برآورد میزان تولید پروتئین A نوترکیب خارج سلولی از باکتری Escherichia coli BL21 با استفاده از روش رگرسیون کمترین مربعات فازی هیبریدی

رحام آرمند[۱]، گرشاسب ریگی[۲،۳*]، طاهره بهرامی[۴]

۱) گروه زیست شناسی، دانشکده علوم پایه، دانشگاه صنعتی خاتم الانبیاء(ص) بهبهان، بهبهان، ایران

۲) گروه ژنتیک، دانشکده علوم پایه، دانشگاه شهرکرد، شهرکرد، ایران

۳) گروه بیوتکنولوژی صنعتی، پژوهشکده بیوتکنولوژی، دانشگاه شهرکرد، شهرکرد، ایران

۴) گروه آمار، دانشکده علوم پایه، دانشگاه صنعتی خاتم الانبیاء(ص) بهبهان، بهبهان، ایران

تاریخ دریافت: ۹۶/۸/۱۴ **تاریخ پذیرش: ۹۸/۳/۲۸**

چکیده

مقدمه: ایمیون پروتئین A یک مولفه با طیف وسیعی از کاربردهای بیوشمیایی، بیوتکنولوژی و پزشکی است. ژن کدکننده این پروتئین از باکتری استافیلوکوکوس اورئوس استخراج و در باکتری اشریشیاکلی کلون و بیان شد. جهت بهینه سازی شرایط بیان به منظور سنجش دقت آزمایش و تضمین صحت کار در آزمایشات بعد، کاهش هزینه و جلوگیری از روش های آزمون و خطا و دستیابی به بیشترین میزان تولید، از روش های آماری مناسب استفاده می شود.

مواد و روش ها: روش رگرسیون آماری معمولی بر پایه فرض دقیق بودن متغیرهای مورد مطالعه و مشاهدات مربوط به آن ها استوار بوده و در نهایت روابط بین متغیرها نیز به طور دقیق مشخص می شود. در مدل سازی طبیعی مانند برآورد میزان پروتئین، عموماً ممکن است با مشاهدات نادقیق و با روابط مبهم روبرو شویم، بنا بر این استفاده از روش های رگرسیون که قادر به تبیین ساختار مبهم میزان پروتئین و در اختیار نهادن الگوهایی که منطبق با واقعیت هستند، ضروری می باشد. در این مقاله از روش رگرسیون فازی هیبریدی با کمترین مربعات که مبتنی بر نظریه مجموعه های فازی می باشد استفاده شد.

یافته های پژوهش: مقدار برآورد شده پروتئین A با اطمینان ۹۰ درصد صحیح محاسبه شده و از طرفی با توجه به این که روش مورد استفاده در این مقاله روش رگرسیون فازی هیبریدی با کمترین مربعات خطا(برآورد با کمترین مربعات خطا) می باشد لذا اگر تمامی داده های مورد استفاده در این مقاله اعداد تردد(غیر فازی) باشند نتایج حاصل از رگرسیون فازی هیبریدی مشابه رگرسیون معمولی است.

بحث و نتیجه گیری: یکی از مزایای رگرسیون فازی هیبریدی این است که نسبت به برآورد نقطه ای دارای سطح اطمینان بالاتری می باشد. در این مطالعه انحراف معیار داده های برآورد شده به روش رگرسیون فازی هیبریدی کمتر از روش رگرسیون معمولی بود که می توان نتیجه گرفت که روش رگرسیون فازی هیبریدی روش بهینه ای جهت برآورد میزان تولید این نوع پروتئین است.

واژه های کلیدی: پروتئین A، پروتئین متصل شونده به ایمونوگلوبین، رگرسیون فازی هیبریدی، بیوتکنولوژی پزشکی

* نویسنده مسئول: گروه ژنتیک، دانشکده علوم پایه، دانشگاه شهرکرد، شهرکرد، ایران-گروه بیوتکنولوژی صنعتی، پژوهشکده بیوتکنولوژی، دانشگاه شهرکرد، شهرکرد، ایران

Email:

Copyright © 2019 Journal of Ilam University of Medical Science. This is an open-access article distributed under the terms of the Creative Commons Attribution international 4.0 International License (https://creativecommons.org/licenses/by-nc/4.0/) which permits copy and redistribute the material, in any medium or format, provided the original work is properly cited.

[Downloaded from sjimu.medilam.ac.ir on 2026-10-10]

[DOI: 10.29252/sjimu.27.3.1]

مقدمه

پروتئین A اولین بار توسط جانسن به عنوان آنتی ژنی معرفی شد که آنتی بادی آن به صورت طبیعی در سرم همه انسان ها وجود دارد. با این وجود به دلیل مقدار زیاد IgG در سرم انسان که قادر است به این پروتئین متصل شود، دانشمندان متوجه شدند که این برهم کنش غیر اختصاصی بوده و از نوع برهم کنش آنتی ژن-آنتی بادی نیست(۱،۲).

ساختارهای NMR از دمین های B، E و Z (دمین B جهش یافته) نشان می دهد که مدول های منفرد از ساختارهای باندلی سه مارپیچی تشکیل شده اند. باقی مانده های دو مارپیچ اول برای اتصال به Fc (قطعه با قابلیت کریستاله شدن) آنتی بادی ها ضروری هستند و نشان داده شده است که با Fc در دمین B ساختار کریستالی تماس پیدا می کند. علاوه بر این یک نوع دمین پروتئین A کوچک شده شامل تنها دو مارپیچ اول است که نشان می دهد این ناحیه نه تنها ضروری است بلکه برای تمایل بالای اتصال به Fc نیز کافی می باشد(۷-۳).

پروتئین A واکنشگر متصل شونده به آنتی بادی می باشد که به FC همه آنتی بادی های انسانی(به جزء ۳ IgG) و برخی انواع آنتی بادی های دیگر گونه های حیوانی مثل خوک، سگ، خرگوش، بز و موش متصل می شود(۱۲-۸).

پروتئین A یک مولفه با طیف وسیعی از کاربردهای بیوشمیایی، بیوتکنولوژی و پزشکی است. پروتئین A توانایی اتصال به نگهدارنده هایی مثل سیلیکا، آگارز، سلولز یا کراس لینک دکستران را جهت تهیه رزین های کروماتوگرافی تمایلی برای تخلیص آنتی بادی ها، پروتئین ها، پپتیدها و غیره را ندارد.

پروتئین A معمول ترین مولفه ای است که جهت روش های آزمونی مثل تجمع، ELISA و وسترن بلات به کار می رود(۱۲-۱۰).

