



پیش‌بینی کیفیت فرآیندهای تولیدی با بُعد بالا به کمک رگرسیون تنسوری

رسول نورالسنا^{*}، فرشاد رمضانخانی^۲

۱. استاد گروه صنایع، دانشکده مهندسی صنایع، دانشگاه علم و صنعت، تهران، ایران
۲. دانشجوی دکتری گروه صنایع، دانشکده مهندسی صنایع، دانشگاه علم و صنعت، تهران، ایران

خلاصه

با پیشرفت فناوری‌های رایانه‌ای مدرن و سیستم‌های سنجش تعبیه شده در حوزه‌های مختلف، رشد چشم‌گیری در حجم، تنوع و سرعت داده‌های تولید شده وجود دارد که منبع غنی از اطلاعات هستند. داده‌های تنسوری یکی از مهم‌ترین نوع داده‌های پیچیده ساختاریافته با ابعاد بالا هستند که روش‌های سنتی قادر به حل این نوع از داده‌ها نیستند. مدل‌های رگرسیونی تنسوری کاربرد زیادی در بهینه‌سازی مسائلی دارد که در آن‌ها تعدادی متغیر مستقل عددی به‌عنوان ورودی مسأله وجود دارند که با تغییر آن‌ها می‌توان خروجی بهتری را ایجاد کرد. در این مطالعه، مدل رگرسیونی ارائه شده است که متغیرهای مستقل آن به‌صورت عددی و متغیر پاسخ آن به‌صورت تنسوری است. مجموعه‌ای از روش‌های جبر خطی و رویکردهای تنسوری به‌منظور پیدا کردن الگوهای درون مجموعه نقاط در فضا و ارتباط آن‌ها با متغیرهای فرآیند پیشنهاد شده است. جوشکاری اصطکاکی اغتشاشی به‌عنوان مثال انگیزشی این مطالعه برای تأیید نتایج در نظر گرفته شده است. مدل پیشنهادی در نرم‌افزار R کدنویسی شده است. نتایج نشان داد که مدل پیشنهادی عملکرد خوبی در پیش‌بینی خروجی فرآیندها با تنظیم پارامترهای آن به‌صورت پویا دارد.

اطلاعات مقاله

تاریخچه مقاله:

دریافت ۱۴۰۰/۶/۱۳

پذیرش ۱۴۰۰/۹/۲۷

(مقاله پژوهشی)

کلمات کلیدی:

رگرسیون تنسوری

تجزیه تنسور

مدل‌سازی فرآیند

تحلیل داده‌های ساختار

پیچیده

۱. مقدمه

امروزه رشد قابل توجهی در حجم، تنوع و سرعت داده‌های تولید شده توسط انسان‌ها، ماشین‌ها، تجهیزات، حسگرها، دوربین‌ها و ابزارها در بسیاری از فرآیندها به چشم می‌خورد که سرشار از اطلاعات ارزشمند هستند. این داده‌ها صرفاً به‌صورت عددی نیستند و می‌توانند انواع مختلفی از داده‌ها از جمله تصویر، ویدئو، صدا، متن و دیگر داده‌ها باشند که به‌دلیل وجود پیچیدگی‌های ذاتی، در سال‌های اخیر مورد توجه بسیاری از محققان قرار گرفته است. مطالعات بسیار زیادی در زمینه‌های مختلف منتشر شده‌اند که بر اهمیت داده‌های بزرگ تأکید می‌کنند [۱]. در صنایع تولیدی، تعداد زیادی نرم‌افزار و سخت‌افزار وجود دارند که داده‌های ارزشمندی را برای افزایش کارایی و بهره‌وری سیستم‌ها ارائه می‌دهند [۲]، اما به‌دلیل عدم وجود روش کافی برای

تجزیه و تحلیل، بسیاری از این داده‌ها جمع‌آوری نمی‌شوند و برای همیشه از دست می‌روند. تقریباً تمام این داده‌ها دارای ابعاد بالا و پیچیدگی ساختاری هستند که چالش‌های جدیدی را از لحاظ نظری و محاسباتی نسبت به روش‌های کلاسیک به‌وجود می‌آورند.

