو عدد حقیقی است.

نتیجه ۲.۱.آ. فرض کنید A و B دو عدد فازی مثلثی باشند آن‌گاه بر پایه تابع وزنی $f(h) = h$ داریم:

$$D_3^2(A, B) = (a_m - b_m)^2 + \frac{1}{3}(a_m - b_m)[(a_r - b_r) - (a_l - b_l)] + \frac{1}{12}[(a_l - b_l)^2 + (a_r - b_r)^2]$$

فاصله کپی و همکاران

تعریف ۱۴.۱.آ. نویسندگان در [۳۷] فاصله میان دو عدد فازی LR ، A و B را به صورت زیر تعریف کرده‌اند:

$$D_4^2(A, B) = (a_m - b_m)^2 + ((a_m - \gamma a_l) - (b_m - \gamma b_l))^2 + ((a_m + \rho a_u) - (b_m + \rho b_u))^2$$

که در آن $\gamma = \int_0^1 L^{-1}(w) dw$ و $\rho = \int_0^1 R^{-1}(w) dw$ تنوع تابع عضویت در مدل را در نظر می‌گیرند و هم‌زمان تأثیر پهناها را کم می‌کنند (برای اطلاعات بیشتر به [۷۹] مراجعه شود).

پیوست ب

اثبات قضایا و گزاره‌ها

ب.۱ فصل ۲

قضیه ۱.۲.۲:

برهان. فرض کنید $y^* = (A^{*\top}, U^{*\top})^\top$ ، یک نقطه تعادل از (۴۲.۲) و (۴۳.۲) باشد. بنابراین نتیجه می‌شود که

$$QA^* + B + G^\top (U^* + GA^*)^+ = 0 \quad (\text{ب.۱})$$

$$(U^* + GA^*)^+ = U^* \quad (\text{ب.۲})$$

از لم ۲.۳.۱، داریم: $(U^* + GA^*)^+ = U^*$ اگر و تنها اگر

$$U^* \geq 0, \quad GA^* \leq 0, \quad U^{*\top}(GA^*) = 0 \quad (\text{ب.۳})$$

علاوه بر این با جایگذاری (ب.۲) در (ب.۱) داریم:

$$QA^* + B + G^\top U^* = 0 \quad (\text{ب.۴})$$

بنابراین دیده می‌شود که $y^* = (A^{*\top}, U^{*\top})^\top$ در شرایط کروش-کان-تاکر (۴۱.۲) صدق می‌کند.

□

اثبات حالت برگشت سر راست است.

لم ۱.۲.۲:

برهان. بدون کاستن از کلیت فرض کنید وجود دارد $0 < m' < 3p' + 3$ به طوری که،

$$(\mathbf{U} + \mathbf{GA})^+ = \left((\mathbf{U} + \mathbf{GA})_1, (\mathbf{U} + \mathbf{GA})_2, \dots, (\mathbf{U} + \mathbf{GA})_{m'}, \underbrace{0, \dots, 0}_{3p'+3-m'} \right)^\top$$

پس داریم:

$$\nabla \phi(\mathbf{y}) = \begin{pmatrix} -(\mathbf{Q} + (\mathbf{G}^*)^\top \mathbf{G}^*) & -(\mathbf{G}^*)^\top \\ \mathbf{G}^* & \mathcal{W}_{(3p'+3) \times (3p'+3)} \end{pmatrix}$$

که در آن

$$\mathbf{G}^* = \begin{pmatrix} \mathbf{R}_{m' \times (3p'+3)} \\ \mathbf{O}_{(3p'+3-m') \times (3p'+3)} \end{pmatrix} = \begin{pmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \\ \dots \\ \dots \\ \mathbf{G}_{m'} \\ \mathbf{O}_{1 \times (3p'+3)} \\ \mathbf{O}_{1 \times (3p'+3)} \\ \dots \\ \dots \\ \mathbf{O}_{1 \times (3p'+3)} \end{pmatrix}$$

$$\mathcal{W}_{(3p'+3) \times (3p'+3)} = \begin{pmatrix} \mathbf{O}_{m' \times m'} & \mathbf{O}_{m' \times (3p'+3-m')} \\ \mathbf{O}_{(3p'+3-m') \times m'} & -I_{(3p'+3-m') \times (3p'+3-m')} \end{pmatrix}$$

و $\mathbf{O}$ نشان دهنده ماتریس صفر است. بنابراین نتیجه می‌شود که $\nabla \phi(\mathbf{y})$ یک ماتریس نیمه معین منفی است.