در مطالعات قبلی، پروتئین A نوترکیب در شرایط آزمایشگاهی در میزبان باکتریایی Escherichia coli سویه بیانی BL21 به صورت خارج سلولی تولید شد و تعدادی شرایط آزمایشگاهی نیز جهت افزایش میزان این نوع تولید و خالص سازی آن انجام شد(۱۶-۱۳). از طرفی این پروتئین روی سطح کیتوزان جهت بهره برداری به عنوان خالص سازی ایمونوگلوبولین G تثبیت شد(۱۷) و سیگنال پپتید طبیعی و تغییر یافته این پروتئین برای ترشحی نمودن نوعی لیپاز در باکتری اشریشیاکلی مورد استفاده قرار گرفت(۱۸). جهت سنجش دقت آزمایش فعلی و تضمین صحت کار در طراحی آزمایشگاهی این پروتئین مهم و دستیابی به بهترین و بهینه ترین شرایط تولید و ترشح، باید ابتدا به روش آماری میزان تولید آن محاسبه شود. در یک مطالعه آماری با استفاده از روش رگرسیون فازی این محاسبات انجام شد(۱۹) و در این مقاله روش دیگری جهت برآورد میزان تولید پروتئین خارج سلولی ارائه می شود که به روش رگرسیون فازی هیبریدی با کمترین مربعات موسوم می باشد. دارای مزایایی نست به روش قبلی است که در ادامه تشریح می شود.

یکی از جنبه های مهم برآورد میزان تولید پروتئین، ایجاد همبستگی بین میزان دما، غلظت گلایسین، غلظت لاکتوز، مدت زمان القاء و تراکم سلولی باکتری به عنوان فاکتورها و شرایط و بیان میزان کمی آن ها در قالب های آماری است، که عمدتاً شامل مدل های رگرسیون آماری می باشد. توسط این مدل ها میزان هر یک از مواد مورد نیاز جهت تولید پروتئین که هم زمان بر است و هم هزینه بردار، به صورت تابعی از موارد مورد نیاز جهت تولید پروتئین A تعیین می شود. اخیراً تلاش هایی به منظور به کارگیری روش ها و فنون جدید در مدل سازی چنین توابعی صورت گرفته است که می توان برای نمونه به استفاده از شبکه های عصبی مصنوعی اشاره کرد(۲۰).

تمامی این روش ها بر پایه فرض دقیق بودن متغیرهای مورد مطالعه و مشاهدات مربوط به آن ها، و البته دقیق بودن روابط بین متغیرها فرض می شود. حال آن که در مواردی از قبیل تولید پروتئین، عموماً با مشاهدات یا روابط نادقیق بین متغیرها روبرو هستیم. و لذا در صورتی که هر یک از شرایط رگرسیون معمولی برقرار نباشد، برآورد با این روش با مشکل روبرو می شود و در حقیقت برآوردی نادرست ارائه می کند.

در چنین شرایطی از مدل هایی باید بهره گرفت که قادر به ارائه الگوی مناسبی هستند به گونه ای که این

[Downloaded from sjimu.medilam.ac.ir on 2026-10-10]

[DOI: 10.29252/sjimu.27.3.1]

مدل ها بتواند انطباق بیشتری با دنیای واقعی داشته باشد. جهت غلبه بر این مشکل می توان از نظریه مجموعه های فازی جهت برآورد میزان تولید پروتئین A استفاده نمود(۲۰). کاربردهای مدل رگرسیون فازی از سال ۱۹۸۲ با معرفی مدل های رگرسیونی و داده های فازی که بر پایه نوعی برنامه ریزی خطی می باشند، گسترش پیدا کرده اند(۲۱).

در روش های رگرسیون معمولی، اختلاف بین مقادیر مشاهده شده متغیر خروجی یا پاسخ ایجاد شده و مقادیر برآورد شده آن به عنوان خطا در نظر می گیریم در ساده ترین حالت، این خطا می توانند به عنوان متغیرهای تصادفی نرمال با میانگین صفر و واریانس ثابت فرض شوند اما در برخی از معادلات به عنوان مثال وقتی مشاهدات یا عباراتی مانند کوتاه، بزرگ و تقریباً بررسی می شوند، مشاهدات فازی بوده و رابطه بین متغیر پاسخ و متغیرهای ورودی یا پیش بینی شده، دقیق نیستند بنا بر این، خصوصیت این روش رگرسیونی آن است که برآورد که از این طریق حاصل می شود یک برآورد فاصله ای است که از لحاظ سطح اطمینان، دارای سطح اطمینان بسیار بالا است(۲۲).

ساختار مقاله بدین شرح است که ابتدا مدل رگرسیون فازی هیبریدی را شرح داده سپس کاربرد عددی این مدل در برآورد میزان ترشح پروتئین A را که از یک نمونه آزمایشگاهی حاصل شده اند(۱۵) تخمین زده می شود و در نهایت مهم ترین نتایج مقاله ذکر می شود. هدف از مقاله حاضر معرفی رگرسیون فازی هیبریدی و نشان دادن قابلیت های کاربردی آن در بسط و توسعه مدل های خطی(رگرسیون) به منظور برآورد میزان تولید پروتئین A است.

مواد و روش ها

رگرسیون فازی

بررسی رگرسیون فازی هیبریدی با کمترین مربعات خطا: رگرسیون فازی یک روش ناپارامتریک می باشد که جهت برآورد رابطه بین متغیرها در مواردی که داده ها محدود، مبهم و نامعین می باشند می تواند مورد استفاده قرار گیرد(۲۲).

تصادفی و فازی بودن دو گونه متفاوت از نامعین بودن می باشد. که در اولی نقاطی از فضای نمونه مدنظر است که می تواند اتفاق بیفتد. یعنی در حقیقت در تصادفی بودن با تابع نشانگر مواجه هستیم. اگر نقطه متعلق به فضای نمونه باشد مقدار تابع مشخصه یک بوده و پیشامد اتفاق می افتد و اگر نقطه متعلق به فضا نباشد مقدار تابع مشخصه صفر بوده و پیشامد اتفاق نمی افتد(۲۳،۲۴).

اما راجع به فازی بودن تابع عضویت که به صورت نامعین از بازه(۱ و ۰) انتخاب می شود جایگزین تابع مشخصه می شود. چانگ مفهوم رگرسیون فازی هیبریدی را به خوبی توصیف کرده است(۲۵).