استفاده از روش‌های یادگیری ماشین رویکردی قدرتمند برای حل مسائلی است که در آن داده‌های با ساختار پیچیده وجود دارد. در مطالعه مولود و عبدالعزیز [۳] به این موضوع اشاره می‌شود که رگرسیون یکی از رایج‌ترین و جامع‌ترین روش‌های آماری و یادگیری ماشین است. در مطالعه موربلا و همکاران [۴] نیز روش رگرسیون یکی از مهم‌ترین الگوریتم‌هایی است که به‌عنوان ابزار یادگیری ماشین جهت پیش‌بینی عملکرد مورد استفاده قرار گرفته است. رگرسیون یک تکنیک آماری در یادگیری ماشین است که به‌صورت گسترده برای پیش‌بینی و

* نویسنده مسئول: رسول نورالسنا

پست الکترونیکی: rassoul@iust.ac.ir

چندصد متغیر است که اگر مدل‌های رگرسیونی با تعدادی بیش از یک مقدار مشخص متغیر کار کنند، با شکست مواجه می‌شوند. این موضوع، استفاده از تجزیه تنسور را برای کاهش متغیر در یادگیری ماشین ضروری می‌کند. در این راستا تجزیه‌های تنسوری که تعمیمی از تجزیه‌های ماتریسی هستند ارائه داده شده و مورد استفاده قرار می‌گیرند. روش پیشنهادی تنسوری ارائه شده در این مطالعه نسبت به سایر مطالعات انجام شده در پیش‌بینی کیفیت، می‌تواند به سه ویژگی بسیار مهم ابعاد فضای پارامتری بسیار پایین، پیچیدگی محاسباتی کم و همچنین بهبود قابلیت تفسیر مدل اشاره کرد.

ادامه مقاله به صورت زیر سازماندهی شده است. مرور ادبیات موضوع در بخش ۲ معرفی شده است. در بخش ۳ مدل رگرسیون با متغیر پاسخ تنسوری و متغیر مستقل عددی و نشانه‌گذاری پایه‌ای آن ارائه شده است. مثال انگیزشی مطالعه و مطالعات شبیه‌سازی در بخش ۴ انجام شده است. نتیجه‌گیری و مطالعات آتی در بخش ۵ ارائه شده است.

۲. پیشینه تحقیق

برای تجزیه تنسورهای با ابعاد بالا روش‌های زیادی توسعه داده شده است که تعداد زیادی از آن‌ها در [۲۳] معرفی شده است. در بین انواع زیادی از روش‌های تجزیه تنسور، تجزیه توکر [۲۴] و تجزیه CP [۲۵] بیشترین توجه را داشته‌اند. مقاله [۲۲] به تفصیل این روش‌ها را توضیح داده است. همچنین مقاله [۲۶] مروری بر مطالعاتی که از روش‌های تجزیه توکر و تجزیه CP برای داده‌های تنسوری مرتبط با داده‌های مراقبت‌های بیمارستانی استفاده کرده‌اند انجام داده است. روش‌های تجزیه توکر و CP نسخه تعمیم‌یافته روش تحلیل مؤلفه اصلی یا به طور خاص، تعمیم روش تجزیه مقادیر منفرد هستند [۲۳]. در تعدادی از مطالعات روش تجزیه توکر با نام تجزیه مقدار منفرد مرتبه بالا معرفی شده است [۲۷]. در مقاله [۲۸] یک روش تجزیه مرتبه بالا به نام تجزیه مقدار منفرد مرتبه بالا برای داده‌های تنسوری مربوط به هندسه ابزار جوشکاری فراصوتی ارائه شده است که با استفاده از آن مجموعه‌ای با ابعاد کمتر از داده‌های مرتبه بالا استخراج شده است. مقاله [۲۹] روش‌های انتقال فوریه سریع، تجزیه موجک، تحلیل مؤلفه اصلی و شناسایی ویژگی‌های غالب را برای استخراج ویژگی‌ها از داده‌ها ارائه داده است.