اگر $m' = 3p' + 3$ ، آن‌گاه $(\mathbf{U} + \mathbf{GA})^+ = \left((\mathbf{U} + \mathbf{GA})_1, (\mathbf{U} + \mathbf{GA})_2, \dots, (\mathbf{U} + \mathbf{GA})_{3p'+3} \right)^\top$ ، بنابراین

$$\nabla \phi(\mathbf{y}) = \begin{pmatrix} -(\mathbf{Q} + \mathbf{G}^\top \mathbf{G}) & -\mathbf{G}^\top \\ \mathbf{G} & \mathbf{O}_{(3p'+3) \times (3p'+3)} \end{pmatrix}$$

مشابه با حالت قبلی، ثابت می‌شود که $\nabla \phi(\mathbf{y})$ نیمه معین منفی است. نهایتاً اگر $m' = 0$ آن‌گاه

$$(\mathbf{U} + \mathbf{GA})^+ = \left(\underbrace{0, 0, \dots, 0}_{3p'+3} \right)^\top$$

بنابراین داریم

$$\nabla \phi(\mathbf{y}) = \begin{pmatrix} -\mathbf{Q} & \mathbf{O}_{(\mathfrak{r}_{p'}+\mathfrak{r}) \times (\mathfrak{r}_{p'}+\mathfrak{r})} \\ \mathbf{O}_{(\mathfrak{r}_{p'}+\mathfrak{r}) \times (\mathfrak{r}_{p'}+\mathfrak{r})} & -I_{(\mathfrak{r}_{p'}+\mathfrak{r}) \times (\mathfrak{r}_{p'}+\mathfrak{r})} \end{pmatrix}$$

در این حالت نیز نتیجه می‌شود که $\nabla \phi(\mathbf{y})$ یک ماتریس نیمه معین منفی است و این اثبات را کامل می‌کند. $\square$

قضیه ۲.۲.۲:

برهان. تابع لیپانوف $E: \mathbb{R}^{\mathfrak{e}_{p'}+\mathfrak{e}} \rightarrow \mathbb{R}$ را به صورت زیر در نظر بگیرید:

$$E(\mathbf{y}) = E_1(\mathbf{y}) + E_2(\mathbf{y}) \quad (۵.ب)$$

که در آن $E_1(\mathbf{y}) = \|\phi(\mathbf{y})\|^2$ و $E_2(\mathbf{y}) = \frac{1}{\mathfrak{r}} \|\mathbf{y} - \mathbf{y}^*\|^2$ از [۶۲] می‌دانیم که $E_1(\mathbf{y})$ یک تابع مشتق‌پذیر است. با توجه به (۴۴.۲)، دیده می‌شود که

$$\frac{d\phi}{dt} = \frac{\partial \phi}{\partial \mathbf{y}} \frac{d\mathbf{y}}{dt} = \nabla \phi(\mathbf{y}) \phi(\mathbf{y})$$

با محاسبه مشتق $E(t)$ در نقطه $\mathbf{y}(t)$ از مدل شبکه عصبی (۱۹.۲) و (۲۰.۲)، داریم:

$$\begin{aligned} \frac{dE(\mathbf{y}(t))}{dt} &= \left(\frac{d\phi}{dt} \right)^\top \phi + \phi^\top \left(\frac{d\phi}{dt} \right) + (\mathbf{y} - \mathbf{y}^*)^\top \frac{d\mathbf{y}(t)}{dt} \\ &= \phi^\top (\nabla \phi(\mathbf{y}))^\top + \nabla \phi(\mathbf{y}) \phi + (\mathbf{y} - \mathbf{y}^*)^\top \phi(\mathbf{y}) \end{aligned} \quad (۶.ب)$$

با استفاده از لم ۱.۲.۲، نتیجه می‌شود که

$$\phi^\top(\mathbf{y})(\nabla \phi(\mathbf{y}))^\top + \nabla \phi(\mathbf{y}) \phi(\mathbf{y}) \leq 0, \quad \forall \mathbf{y} \neq \mathbf{y}^* \quad (۷.ب)$$