مانند هر تکنیک رگرسیون، هدف از رگرسیون فازی تعیین رابطه تابعی بین متغیر وابسته و مجموعه متغیرهای مستقل است، که می تواند هر دو عدد ترد (غیر فازی) باشد یا متغیر مستقل ترد(غیر فازی) بوده و متغیر وابسته عدد فازی باشد. روش غیر پارامتری رگرسیون فازی در برآورد روابط میان متغیرها می تواند بسیار مفید باشد.

به عبارتی دیگر می توان گفت در رگرسیون نقطه ای یک نقطه به عنوان نقطه برآورد برای متغیر وابسته معرفی می شود اما در رگرسیون فازی به جای یک نقطه ما با بازه ای از اعداد مواجه هستیم. یعنی هر نقطه برآورد شده در رگرسیون فازی معادل با یک بازه در رگرسیون فازی هست که کران پایین بازه را اصطلاحاً نیم پهنای چپ و کران بالای بازه را نیم پهنای راست عدد فازی می نامند و در حقیقت مزیت عدد فازی نسبت به عدد غیر فازی در برآوردها همین نکته است و به بیان ساده تر احتمال این که عددی در یک بازه برآورد قرار بگیرد بیشتر از زمانی است که پیش بینی کنیم آن عدد برابر با مقدار مشخصی شود.

مدل رگرسیون فازی هیبریدی حل دستگاه های معادلات مربوطه(۲۵) در قسمت ضمیمه ارائه شده است. نوع مطالعه حاضر، یک پژوهش در زمینه آمار زیستی جهت برآورد میزان ترشح پروتئین مهم از نظر اقتصادی، دارویی و پزشکی به روش رگرسیون با تعداد نمونه ۲۶ موردی است.

یافته های پژوهش

کاربردی عددی رگرسیون فازی هیبریدی در برآورد میزان ترشح پروتئین A: در این بخش از مقاله میزان

[Downloaded from sjimu.medilam.ac.ir on 2026-10-10]

[DOI: 10.29252/sjimu.27.3.1]

ترشح پروتئین A را که تحت فاکتورهای میزان دما، غلظت گلایسین، غلظت لاکتوز، مدت زمان القاء و تراکم سلولی باکتری در شرایط آزمایشگاهی در یک نمونه ۲۶ تایی حاصل شد(۱۵) ملاحظه می شود (جدول شماره ۱). جدول شماره ۱ نشان دهنده تجزیه و تحلیل اثرات این پنج متغیر بر عملکرد بیان و بهره وری پروتئین A نوترکیب محلول و بالغ خارج سلولی متشکل از تمام ترکیبات از پنج متغیر است که با استفاده از طرح آماری Central Composite Design (CCD) و نرم افزار Design-Expert v. 8.0.6, Stat-Ease, Inc., Minneapolis, MN, USA انجام شد(جدول شماره ۱). در ادامه با استفاده از رگرسیون فازی هیبریدی به برآورد میزان تولید خارج سلولی این پروتئین پرداخته می شود.

جدول شماره ۱. متغیرهای نشان دهنده مقادیر تولید پروتئین A نوترکیب خارج سلولی*

میزان تولید(میلی گرم/میلی لیتر)	دمای القاء(ساعت)	تراکم سلول(OD.600)	غلظت گلاسین(٪)	غلظت لاکتوز(٪)	دما(°C)	دور
۲۵/۳۴۲	۱۱	۱/۳	۱/۲۳	۱۲/۷۵	۳۰	۱
۱۳/۳۵۱	۱۱	۱/۳	۱/۲۳	۷/۲۵	۳۵	۲
۱۹/۱۷۳	۱۱	۰/۸	۱/۲۳	۱۲/۷۵	۳۵	۳
۲۸/۵۹۴	۱۰	۱/۳	۱/۲۳	۱۲/۷۵	۳۵	۴
۲۸/۶۴۷	۱۰	۰/۸	۱/۷۷	۱۲/۷۵	۳۵	۵
۲۵/۹۰۲	۱۱	۱/۱	۱/۵۰	۱۰/۰۰	۳۳	۶
۳۴/۲۷۰	۱۰	۱/۳	۱/۷۷	۱۲/۷۵	۳۰	۷
۵۰/۸۴۶	۱۱	۱/۱	۱/۵۰	۱۰/۰۰	۲۸	۸
۲۵/۰۸۱	۱۱	۱/۱	۱/۵۰	۱۰/۰۰	۳۳	۹
۱۶/۳۱۴	۱۲	۱/۱	۱/۵۰	۱۰/۰۰	۳۳	۱۰
۶۱/۴۱۸	۱۰	۱/۳	۱/۷۷	۷/۲۵	۳۵	۱۱
۵۲/۳۳۹	۹	۱/۱	۱/۵۰	۱۰/۰۰	۳۳	۱۲
۳۸/۱۳۷	۱۱	۱/۳	۱/۷۷	۷/۲۵	۳۰	۱۳
۳۶/۲۷۴	۱۱	۱/۱	۱/۵۰	۱۵/۰۰	۳۳	۱۴
۱۹/۰۰۶	۱۰	۰/۸	۱/۲۳	۷/۲۵	۳۰	۱۵
۴۲/۱۱۷	۱۱	۱/۱	۲/۰۰	۱۰/۰۰	۳۳	۱۶
۳۰/۰۰۸	۱۱	۱/۵	۱/۵۰	۱۰/۰۰	۳۳	۱۷
۲۱/۲۲۵	۱۱	۱/۱	۱/۵۰	۵/۰۰	۳۳	۱۸
۲۳/۵۰۳	۱۱	۱/۱	۱/۵۰	۱۰/۰۰	۳۳	۱۹
۳۶/۹۴۸	۱۱	۰/۸	۱/۷۷	۷/۲۵	۳۳	۲۰
۳۱/۲۱۷	۱۱	۱/۱	۱/۵۰	۱۰/۰۰	۳۳	۲۱
۲۰/۹۸۹	۱۱	۱/۱	۱/۵۰	۱۰/۰۰	۳۷	۲۲
۱۵/۴۳۹	۱۱	۱/۱	۱/۰۰	۱۰/۰۰	۳۳	۲۳
۲۵/۰۸۰	۱۱	۱/۱	۱/۵۰	۱۰/۰۰	۳۳	۲۴
۱۹/۶۴۱	۱۱	۰/۶	۱/۵۰	۱۰/۰۰	۳۳	۲۵
۳۰/۸۲۹	۱۱	۰/۸	۱/۷۷	۱۲/۷۵	۳۰	۲۶

* تجزیه و تحلیل اثرات پنج متغیر(غلظت لاکتوز، غلظت گلایسین و OD باکتریایی در زمان القا، دما و مدت زمان پس از القاء) بر عملکرد بیان و بهره وری پروتئین A نوترکیب محلول و بالغ خارج سلولی. جهت اطلاع بیشتر لطفا به مقاله ریگی و همکاران ۲۰۱۳ مراجعه شود(۱۵).