مقاله [۳۰] چارچوب تحلیل مؤلفه اصلی چندخطی را برای استخراج ویژگی از داده‌های تنسوری ارائه کرده است که الگوریتمی تکراری است و در آن مسأله اصلی به چندین زیر مسأله تبدیل می‌شود. مقاله [۳۱] از تجزیه تنسور برای استخراج ویژگی از داده‌های مربوط به تصاویر خاکستری و رنگی پرداخته است. استفاده از تجزیه تنسور در مطالعات و زمینه‌های مختلفی به چشم می‌خورد که نشان از اهمیت و مفید بودن آن است. از جمله این زمینه‌ها می‌توان به پردازش سیگنال، جبرخطی عددی، بینایی کامپیوتر، تجزیه و تحلیل عددی، داده‌کاوی، تجزیه و تحلیل گراف و علوم اعصاب اشاره کرد که در مقاله [۳۲] برای

شناخت ارتباط بین متغیرها و شکل این روابط با استفاده از برازش خط، منحنی، صفحه یا ابرصفحه روی داده‌ها استفاده می‌شود؛ طوری که به راحتی می‌توان تغییر در مقدار متغیر وابسته را با استفاده از تغییر در مقدار یک متغیر مستقل و ثابت بودن سایر متغیرهای مستقل شناسایی کرد. منظور از برازش توسط مدل رگرسیونی، برآورد پارامترهای مدل توسعه داده شده است. رگرسیون تقریباً در بسیاری از زمینه‌ها از جمله مهندسی، فیزیک، اقتصاد، مدیریت، علوم زیستی، بیولوژی و علوم اجتماعی برای برآورد و پیش‌بینی مورد نیاز است. در طول سال‌های گذشته مدل‌های مختلف رگرسیونی جهت پیاده‌سازی بر روی مسائل گوناگون توسعه داده شده است. از جمله این مدل‌های رگرسیونی می‌توان به رگرسیون خطی [۵]، لجستیک [۶]، چندکی [۷]، ستیغی [۹]، لاسو [۱۰]، شبکه الاستیک [۱۱]، مؤلفه‌های اصلی [۱۲]، کمترین مربعات جزئی [۱۳]، بردار پشتیبان [۱۴]، ترتیبی [۱۵]، پواسون [۱۶]، دو جمله‌ای منفی [۱۷]، شبه پواسون [۱۸]، کاکس [۱۹]، کمترین زاویه [۲۰] و تابعی [۲۱] اشاره کرد. یکی از بزرگترین اشتباهاتی که در به کارگیری از این مدل‌ها می‌تواند رخ دهد این است که به دلیل آگاهی نداشتن یا ساده کردن حل مسأله، یک مدل اشتباه و مغایر با مدلی که ذات اصلی مسأله نیاز دارد انتخاب شود. در واقع، مدل رگرسیونی باید دقیقاً با توجه به مسأله و نوع داده آن انتخاب شود. یکی از داده‌های بسیار مهمی که در ادبیات موضوع رگرسیون توجه زیادی به آن نشده است یا در مطالعات موردی صنعتی تعداد کمی به کار گرفته شده است، مدل‌های رگرسیونی برای داده‌های تنسوری با مرتبه بالا است. تنسورها یکی از انواع داده‌های با ابعاد بالا هستند که فرصت‌های زیادی را برای تجزیه و تحلیل به‌ارمغان می‌آورند. تنسور عنصری هندسی است که به منظور تعمیم مفاهیم عددی، بردارها و ماتریس‌ها به ابعاد بالاتر معرفی می‌شوند. تنسورها می‌توانند از مرتبه صفر تا بی‌نهایت تعریف شوند. عدد، بردار و ماتریس به ترتیب تنسورهای مرتبه صفر، ۱ و ۲ هستند. تنسورهای مرتبه ۳ و بالاتر را با نام آرایه‌های چندبعدی نیز می‌شناسند. در واقع، تنسور آرایه‌ای چندبعدی از اعداد است که در یک جدول با یک ساختار مشخص در کنار هم چیده شده‌اند. تنسور از ابزارهای مفید برای ذخیره‌سازی و کار با داده‌های زیاد که چند ویژگی از آن‌ها در دسترس است می‌باشد. مفاهیم جبری داده‌های تنسوری در مطالعه گُلدا و بَدِر [۲۲] به تفصیل بیان شده است که در این مطالعه نیز از آن‌ها استفاده شده است.