علاوه بر این، با توجه به تعریف (۶.۳.۱) و لم (۱.۳.۱) داریم:

$$(\mathbf{y} - \mathbf{y}^*)^\top (\phi(\mathbf{y}) - \phi(\mathbf{y}^*)) = (\mathbf{y} - \mathbf{y}^*)^\top \phi(\mathbf{y}) \leq 0, \quad \forall \mathbf{y} \neq \mathbf{y}^*$$

بنابراین

$$\frac{dE(\mathbf{y}(t))}{dt} \leq 0 \quad (۸.ب)$$

و این به این معنی است که شبکه عصبی (۴۲.۲) و (۴۳.۲) پایدار لیپانوف است. $\square$

لم ۲.۲.۲:

برهان. (i) واضح است که $\{(\mathbf{U} + \mathbf{GA})^+ - \mathbf{U}\}$ و $\{\mathbf{QA} + \mathbf{B} + \mathbf{G}^\top(\mathbf{U} + \mathbf{GA})^+\}$ پیوسته لپیشیتس محلی روی مجموعه محدب باز $D \subseteq \mathbb{R}^{\mathfrak{e}_{p'}+\mathfrak{e}}$ هستند. با توجه به قضیه ۱.۳.۱ شبکه

عصبی (۴۲.۲) و (۴۳.۲) یک جواب پیوسته یکتا $y(t), t \in [t_0, \eta)$ برای بعضی $\eta > t_0$ دارد. از اثبات قضیه ۲.۲.۲ می‌دانیم که E روی t ، یک تابع غیر صعودی است. بنابراین

$$\frac{1}{\gamma} \|y - y^*\|^2 \leq E(y(t)) \leq E(y(t_0)), \quad \forall t \geq t_0 \quad (۹.ب)$$

این نشان می‌دهد که مسیر شبکه عصبی (۴۲.۲) و (۴۳.۲) کراندار است. بنابراین $\eta \rightarrow +\infty$. (ii) برای نقطه ابتدایی داده شده $y(t_0)$ با $U(t_0) \geq 0$ داریم:

$$\begin{aligned} \frac{dU}{dt} + U &= (U + GA)^+ \\ \int_{t_0}^t \left(\frac{dU}{ds} + U \right) e^s ds &= \int_{t_0}^t e^s (U + GA)^+ ds \end{aligned}$$

همچنین

$$U(t) = e^{-(t-t_0)} U(t_0) + e^{-t} \int_{t_0}^t e^s (U + GA)^+ ds$$

در نتیجه $(U + GA)^+ \geq 0$ و $U(t) \geq 0$ برای هر $t \geq t_0$. $\square$

قضیه ۳.۲.۲:

برهان. از اثبات لم ۲.۲.۲، داریم که مسیر $((A(t))^T, U(t)^T)^T$ از شبکه عصبی (۴۲.۲) و (۴۳.۲) کراندار است. بنابراین، یک دنباله صعودی $\{t_k\}$ با $t_k \rightarrow \infty$ که $k \rightarrow \infty$ و نقطه حدی $(\bar{A}^T, \bar{U}^T)^T$ وجود دارد به طوری که:

$$\lim_{k \rightarrow \infty} (A(t_k)^T, U(t_k)^T)^T = (\bar{A}^T, \bar{U}^T)^T$$

با استفاده از قضیه ۵.۳.۱ داریم: $\{(A(t)^T, U(t)^T)^T \rightarrow H\}$ که $t \rightarrow \infty$ و H بزرگ‌ترین مجموعه ثابت در $K = \{(A(t)^T, U(t)^T)^T \mid \frac{dE(y(t))}{dt} = 0\}$ است. از (۴۲.۲) و (۴۳.۲) و (۸.ب)، نتیجه می‌شود که $\frac{dE(y(t))}{dt} = 0 \Leftrightarrow \frac{dU}{dt} = 0$ و $\frac{dA}{dt} = 0$ که $(\hat{A}^T, \hat{U}^T)^T \in D^*$ بنابراین $H \subseteq K \subseteq D^*$. از طرف دیگر، اثبات کردیم که مسیر $(A(t)^T, U(t)^T)^T$ همگرای سراسری به نقطه تعادل $(\bar{A}_0^T, \bar{U}^T)^T$ است. اکنون تابع لیاپانوف دیگری تعریف می‌کنیم:

$$\bar{E}(y) = \|\phi(y)\|^2 + \frac{1}{\gamma} \|y - \bar{y}\|^2 \quad (۱۰.ب)$$

با جایگذاری $A^* = \bar{A}$ و $U^* = \bar{U}$ در (۵.ب) نتیجه می‌شود که $\bar{E}(y)$ به طور پیوسته مشتق پذیر است و $\bar{E}(\bar{y}) = 0$ توجه شود که $\lim_{k \rightarrow \infty} (A(t_k)^T, U(t_k)^T)^T = (\bar{A}^T, \bar{U}^T)^T$ بنابراین داریم $\lim_{k \rightarrow \infty} \bar{E}(A(t_k)^T, U(t_k)^T)^T = \bar{E}(\bar{A}^T, \bar{U}^T)^T$. طوری که برای هر $t \geq t_q$ داریم $\bar{E}(y(t)) < \epsilon$. به طریق مشابه $\frac{d\bar{E}(y(t))}{dt} \leq 0$ در نتیجه برای $t \geq t_q$,

$$\frac{1}{\gamma} \|y(t) - \bar{y}\|^2 \leq \bar{E}(y(t)) \leq \epsilon$$

. همچنین $\lim_{t \rightarrow \infty} y(t) = \bar{y}$ و $\lim_{t \rightarrow \infty} \|y(t) - \bar{y}\| = 0$. بنابراین شبکه (۴۲.۲) و (۴۳.۲) همگرایی سراسری به نقطه تعادل $\hat{y} = (\bar{A}^\top, \bar{U}^\top)^\top$ است. که در آن $\bar{A}$ جواب بهینه مسئله (۳۹.۲) و (۴۰.۲) است.

در حالت خاص، اگر $D^* = \{(A^{*\top}, U^{*\top})^\top\}$ ، بنابراین مسیر $(A(t)^\top, U(t)^\top)^\top$ با هر نقطه ابتدایی $(A(t_0)^\top, U(t_0)^\top)^\top$ پایدار مجانبی سراسری به نقطه تعادل $y^* = (A^{*\top}, U^{*\top})^\top$ است. $\square$

قضیه ۴.۲.۲:

برهان. به برهان قضیه ۹.۳.۱ مراجعه شود. $\square$

ب.۲ فصل ۴

قضیه ۱.۴.۴:

برهان. فرض کنید $y^* = (A^{*\top}, U^{*\top}, U_q^{*\top})^\top$ ، یک نقطه تعادل از (۴۱.۴) و (۴۲.۴) باشد. بنابراین به راحتی نتیجه می‌شود که

$$2\bar{Q}\bar{A} + \bar{B} + \bar{G}^\top(\bar{U} + \bar{G}\bar{A})^+ + \nabla \bar{g}_q(\bar{A})^\top(\bar{U}_q + \bar{g}_q(\bar{A}))^+ = 0 \quad (\text{ب.۱۱})$$

$$(\bar{U} + \bar{G}\bar{A})^+ - \bar{U} = 0 \quad (\text{ب.۱۲})$$

$$(\bar{U}_q + \bar{g}_q(\bar{A}))^+ - \bar{U}_q = 0 \quad (\text{ب.۱۳})$$

با توجه به لم ۲.۳.۱ و با جایگذاری (ب.۱۲) و (ب.۱۳) در (ب.۱۱) داریم:

$$QA^* + B + G^\top U^* + \nabla \bar{g}_q(\bar{A})^\top \bar{U}_q = 0 \quad (\text{ب.۱۴})$$

بنابراین $y^* = (A^{*\top}, U^{*\top}, U_q^{*\top})^\top$ در شرایط کروش-کان-تاگر (۴۰.۴) صدق می‌کند. اثبات حالت برگشت سر راست است. $\square$

لم ۱.۴.۴:

برهان. با جایگذاری

$$G = \begin{pmatrix} \bar{G} \\ \nabla g_q(\bar{A}) \end{pmatrix}$$

و $U = (\bar{U}^\top, \bar{U}_q^\top)^\top$ در اثبات لم ۱.۱.۲ نتیجه به دست می‌آید. $\square$

قضیه ۲.۴.۴:

برهان. اثبات مشابه اثبات ارائه شده برای قضیه ۲.۱.۲ می‌باشد. $\square$

لم ۲.۴.۴:

برهان. (i) از آنجایی که $\{(U + \{2\bar{Q}\bar{A} + \bar{B} + \bar{G}^\top(\bar{U} + \bar{G}\bar{A})^+ + \nabla\bar{g}_q(\bar{A})^\top(\bar{U}_q + \bar{g}_q(\bar{A}))^+ - \bar{U}_q\} + \{(\bar{U}_q + \bar{g}_q(\bar{A}))^+ - \bar{U}_q\} + \{GA\}^+ - U\}$ پیوسته لیپشیتس محلی روی مجموعه محدب باز $D \subseteq \mathbb{R}^{11p+11}$ هستند، با توجه به قضیه ۱.۳.۱ شبکه عصبی (۴۱.۴) و (۴۲.۴) یک جواب پیوسته یکتا $y(t), t \in [t_0, \eta]$ برای بعضی $\eta > t_0$ دارد. از اثبات قضیه ۲.۴.۴ می‌دانیم که E روی t ، یک تابع غیر صعودی است. بنابراین

$$\frac{1}{\varphi} \|y - y^*\|^2 \leq E(y(t)) \leq E(y(t_0)), \quad \forall t \geq t_0 \quad (ب.۱۵)$$

این نشان می‌دهد که مسیر شبکه عصبی (۴۱.۴) و (۴۲.۴) کراندار است. بنابراین $\eta \rightarrow +\infty$. (ii) برای نقطه ابتدایی داده شده $y(t_0) \geq 0$ و $U(t_0) \geq 0$ داریم:

$$\begin{aligned} \frac{dU}{dt} + U &= (U + GA)^+ \\ \int_{t_0}^t \left(\frac{dU}{dt} + U \right) e^s ds &= \int_{t_0}^t e^s (U + GA)^+ ds \end{aligned}$$

همچنین

$$U(t) = e^{-(t-t_0)} U(t_0) + e^{-t} \int_{t_0}^t e^s (U + GA)^+ ds$$

در نتیجه $U(t) \geq 0$ و $(U + GA) \geq 0$ برای هر $t \geq t_0$.

به طریق مشابه داریم: $U_q(t) \geq 0$ و $(U_q + \bar{g}_q(\bar{A})) \geq 0$ برای هر $t \geq t_0$. $\square$

قضیه ۳.۴.۴:

برهان. از اثبات لم ۲.۴.۴، داریم که مسیر

$$((A(t))^\top, U(t)^\top, U_q(t)^\top)^\top$$

از شبکه عصبی (۴۱.۴) و (۴۲.۴) کراندار است. بنابراین، یک دنباله صعودی $\{t_k\}$ با $t_k \rightarrow \infty$ که $k \rightarrow \infty$ و نقطه حدی $(\bar{A}^\top, \bar{U}^\top, \bar{U}_q^\top)^\top$ وجود دارد به طوری که:

$$\lim_{k \rightarrow \infty} (A(t_k)^\top, U(t_k)^\top, U_q(t_k)^\top)^\top = (\bar{A}^\top, \bar{U}^\top, \bar{U}_q^\top)^\top$$

با استفاده از قضیه مجموعه ثابت لاسال [۵۹]، داریم: $\{(A(t)^\top, U(t)^\top, U_q(t)^\top)^\top \rightarrow H\}$ که $H = \{(A(t)^\top, U(t)^\top, U_q(t)^\top)^\top \mid \frac{dE(y(t))}{dt} = 0\}$ و $t \rightarrow \infty$ ، بزرگ‌ترین مجموعه ثابت در H است. از (۴۱.۴) و (۴۲.۴) نتیجه می‌شود که $\frac{dA}{dt} = 0$ و $\frac{dU}{dt} = 0 \Leftrightarrow \frac{dE(y(t))}{dt} = 0$ بنابراین $H \subseteq K \subseteq D^*$ که $(\bar{A}^\top, \bar{U}^\top, \bar{U}_q^\top)^\top \in D^*$.