محاسبه رگرسیون فازی هیبریدی برای داده های فوق با استفاده از برنامه نویسی متلب منجر به برآوردهای موجود در جدول شماره ۲ شد. برای داده های آزمایشگاهی برآورد شده در جدول شماره ۱، با استفاده از برنامه نویسی متلب، رگرسیون فازی هیبریدی محاسبه شد که برآوردهای مربوطه به این روش در جدول شماره ۲ مشاهده می شود. ضمناً یک عدد فازی تعمیمی از اعداد معمولی و حقیقی است به صورتی که یک عدد فازی نه به یک مقدار بلکه به مجموعه همبندی از مقادیر ممکن اشاره می‌کند و عدد فازی مثلثی، یک عدد فازی است که با سه عدد حقیقی شامل مرکز عدد، گستره چپ و گستره راست نمایش داده شد.

[Downloaded from sjimu.medilam.ac.ir on 2026-10-10]

[DOI: 10.29252/sjimu.27.3.1]

جدول شماره ۲. محاسبه رگرسیون فازی هیبریدی برای داده های آزمایشگاهی*

پروتئین A	میزان پروتئین خارج سلولی		گستره راست عدد فازی مثلثی
	مرکز عدد فازی مثلثی	گستره چپ عدد فازی مثلثی	
y_1	۲۶/۹۴	۲۲/۸۷	۳۲/۶۹
y_2	۲۷/۰۲	۲۲/۵۳	۳۱/۹۰
y_3	۲۱/۱۷	۱۵/۲۸	۲۸/۶۶
y_4	۲۶/۹۷	۲۲/۶۴	۳۲/۹۳
y_5	۲۳/۳۰	۱۴/۴۸	۳۴/۴۳
y_6	۲۵/۷۶	۱۹/۳۹	۳۳/۳۴
y_7	۲۹/۰۷	۲۲/۰۷	۳۸/۴۶
y_8	۲۵/۵۶	۱۹/۳۳	۳۳/۸۶
y_9	۲۵/۷۶	۱۹/۳۹	۳۳/۳۴
y_{10}	۲۵/۹۴	۱۹/۶۷	۳۲/۵۸
y_{11}	۲۹/۱۵	۲۱/۷۳	۳۷/۶۷
y_{12}	۲۵/۴۰	۱۸/۸۲	۳۴/۸۶
y_{13}	۲۹/۱۳	۲۱/۹۶	۳۷/۴۳
y_{14}	۲۵/۸۸	۱۹/۷۵	۳۳/۵۹
y_{15}	۲۰/۶۵	۱۴/۵۴	۲۹/۶۶
y_{16}	۲۷/۹۰	۱۸/۹۱	۳۷/۹۸
y_{17}	۳۰/۵۵	۲۵/۵۰	۳۶/۱۵
y_{18}	۲۵/۶۵	۱۹/۰۳	۳۳/۰۹
y_{19}	۲۵/۷۶	۱۹/۳۹	۳۳/۰۴
y_{20}	۲۳/۳۵	۱۴/۳۷	۳۳/۳۹
y_{21}	۲۵/۷۶	۱۹/۳۹	۳۳/۳۴
y_{22}	۲۵/۹۳	۱۹/۴۳	۳۲/۹۳
y_{23}	۲۳/۶۲	۱۹/۸۶	۲۸/۷۰
y_{24}	۲۵/۷۶	۱۹/۳۹	۳۳/۳۴
y_{25}	۱۹/۷۸	۱۱/۷۴	۲۹/۸۳
y_{26}	۲۳/۲۷	۱۴/۷۱	۳۴/۱۹

* با محاسبه رگرسیون فازی هیبریدی برای داده های ذکر شده در جدول شماره ۱ با استفاده از برنامه نویسی متلب، برآوردهای موجود در جدول شماره ۲ به دست آمد.

نتایج و برآوردهای ارائه شده در جدول شماره ۲، برآورد فاصله اطمینان در سطح ۹۰ در صد می باشد که این برآورد با این سطح اطمینان برآورد مناسبی نسبت به برآورد نقطه ای است.

میزان تولید واقعی حاصل از آزمایشگاه، مقادیر رگرسیون نقطه ای که با استفاده از نرم افـــزار SPSS محاسبه شده است و رگرسیون فازی هیبریدی حاصل از جدول شماره ۲ به صورت یکجا جهت مقایسه نمایش داده شده اند(جدول شماره ۳). با مشاهده جدول شماره ۳، مقادیر رگرسیون نقـــطه ای و رگرسیون فازی هیبریدی(مراکز برآورد فازی) را می توان مقایسه کرد.

[Downloaded from sjimu.medilam.ac.ir on 2026-10-10]

[DOI: 10.29252/sjimu.27.3.1]