در این مطالعه به دنبال حل مدل رگرسیونی هستیم که متغیرهای مستقل آن به صورت عددی و متغیر پاسخ آن به صورت تنسوری باشد. در واقع نوآوری اصلی این مقاله توسعه یک مدل ریاضی رگرسیونی با متغیرهای مستقل عددی و متغیر پاسخ تنسوری و چگونگی حل آن زمانی که تعداد نمونه‌ها خیلی کمتر از تعداد متغیرها باشد در نظر گرفته می‌شود. این مدل رگرسیونی کاربرد زیادی در بهینه‌سازی مسائلی دارد که در آن‌ها تعدادی متغیر مستقل عددی به عنوان ورودی مسأله وجود دارند که با تغییر آن‌ها می‌توان خروجی بهتری را که از نوع تنسور است ایجاد کرد. یکی چالش‌های ایجاد شده در این مدل رگرسیونی وجود

هرکدام از این زمینه‌ها مقالات و کتب زیادی معرفی شده است. مقاله [۳۳] مفاهیم مربوط به تنسور و تجزیه و تحلیل آن را به صورت کاربردی توضیح داده است. مقاله‌های [۳۴]، [۳۵] روش حداقل مربع جزئی را برای کاهش بعد در نظر گرفته‌اند.

استفاده از رگرسیون تنسوری مزایای زیادی مخصوصاً زمانی که تعداد متغیر تأثیرگذار در مسأله بسیار بیشتر از تعداد نمونه در دسترس باشد دارد. از جمله مزایایی که در ادبیات موضوع اشاره شده است می‌توان به مواردی همچون در نظرگیری تعداد زیادی عامل تأثیرگذار متفاوت در کنار یکدیگر، کاهش بُعد داده‌ها و فضای پارامتری در نظر گرفته شده در مدل، بهبود تفسیرپذیری مدل و کاهش پیچیدگی محاسباتی اشاره کرد [۳۶].

مطالعات کمی رگرسیون تنسوری را بررسی کرده‌اند که بیشتر این مطالعات هم در سال‌های اخیر اتفاق افتاده است. مطالعه [۳۷] مدلی را ارائه داده است که در آن تنها متغیرهای مستقل تنسوری در نظر گرفته شده است که ایده اصلی آن کاهش بُعد فضای پارامترهای رگرسیون با اعمال یک روش تجزیه مانند تجزیه CP و حل مسأله رگرسیون خطی است. مطالعه [۳۸] مدل رگرسیونی ارائه داده است که متغیر مستقل و وابسته تنسوری دارد. مقاله [۳۹] چالش‌ها و راهکارهای موجود برای تولید بدون نقص را با استفاده از مدل‌سازی داده‌های فرآیند تشریح می‌کند که داده‌های تنسوری نیز در آن مطرح شده است. مطالعه [۴۰] چارچوبی کلی را ارائه کرده است که در آن به پیش‌بینی خطی انواع متغیرهای از نوع آرابه چند بُعدی برای متغیرهای ورودی و خروجی مدل رگرسیون از جمله خروجی عدد از ورودی تنسور، ماتریس از ماتریس یا تنسور از عدد پرداخته است.

زره‌ساز [۴۱] در پایان نامه خود مدل رگرسیونی ارائه کرده است که در آن انواع متغیرها از جمله متغیر تنسوری به عنوان متغیر مستقل در نظر گرفته شده است. در این مطالعه نقطه ماکزیمم نمودار تنش- کرنش با استفاده از متغیرهای اولیه، تابع نیرو و تصویر حرارتی پیش‌بینی شده است. یکی از راه‌های کاهش بُعد در داده‌های تنسوری، استفاده از روش‌های تجزیه تنسور است. مطالعات مربوط به رگرسیون تنسوری بیشتر در موضوعات پزشکی انجام شده است. گوا و همکاران [۳۷] مدل رگرسیون برای تخمین ژست سر، ژست بدن و سن انسان با متغیر مستقل تنسوری و متغیر پاسخ به صورت عددی ارائه داده است که در آن به کاهش بُعد پارامترهای مدل و حل آن با مسأله رگرسیون خطی پرداخته است. لی و همکاران [۳۶] مدل رگرسیون برای تصاویر عصب ارائه کرده است که در آن متغیر مستقل به صورت تنسور و متغیر پاسخ به صورت عددی است. از روش تجزیه تنسور برای حل مدل ارائه شده استفاده شده است. اختلالات نوروزئیک و شناسایی بیماری‌های مربوط به نواحی مغز با استفاده از مدل چندمتغیره با ورودی تصویر به فرم تنسوری در مطالعه ژانگ و همکاران [۴۲] انجام شده است. ژو و همکاران [۴۳] به پیش‌بینی نتیجه بالینی با استفاده از مدل خطی تعمیم یافته با ورودی تصویر به فرم تنسوری پرداخته‌اند. هو و همکاران [۴۴] نیز در مطالعه خود از تصاویر مغز به عنوان متغیر مستقل تنسوری