اکنون اثبات می‌کنیم که مسیر $(A(t)^\top, U(t)^\top, U_q(t)^\top)^\top$ همگرای سراسری به نقطه تعادل $(\bar{A}^\top, \bar{U}^\top, \bar{U}_q^\top)^\top$ است. اکنون تابع لیاپانوف دیگری تعریف می‌کنیم:

$$\bar{E}(y) = \|\phi(y)\|^2 + \frac{1}{\varphi} \|y - \bar{y}\|^2 \quad (ب.۱۶)$$

با جایگذاری $\mathbf{U}^* = \bar{\mathbf{U}}$ ، $\mathbf{A}^* = \bar{\mathbf{A}}$ و $\mathbf{U}_q^* = \bar{\mathbf{U}}_q$ در (ب.۱۶). نتیجه می‌شود که $\bar{E}(\mathbf{y})$ به طور پیوسته مشتق‌پذیر است و $\bar{E}(\bar{\mathbf{y}}) = 0$ توجه شود که

$$\lim_{k \rightarrow \infty} (\mathbf{A}(t_k)^\top, \mathbf{U}(t_k)^\top, \mathbf{U}_q(t_k)^\top)^\top = (\bar{\mathbf{A}}^\top, \bar{\mathbf{U}}^\top, \bar{\mathbf{U}}_q^\top)^\top$$

بنابراین داریم

$$\lim_{k \rightarrow \infty} \bar{E}(\mathbf{A}(t_k)^\top, \mathbf{U}(t_k)^\top, \mathbf{U}_q(t_k)^\top)^\top = \bar{E}(\bar{\mathbf{A}}^\top, \bar{\mathbf{U}}^\top, \bar{\mathbf{U}}_q^\top)^\top$$

پس برای هر $\epsilon > 0$ وجود دارد $s > 0$ به طوری که برای هر $t \geq t_s$ داریم $\bar{E}(\mathbf{y}(t)) < \epsilon$ به طریق مشابه $\frac{d\bar{E}(\mathbf{y}(t))}{dt} \leq 0$ در نتیجه برای $t \geq t_s$

$$\frac{1}{\tau} \|\mathbf{y}(t) - \bar{\mathbf{y}}\|^2 \leq \bar{E}(\mathbf{y}(t)) \leq \epsilon$$

همچنین $\lim_{t \rightarrow \infty} \mathbf{y}(t) = \bar{\mathbf{y}}$ و $\lim_{t \rightarrow \infty} \|\mathbf{y}(t) - \bar{\mathbf{y}}\| = 0$ بنابراین شبکه (۴۱.۴) و (۴۲.۴) همگرایی سراسری به نقطه تعادل $\bar{\mathbf{y}} = (\bar{\mathbf{A}}^\top, \bar{\mathbf{U}}^\top, \bar{\mathbf{U}}_q^\top)^\top$ است. که در آن $\bar{\mathbf{A}}$ جواب بهینه مسئله (۱۴.۴)–(۱۶.۴) است. در حالت خاص، اگر $D^* = \{(\mathbf{A}^{*\top}, \mathbf{U}^{*\top}, \mathbf{U}_q^{*\top})^\top\}$ ، بنابراین مسیر $(\mathbf{A}(t)^\top, \mathbf{U}(t)^\top, \mathbf{U}_q(t)^\top)^\top$ با هر نقطه ابتدایی $(\mathbf{A}(t_0)^\top, \mathbf{U}(t_0)^\top, \mathbf{U}_q(t_0)^\top)^\top$ پایدار مجانبی سراسری به نقطه تعادل $\mathbf{y}^* = (\mathbf{A}^{*\top}, \mathbf{U}^{*\top}, \mathbf{U}_q^{*\top})^\top$ است. □

قضیه ۴.۴.۴:

□

برهان. مشابه با برهان قضیه ۹.۳.۱ می‌باشد.

واژه‌نامه فارسی به انگلیسی

Initial	ابتدایی
Extention Principle	اصل گسترش
Cross-Validation	اعتبار سنجی متقابل
Axon	آکسون
h -cut	h - برش
Optimization	بهینه سازی
Constrained Optimization	بهینه سازی مقید
Stability	پایداری
Stability Asymptotic	پایداری مجانبی
Left Spread	پهنای چپ
Right Spread	پهنای راست
Predictor	پیش بین
Continuous	پیوسته
Activation Function	تابع تحریک
Errore Function	تابع خطا
Membership Function	تابع عضویت
Liapunov Function	تابع لیاپانف
Convex Function	تابع محدب
S-shap Function	تابع S -شکل
Objective Function	تابع هدف
Optimal Solution	جواب بهینه
Estimation Error	خطای به دست آمده
Crisp	دقیق
Dendrit	دندریت
Sequence	دنباله