جدول شماره ۳ مقایسه دو روش رگرسیون نقطه ای و فازی هیبریدی به صورت یکجا*

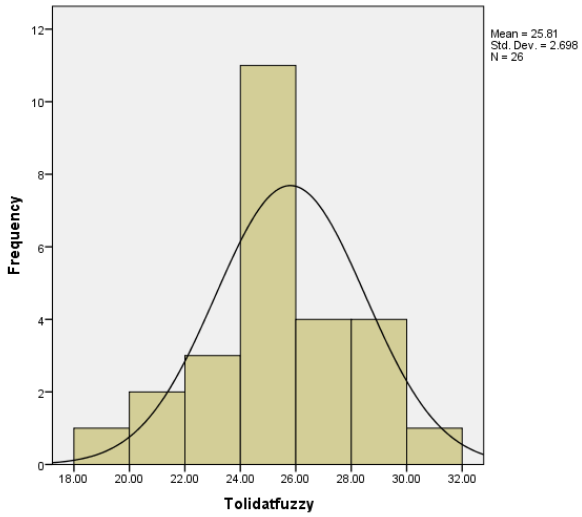
دور آزمایش	میزان تولید واقعی	برآورد نقطه ای	برآورد فازی
۱	۲۵/۳۴۲	۲۶/۹۴	۱۹/۲۶
۲	۱۳/۳۵۱	۲۷/۰۲	۱۳/۱۱
۳	۱۹/۱۷۳	۲۱/۱۷	۶/۱۸
۴	۲۸/۵۹۴	۲۶/۹۷	۲۱/۹۹
۵	۲۸/۶۴۷	۲۳/۳۰	۳۰/۴۰
۶	۲۵/۹۰۲	۲۵/۷۶	۲۰/۴۷
۷	۳۴/۲۷۰	۲۹/۰۷	۴۳/۴۸
۸	۵۰/۸۴۶	۲۵/۵۶	۲۶/۶۲
۹	۲۵/۰۸۱	۲۵/۷۶	۲۰/۴۷
۱۰	۱۶/۳۱۴	۲۵/۹۴	۱۱/۶۰
۱۱	۶۱/۴۱۸	۲۹/۱۵	۳۷/۳۳
۱۲	۵۲/۳۳۹	۲۵/۴۰	۳۸/۲۲
۱۳	۳۸/۱۳۷	۲۹/۱۳	۳۴/۶۱
۱۴	۳۶/۲۷۴	۲۹/۷۷	۲۰/۴۷
۱۵	۱۹/۰۰۶	۲۰/۶۵	۲۱/۲۰
۱۶	۴۲/۱۱۷	۲۷/۹۰	۳۴/۶۸
۱۷	۳۰/۰۰۸	۳۰/۵۵	۲۶/۰۲
۱۸	۲۱/۲۲۵	۲۵/۶۵	۲۰/۴۷
۱۹	۲۱/۵۰۳	۲۵/۷۶	۲۰/۴۷
۲۰	۳۶/۹۴۸	۲۵/۳۸	۲۳/۹۹
۲۱	۳۱/۲۱۷	۲۵/۷۶	۲۰/۴۷
۲۲	۲۰/۹۸۹	۲۵/۹۳	۱۵/۵۵
۲۳	۱۵/۴۳۹	۲۳/۶۲	۶/۲۶
۲۴	۲۵/۰۸۰	۲۵/۷۶	۲۰/۴۷
۲۵	۱۹/۶۴۱	۱۹/۷۸	۱۳/۵۴
۲۶	۳۰/۸۲۹	۲۳/۲۷	۲۷/۶۸

*در جدول شماره ۳، میزان تولید واقعی حاصل از آزمایشگاه، و رگرسیون فازی هیبریدی حاصل از جدول شماره ۲ به صورت یکجا جهت مقایسه نمایش داده شد.

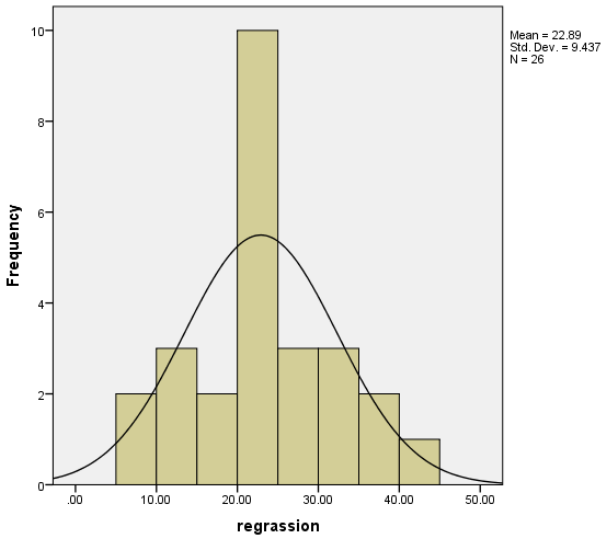
یکی از روش های آماری جهت بررسی مناسب بودن روش آمـــاری مورد استفاده، بررسی نمودار باقی مانده های مدل می باشد. هر چه انحراف معیار باقی مانده ها کمتر و منحنی به منحنی نـــرمال نزدیک تر باشد آن روش، روش مناسب تری خواهد بود. لذا در ادامه انحراف معیار باقی مانده ها(اختلاف مقادیر اصلی از مقادیر برآورد شـــده) و نمودار باقی مانده های مربوط به هر دو روش رگرسیون نقطه ای و رگرسیون فازی ذکر می شود(با استفاده از نرم افزار SPSS). در مقایسه نمودار روش فازی(نمودار شماره ۱) نسبت به نمودار روش نقطه ای(نمودار شماره ۲)، همان طور که ملاحظه می شود روش فازی انحراف کمتری نسبت نمودار نرمال دارد و از طرفی انحراف معیار روش رگرسیون نقطه ای برابر با ۹/۴۳۷ و انحراف معیار رگرسیون فازی برابر با ۲/۶۹۸ است. لذا می توان نتیجه گرفت که روش رگرسیون فازی هیبریدی روش بهینه ای جهت برآورد میزان تولید این نوع پروتئین است.

[Downloaded from sjimu.medilam.ac.ir on 2026-10-10]

[DOI: 10.29252/sjimu.27.3.1]



نمودار شماره ۱. نرمال بودن روش رگرسیون فازی. نمودار، انحراف معیار داده های برآورد شده به روش رگرسیون فازی هیبریدی را برابر با ۲/٦۹۸، نمایش می دهد



نمودار شماره ۲. نرمال بودن باقی مانده ها در روش رگرسیون نقطه ای. نمودار، انحراف معیار داده های برآورد شده به روش رگرسیون نقطه ای را برابر با ۹/٤۳۷، نمایش می دهد

[Downloaded from sjimu.medilam.ac.ir on 2026-10-10]

[DOI: 10.29252/sjimu.27.3.1]

بحث و نتیجه گیری

استفاده از باکتری E.coli به عنوان میزبان بیانی تولید پروتئین های نوترکیب دارای محاسنی است به این دلیل که اطلاعات زیادی در خصوص این میکروب در اختیار است، امکان دستیابی به بهره دهی های حجمی بالا برای تولید محصولات نوترکیب با E.coli وجود دارد و این میکروب را می توان بر روی محیط های کشت ارزان قیمت ساده رشد داد. هم چنین تولید خارج سلولی پروتئین، باعث حذف متیونین ناخواسته در ابتدای زنجیره پلی پپتیدی می شود. پروتیئن نوترکیب را تا حدودی در برابر تجزیه توسط پروتئازها حمایت می‌کند. محیط فضای پری پلاسمی، مناسب تشکیل پل های دی سولفیدی است. پروتیئن های در خارج سلولی را می توان، بدون نیاز به تجزیه سلولی، به دست آورد(۲۶).