برای پیش‌بینی بیماری استفاده کرده است. اسپنسر و همکاران [۴۵]، یک روش بیزی به منظور مدل‌سازی مدل رگرسیونی با متغیر پاسخ عددی و یک متغیر مستقل تنسور ارائه کرده است که در آن از روش تجزیه تنسور توکر استفاده شده است. در این مطالعه تجزیه و تحلیل تصویربرداری عصبی با استفاده از داده‌های آرایمر در نظر گرفته شده است. در سال‌های اخیر مطالعات دیگری نیز وجود دارد که متغیر پاسخ به صورت عددی و متغیر مستقل به صورت تنسوری در نظر گرفته شده است [۴۶]، [۴۷]. ژو و همکاران [۴۸] یک مدل رگرسیون تنسوری دینامیکی ارائه داده‌اند که متغیر پاسخ آن به صورت تنسور و تعدادی متغیر مستقل به عنوان پیش‌بینی‌کننده در نظر گرفته شده است. صحت‌سنجی مدل ارائه شده توسط دو مطالعه موردی در تصاویر عصب‌شناسی و مطالعات تبلیغاتی دیجیتال بررسی شده است.

۳. مدل رگرسیون با متغیر پاسخ تنسوری و متغیر مستقل عددی

به منظور درک آسان معادلات ارائه شده، تعدادی از نمادهای استفاده شده در مقاله در ادامه معرفی می‌شود. مقادیر عددی به صورت حروف کوچک کج (یعنی x)، بردارها به صورت حروف کوچک ضخیم (یعنی \mathbf{x})، ماتریس‌ها به صورت حروف بزرگ ضخیم (یعنی \mathbf{X}) و تنسور (آرایه‌های چندبعدی) به صورت حروف بزرگ خوشنویسی (یعنی \mathbf{X}) در نظر گرفته می‌شود.

به منظور سادگی و درعین حال بدون ازدست دادن کلیت مدل، روش پیشنهادی در این مطالعه با متغیر پاسخ تنسوری دوبعدی ارائه شده است که به راحتی می‌توان با افزودن ابعاد دیگر به تنسورهای مرتبه بالاتر توسعه داد. فرض کنید یک مجموعه آموزشی با اندازه N نمونه در دسترس است که متغیرهای پاسخ تنسوری با $\mathbf{Y}_i \in \mathbb{R}^{D_1 \times D_2}$; $i = 1, \dots, N$ و بردار متغیرهای مستقل مربوطه با $\mathbf{x}_i \in \mathbb{R}^P$; $i = 1, \dots, N$ نشان داده می‌شود. که در آن D_1 و D_2 ابعاد تنسوری مرتبه دوم (D_1 ردیف و D_2 ستون ماتریس) و P تعداد ضرایب رگرسیون در هر نمونه است. در هریک از نمونه‌های در نظر گرفته شده، بردار متغیرهای مستقل به صورت یک بردار با تعداد P عضو است که به صورت x_{ij} ; $i = 1, \dots, N$, $j = 1, \dots, P$ تعریف می‌شوند. هریک از عضوهای متغیرهای مستقل مدل رگرسیونی در یک نمونه خاص، اثری بر روی متغیر پاسخ همان نمونه دارد که این اثر به عنوان ضریب رگرسیونی در نظر گرفته می‌شوند. به بیانی دیگر در یک نمونه مشخص، به ازای هریک از عناصر بردار متغیرهای مستقل \mathbf{x}_{ij} , $i = 1, \dots, N$ یک ضریب تنسوری با ابعادی مانند متغیر پاسخ در نظر گرفته می‌شود که با $\mathbf{A}_{ij} \in \mathbb{R}^{D_1 \times D_2}$; $i = 1, \dots, N$, $j = 1, \dots, P$ در واقع، در نمونه i ام، اثر متغیر مستقل i ام بر متغیر پاسخ به صورت $\mathbf{A}_{ij} x_{ij}$ تعریف می‌گردد. همچنین برای هریک از متغیرهای مستقل خطای تصادفی $\mathbf{E}_i \in \mathbb{R}^{D_1 \times D_2}$, $i = 1, \dots, N$ در نظر گرفته می‌شود. رگرسیون تنسوری که هدف آن پیش‌بینی متغیر پاسخ \mathbf{Y}_i با استفاده از متغیرهای مستقل \mathbf{x}_i به صورت معادله (۱) تعریف می‌شود.