Polynomial Regression	رگرسیون چندجمله‌ای
Linear Regression	رگرسیون خطی
Multiple Linear Regression	رگرسیون خطی چندگانه
Dynamical Systems	سیستم‌های دینامیکی
Biological Neural Network	شبکه عصبی زیستی
Artificial Neural Network	شبکه عصبی مصنوعی
Lagrangian Multipliers	ضرایب لاگرانژ
Variance Inflation Factor	عامل تورم واریانس
Fuzzy Number	عدد فازی
LR Fuzzy Number	عدد فازی LR
Triangular Fuzzy Number	عدد فازی مثلثی
Normal Fuzzy Number	عدد فازی نرمال
Nonlinear	غیرخطی
Bounded	کراندار
Least Square Error	کمترین توان‌های دوم خطا
Output Variable	متغیر خروجی
Input Variable	متغیر ورودی
Differentiable	مشتق پذیر
Positive Definite	معین مثبت
Locally	محلی
Equilibrium Point	نقطه تعادل
Goodness of Fit	نیکویی برازش
Positive Semi-definite	نیمه معین مثبت
Collinearity	هم‌خطی
Convergence	همگرایی

واژه‌نامه انگلیسی به فارسی

Activation Function	تابع تحریک
Artificial Neural Network	شبکه‌های عصبی مصنوعی
Axon	آکسون
Biological Neural Network	شبکه عصبی زیستی
Bounded	کراندار
Collinearity	هم‌خطی
Constrained Optimization	بهبودسازی مقید
Continuous	پیوسته
Convergence	همگرایی
Convex Function	تابع محدب
Crisp	دقیق
Cross-Validation	اعتبار سنجی متقابل
Dendrit	دندریت
Differentiable	مشتق‌پذیر
Dynamical Systems	سیستم‌های دینامیکی
Equilibrium Point	نقطه تعادل
Errore Function	تابع خطا
Estimation Error	خطای به‌دست آمده
Extention Principle	اصل گسترش
Fuzzy Numbers	اعداد فازی
Goodness of Fit	نیکویی برازش
h-cut	h - برش
Initial	ابتدایی
Input Variable	متغیر ورودی
Lagrangian Multipliers	ضرایب لاگرانژ

Least Square Error	کمترین توان‌های دوم خطا
Left Spread	پهنای چپ
Liapunov Function	تابع لیاپانوف
Linear Regression	رگرسیون خطی
Locally	محلی
LR Fuzzy Numbers	اعداد فازی LR
Membership Function	تابع عضویت
Multiple Linear Regression	رگرسیون خطی چندگانه
Nonlinear	غیرخطی
Normal Fuzzy Number	عدد فازی نرمال
Objective Function	تابع هدف
Optimization	بهینه‌سازی
Output Variable	متغیر خروجی
Polynomial Regression	رگرسیون چندجمله‌ای
Positive Definite	معین مثبت
Positive Semi-definite	نیمه معین مثبت
Predictor	پیش‌بین
Right Spread	پهنای راست
S-shap Function	تابع S-شکل
Stability Asymptotic	پایداری مجانبی
Stability	پایداری
Sequence	دنباله
Trajectory	مسیر
Variance Inflation Factor	عامل تورم واریانس

Aabstract

In this thesis, a hybrid scheme based on the recurrent neural networks for approximate fuzzy coefficients (parameters) of fuzzy linear and polynomial regression models is presented. When multicollinearity happens (As in regression models in traditional statistical theory), the model may lead to poor prediction. To overcome the problem, Bridge regression methodology which combines Bridge regression with the fuzzy regression models in order to reduce the effect of multicollinearity, for the first time is considered in this thesis. Then, the proposed model and its specific formulations (Ridge and Lasso regression) are solved by using a capable neural network. The proposed neural network is first constructed based on some concepts of convex optimization and stability theory. The presented neural network framework guarantees to find the approximate parameters of the fuzzy regression models. The existence and convergence of the trajectories of the neural network are studied. The Lyapunov stability for the neural network is also shown.

Keywords: Fuzzy linear regression; Fuzzy polynomial regression; Fuzzy Bridge regression; Recurrent neural network; Stability; Convergence



Faculty Of Mathematical Sciences

PhD Thesis in Operation Research

A Dynamic Optimization Scheme for Solving Fuzzy Regression Models

By: Delara Karbasi

Supervisors:

Dr. Alireza Nazemi
Dr. Mohammad Reza Rabiei

February 2021