روش های آماری برای مطالعات بهینه سازی فرآیند در بیوتکنولوژی پزشکی در مقیاس صنعت روشی ایده آل می باشد. در مرحله بهینه سازی فرآیند، ما برخی از پارامترهای مشخص از یک مجموعه را بدون نقض برخی از محدودیت، بهینه سازی می کنیم. از مهم ترین اهداف آن، به حداقل رساندن هزینه، به حداکثر رساندن توان و کارایی، برای مرحله بعدی مطالعات مقدماتی و در مقیاس بزرگ است. این یکی از ابزارهای کمی مهم در تصمیم گیری های صنعتی است(۲۶،۲۷).

جهت تولید پروتئین های درمانی توجه به یک سری نکات ضروری است از جمله برای پروتئین های درمانی تزریقی خلوص و درست بودن ساختار پروتئین بسیار مهم می باشد. هم چنین در تولید صنعتی این محصولات کیفیت و ایمنی محصول در درجه اول اهمیت قرار دارد. کارایی فرآیند از نقطه نظر حداقل کردن هزینه ها در درجه دوم اهمیت است(۲۷) که به منظور دستیابی به این هدف یعنی کاهش هزینه و جلوگیری از روش های آزمون و خطا باید از روش های آماری برای مطالعات بهینه سازی فرآیند در بیوتکنولوژی صنعتی و پزشکی استفاده کرد(۲۸،۲۹). در یک مطالعه مروری توسط نگارنده مقاله، تاثیر بهینه سازی فرآیند تولید پروتئین A بیانگر افزایش نتیجه پنج برابری تولید این پروتئین است(۳۰).

جهت طراحی آزمایشگاهی این پروتئین مهم و دستیابی به بهترین و بهینه ترین شرایط تولید، باید ابتدا به روش آماری میزان تولید آن محاسبه شود. لذا ابتدا باید بدانیم از کدام روش آماری برای این منظور استفاده کرد؟ و آیا روش آماری خاص مورد استفاده در سطح اطمینان مناسبی این برآورد را انجام خواهد داد؟

در این مطالعه از روش رگرسیون فازی هیبریدی جهت برآورد میزان ترشح پروتئین A استفاده شده ، با توجه به خصوصیات ریاضیات فازی و عدد مورد استفاده در این مقاله، یعنی عدد فازی مثلثی، برآورد حاصل یک برآورد فاصله ای می باشد که به مراتب دارای سطح اطمینان بالاتری نسبت به برآورد نقطه ای می باشد و از طرفی برآوردهای حاصل دارای سطح اطمینان ۸۰ در صد می باشد که می تواند سطح اطمینان خوبی برای این برآورد باشد. با استفاده از رگرسیون معمولی و با وجود شرایط مورد نظر می توان برآوردی جهت تولید پروتئین مورد نظر انجام داد.

اما در مطالعه حاضر، دلیل ما جهت برآورد میزان پروتئین A با استفاده از رگرسیون فازی هیبریدی اولاً جهت بالا بردن سطح اطمینان برآورد مورد نظر بوده و در ثانی برآورد فازی هیبریدی میزان تولید این پروتئین دارای واریانس کمتر و در حقیقت پراکندگی کمتر نسبت به روش رگرسیون معمولی باشد و از دیگر مزایای استفاده از روش رگرسیون فازی هیبریدی با کمترین مربعات آن است که اگر همه داده ها تردد(غیر فازی) باشند، نتایج حاصل از روش رگرسیون فازی هیبریدی دقیقاً مشابه رگرسیون معمولی می باشد که مقایسه کاملی از این دو روش در بخش نتایج انجام شده است و در جدول شماره ۳ و نمودارهای شماره ۱ و ۲ به خوبی نمایش داده شده است.

سپاسگزاری

نویسندگان از مدیریت پژوهشی دانشگاه صنعتی خاتم الانبیاء(ص) بهبهان قدردانی می نمایند. این مقاله مستخرج از طرح پژوهشی با عنوان«برآورد میزان ترشح پروتئین A نوترکیب با استفاده از رگرسیون فازی به

[Downloaded from sjimu.medilam.ac.ir on 2026-10-10]

[DOI: 10.29252/sjimu.27.3.1]

منظور تعیین بهینه تـــرین شرایط تولیـد» بـه شـماره ۰۰-۱۰-۹۴ مــصوب این دانشگاه می باشد.

References

1.Janzon L, Arvidson S. The role of the lysin gene in the regulation of virulence genes by the accessory gene regulator agr in Staphylococcus aureus. EMBO J1990;9:1391–9. doi:10.1002/j.1460-2075.1990.tb08254.x.

2.Ji G, Beavis RC, Novick RP. Cell density control of staphylococcal virulence mediated by an octapeptide pheromone. Proc Natl Acad Sci USA1995;92:12055-9. doi:10.1073/pnas.92.26.12055

3.Cheung AL, Bayer AS, Zhang G, Gresham H, Xiong YQ. Regulation of virulence determinants in vitro and in vivo in Staphylococcus aureus. FEMS Immunol Med Microbiol2004;40:1-9. doi:10.1016/S0928-8244(03)00309-2

4.Cheung AL, Eberhardt K, Heinrichs JH. Regulation of protein A synthesis by the sar and agr loci of Staphylococcus aureus. Infect Immun1997;65:2243-9. doi: 10. 0019-9567/97/$04.0010+0

5.Langone JJ. Protein A of Staphylococcus aureus and related immunoglobulin receptors produced by streptococci and pneumonococci. Adv Immunol 1982;32:157-252. doi:10.1016/S0065-2776(08)60722-1

6.Moks T, Abrahmsen L, Nilsson B, Hellman U, Sjoquist J, Uhlen M. Staphylococcal protein A consists of five IgG binding domains. Eur J Biochem 1986;156:637-43. doi:10.1111/j.1432-1033.1986.tb09625.x

7.Cheung AL, Schmidt K, Bateman B, Manna AC. SarS, a SarA homolog repressible by agr, is an activator of protein A synthesis in Staphylococcus aureus. Infect Immun 2001;69:2448-55. doi: 10.1128/IAI.69.4.2448-2455.2001

8.Graille M, Stura EA, Corper AL, Sutton BJ, Taussig MJ, Charbonnier JB, et al. Crystal structure of a Staphylococcus aureus protein A domain complexed with the Fab fragment of a human IgM antibody Structural basis for recognition of B cell receptors and superantigen activity. Proce National Acad Sci 2000;97:5399-404. doi:10.1073/pnas.97.10.5399

9.Lofdahl S, Guss B, Uhlen M, Philipson L, Lindberg M. Gene for staphylococcal protein A. Proc Natl Acad Sci USA. 1983;80:697-701. doi:10.1073/pnas.80.3.697

10.Starovasnik MA, Oconnell MP, Fairbrother WJ, Kelley RF. Antibody variable region binding by Staphylococcal protein A thermodynamic analysis and location of the Fv binding site on E domain. Prote Sci1999;8:1423-31.

11.Tashiro M, Montelione GT. Structures of bacterial immunoglobulin binding domains and their complexes with immunoglobulins. Curr Opin Struct Biol1995;5:471-81. doi:10.1016/0959-440X(95)80031-X

12.Loghem E, Frangione B, Recht B, Franklin EC. Staphylococcal protein A and human IgG subclasses and allotypes. Scand J Immunol1982;15:275-8. doi:10.1111/j.1365-3083.1982.tb00649.x

13.Ghaedmohammadi S, Rigi G, Zadmard R, Ricca E, Ahmadian G. Immobilization of Bioactive protein a from Staphylococcus aureu SpA on the surface of Bacillus subtilis spores. Mol Biotechnol 2015;57:756-66. doi: 10.1007/s12033-015-9868-z

14.Rigi G, Beyranvand P, Ghaedmohammadi S, Heidarpanah S, Akbari Noghabi K, Ahmadian G. Comparison of the extracellular full length and truncated recombinant protein A production in Escherichia coli BL21 DE3. J Paramed Sci 2015;6:35-45. doi:10.1.1.845.4772

15.Rigi G, Mohammadi SG, Arjomand MR, Ahmadian G, Noghabi KA. Optimization of extracellular truncated Staphylococcal protein A expression in Escherichia coli BL21 DE3. Biotechnol Appl Biochem 2013;61: 217-225 . doi:10.1002/bab.1157

16. Enein HY, Rigi G, Farhadpour M, Ghasempour A, Ahmadian G. Per aqueous liquid chromatography as a simple method for native separation of protein A. Chromatographia. 2017; 80: 1633-19390. doi: 10.1007/s10337-017-3412-1

17.Khodaei S, Ghaedmohammadi S, Mohammadi M, Rigi G, Ghahremanifard P, Zadmard R, et al. Covalent immobilization of protein A on chitosan and aldehyde

[Downloaded from sjimu.medilam.ac.ir on 2026-10-10]

[DOI: 10.29252/sjimu.27.3.1]

double branched chitosan as biocompatible carriers for Immunoglobulin G purification. J Chrom Sci2018; 56:933-40. doi:10.1093/chromsci/bmy070
18.Ghahremanifard P, Rezaeinezhad N, Rigi G, Ramezani F, Ahmadian G. Designing a novel signal sequence for efficient secretion of Candida antarctica lipase B in E. coli the molecular dynamic simulation codon optimization and statistical analysis approach. Int J Biol Macromol 2018; 119:291-305. doi:10.1016/j.ijbiomac.2018.07.150
19.Rigi G, Bahrami T, Armand R, Piruzeh Z. Estimation of the amount of recombinant protein A secretion using fuzzy regression. Iran J Health Sci 2016;4:35-44. doi: 10.18869/acadpub.jhs.4.3.35
20.Bahrami T, Chashiani S. Estimating daily peak load using symmetrical doubly linear adaptive-fuzzy regression model. Cumhu Sci J2015;36:715-20. doi: 10.1.1.922.7288
21.Tanaka H, Watada J. Possibilistic linear systems and their application to the linear regression model. Fuz Set Sys 1988;27:275-89. doi:10.1016/0165-0114(88)90054-1
22.Moskowitz H, Kim K. On assessing the H value in fuzzy linear regression. Fuz Set Sys1993;58:303-27. doi:10.1016/0165-0114(93)90505-C
23.Chang PT. Fuzzy seasonality forecasting. Fuz Set Sys 1997;90:1-10. doi:10.1016/S0165-0114(96)00138-8
24.Bahrami T, Chashiani S. Estimation of daily peak of a house feeder with least squares regression of Fuzzy hybrid. MSE2015;114:1142-6.
25.Chang YH. Hybrid fuzzy least squares regression analysis and its reliability measures. Fuz Set Sys 2001;119:225-46. doi:10.1016/S0165-0114(99) 00092-5
26.Oskouie SFG, Tabandeh F, Yakhchali B, Eftekhar F. Response surface optimization of medium composition for alkaline protease production by Bacillus clausii. Biochem Eng J 2008;39:37-42. doi:10.1016/j.bej.2007.08.016
27.Singh AK, Mehta G, Chhatpar HS. Optimization of medium constituents for improved chitinase production by Paenibacillus sp. D1 using statistical approach. Let Appl Microbiol 2009;49:708-14. doi: 10.1111/j.1472-765X.2009.02731.x
28.Benito Y, Kolb FA, Romby P, Lina G, Etienne J, Vandenesch F. Probing the structure of RNAIII the Staphylococcus aureus agr regulatory RNA and identification of the RNA domain involved in repression of protein A expression. Rna2000;6:668-79. doi: 10.1017/S1355838200992550
29.Novick RP. Autoinduction and signal transduction in the regulation of staphylococcal virulence. Mol Microbiol2003;48:1429-49. doi:10.1046/j.1365-2958.2003.03526.x
30.Rigi G, Ghaedmohammadi S, Ahmadian G. A comprehensive review on staphylococcal protein A its production and applications. Biotechnol Appl Biochem 2019;2:231-6. doi: 10.1002/bab.1742

[Downloaded from sjimu.medilam.ac.ir on 2026-10-10]

[DOI: 10.29252/sjimu.27.3.1]

Fuzzy Hybrid least-Squares Regression Approach to Estimating the amount of Extra Cellular Recombinant Protein A from Escherichia coli BL21

Armand R[1], Rigi G[2,3], Bahrami T[4]*

(Received: November 5, 2017 Accepted: June 18, 2019)

Abstract

Introduction: Immune Protein A is a component with a vast spectrum of biochemical, biological and medical usages. The coding gene of this protein was extracted from Staphylococcus aureus and was cloned and expressed in Escherichia coli bacteria. Suitable statistical methods are utilized to optimize expression conditions for evaluating experiment accuracy , guarantee the accuracy of subsequent experiments, reduce cost , prevent trial and error method, and obtain the highest production level.

Materials & Methods: Normal statistical regression method is based on the assumption of accuracy of variables and their observations, and finally the relationship among the variables are precisely specified .In normal modeling ,such as estimation of protein level inaccurate observations and vague relationships may be encountered; therefore,utilization of regression methods capable of explaining the vague structure of protein level and providing the models attunes to reality is necessary. In this article, hybrid fuzzy regression method with the least square, based on fuzzy-set theory was utilized.

Findings: According to the results, Protein A production level was estimated at 90 % level .On the other hand, since the method utilized in this article was hybrid fuzzy linear regression method with the least square errors, we can conclude that if all the data utilized in this article are crisp numbers,the hybrid fuzzy regression will produce results similar to normal regression.

Discussion & Conclusions: One of the advantages of hybrid fuzzy regression is higher level of certainty than point estimation. In this study, the standard deviation of estimated data by fuzzy hybrid regression method was lower than conventional regression method.Therefore, it can be concluded that fuzzy hybrid regression method is an optimal method for estimating the production of this type of protein.

Keywords: Protein A, Immunoglobin binding protein, Hybrid fuzzy regression, Medical biotechnology

1. Dept of Biology, Faculty of Science, Behbahan Khatam Alanbia University of Technology, Behbahan, Iran
2. Dept of Genetics, Faculty of Basic Sciences, University of Shahrekord, Shahrekord, Iran
3. Dept of Industrial Biotechnology, Research Institute of Biotechnology, Shahrekord University, Shahrekord, Iran
4.Dept of Statistics, Faculty of Science, Behbahan Khatam Alanbia University of Technology, Behbahan, Iran

[Downloaded from sjimu.medilam.ac.ir on 2026-10-10]

[DOI: 10.29252/sjimu.27.3.1]

ضمیمه(Appendix):

مدل رگرسیون فازی هیبریدی به شرح زیر می باشد:

$$\tilde{Y}_i = (a_0, a_{0l}, a_{0r}) + (a_1, a_{1l}, a_{1r})X_1 + (a_2, a_{2l}, a_{2r})X_2 + \cdots + (a_p, a_{pl}, a_{pr})X_p \quad (1)$$

که در آن X_i ها متغیر مستقل، a_i, a_{ir}, a_{il} $(i = 0, 1, \ldots, p)$ به ترتیب نیم پهنای چپ، نیم پهنای راست و مرکز ضرایب رگرسیون می باشند و P تعداد متغیرهای مستقل می باشد.

بنا بر این می توان برآورد متغیر وابسته را با استفاده از حل دستگاه های معادلات زیر محاسبه می شود:

$$na_0 + \left(\sum_{i=1}^{n} X_{1,i}\right)a_1 + \left(\sum_{i=1}^{n} X_{2,i}\right)a_2 + \cdots + \left(\sum_{i=1}^{n} X_{p,i}\right)a_p = \sum_{i=1}^{n} Y_i \quad (2)$$

$$\left(\sum_{i=1}^{n} X_{1,i}\right)a_0 + \left(\sum_{i=1}^{n} {X^2}_{1,i}\right)a_1 + \left(\sum_{i=1}^{n} X_{1,i}X_{2,i}\right)a_1 + \cdots + \left(\sum_{i=1}^{n} X_{1,i}X_{p,i}\right)a_p = \sum_{i=1}^{n} X_{1,i}Y_i \quad (3)$$

$$\vdots$$

$$\left(\sum_{i=1}^{n} X_{p,i}\right)a_0 + \left(\sum_{i=1}^{n} X_{p,i}X_{1,i}\right)a_1 + \left(\sum_{i=1}^{n} X_{p,i}X_{2,i}\right)a_1 + \cdots + \left(\sum_{i=1}^{n} {X^2}_{p,i}\right)a_p = \sum_{i=1}^{n} X_{p,i}Y_i \quad (4)$$

$$na_{0l} + \left(\sum_{i=1}^{n} X_{1,i}\right)a_{1l} + \left(\sum_{i=1}^{n} X_{2,i}\right)a_{2l} + \cdots + \left(\sum_{i=1}^{n} X_{p,i}\right)a_{pl} = \sum_{i=1}^{n} Y_{il} \quad (5)$$

$$\left(\sum_{i=1}^{n} X_{1,i}\right)a_{0l} + \left(\sum_{i=1}^{n} {X^2}_{1,i}\right)a_{1l} + \left(\sum_{i=1}^{n} X_{1,i}X_{2,i}\right)a_{1l} + \cdots + \left(\sum_{i=1}^{n} X_{1,i}X_{p,i}\right)a_{pl} = \sum_{i=1}^{n} X_{1,i}Y_{il} \quad (6)$$

$$\vdots$$

$$\left(\sum_{i=1}^{n} X_{p,i}\right)a_{0l} + \left(\sum_{i=1}^{n} X_{p,i}X_{1,i}\right)a_{1l} + \left(\sum_{i=1}^{n} X_{p,i}X_{2,i}\right)a_{1l} + \cdots + \left(\sum_{i=1}^{n} {X^2}_{p,i}\right)a_{pl} = \sum_{i=1}^{n} X_{p,i}Y_{il} \quad (7)$$

$$\vdots$$
$$\vdots$$
$$\vdots$$

[Downloaded from sjimu.medilam.ac.ir on 2026-10-10]

[DOI: 10.29252/sjimu.27.3.1]

$$na_{0r} + \left(\sum_{i=1}^{n} X_{1,i}\right) a_{1r} + \left(\sum_{i=1}^{n} X_{2,i}\right) a_{2r} + \cdots + \left(\sum_{i=1}^{n} X_{p,i}\right) a_{pr} = \sum_{i=1}^{n} Y_{ir} \qquad (8)$$

$$\left(\sum_{i=1}^{n} X_{1,i}\right) a_{0r} + \left(\sum_{i=1}^{n} X^2{}_{1,i}\right) a_{1r} + \left(\sum_{i=1}^{n} X_{1,i}X_{2,i}\right) a_{1r} + \cdots + \left(\sum_{i=1}^{n} X_{1,i}X_{p,i}\right) a_{pr} = \sum_{i=1}^{n} X_{1,i}Y_{ir} \qquad (9)$$

$$\vdots$$

$$\left(\sum_{i=1}^{n} X_{p,i}\right) a_{0r} + \left(\sum_{i=1}^{n} X_{p,i}X_{1,i}\right) a_{1r} + \left(\sum_{i=1}^{n} X_{p,i}X_{2,i}\right) a_{1r} + \cdots + \left(\sum_{i=1}^{n} X^2{}_{p,i}\right) a_{pr} = \sum_{i=1}^{n} X_{p,i}Y_{ir} \qquad (10)$$

Y_i و Y_{il} و Y_{ir} به ترتیب مرکز، گستره چپ و گستره راست عدد فازی وابسته می باشد.

[Downloaded from sjimu.medilam.ac.ir on 2026-10-10]

[DOI: 10.29252/sjimu.27.3.1]