آن زیاد باشد، می‌تواند منجر به بیش‌برازش بشود. بنابراین در این مطالعه از روش تجزیه توکر برای کاهش بُعد تنسور استفاده می‌شود.

تجزیه توکر نوعی تحلیل مؤلفه‌های اصلی مرتبه بالا است که در آن یک تنسور به مجموعه‌ای از ماتریس‌های عامل و یک تنسور هسته‌ای تجزیه می‌شود. تجزیه توکر یک تنسور مرتبه سوم مانند $\mathcal{A} \in \mathbb{R}^{D_1 \times D_2 \times P}$ که در این مطالعه نیز مورد بررسی است به صورت معادله (۷) تجزیه می‌شود.

$$\mathcal{A} \approx \mathcal{Q} \times_1 \mathbf{S} \times_2 \mathbf{T} \times_3 \mathbf{U} \quad (7)$$

که در آن، $\mathbf{S} \in \mathbb{R}^{D_1 \times I}$ ، $\mathbf{T} \in \mathbb{R}^{D_2 \times J}$ و $\mathbf{U} \in \mathbb{R}^{P \times K}$ به عنوان ماتریس‌های عامل و $\mathcal{Q} \in \mathbb{R}^{I \times J \times K}$ به عنوان تنسور هسته‌ای در نظر گرفته می‌شوند. I ، J و K به ترتیب تعداد اجزای ماتریس‌های عامل \mathbf{S} ، \mathbf{T} و \mathbf{U} هستند. باتوجه به معادله (۵)، مدل ماتریس شده معادله (۷) به‌ازای هر ۳ حالت تنسور مرتبه ۳ ارائه شده به صورت ۳ معادله (۸)، (۹) و (۱۰) در نظر گرفته می‌شود.

$$\mathbf{A}_{(1)} \approx \mathbf{S} \mathbf{Q}_{(1)} (\mathbf{U} \otimes \mathbf{T})' \quad (8)$$

$$\mathbf{A}_{(2)} \approx \mathbf{T} \mathbf{Q}_{(2)} (\mathbf{U} \otimes \mathbf{S})' \quad (9)$$

$$\mathbf{A}_{(3)} \approx \mathbf{U} \mathbf{Q}_{(3)} (\mathbf{T} \otimes \mathbf{S})' \quad (10)$$

که در آن، علامت ' نشان‌دهنده‌ی ترانپوذه یک ماتریس و نماد \otimes نشان‌دهنده‌ی ضرب کرانکر دو ماتریس است. ضرب کرانکر دو ماتریس $\mathbf{W} \in \mathbb{R}^{m_1 \times m_2}$ و $\mathbf{Q} \in \mathbb{R}^{l_1 \times l_2}$ به صورت $\mathbf{Q} \otimes \mathbf{W}$ است که یک ماتریس با ابعاد $l_1 m_1 \times l_2 m_2$ ایجاد می‌شود. مقادیر ماتریس به‌دست آمده به صورت معادله (۱۱) محاسبه می‌شود.

$$\mathbf{Q} \otimes \mathbf{W} = \begin{bmatrix} q_{11} \mathbf{W} & \dots & q_{1l_2} \mathbf{W} \\ \vdots & \ddots & \vdots \\ q_{l_1 1} \mathbf{W} & \dots & q_{l_1 l_2} \mathbf{W} \end{bmatrix} \quad (11)$$

معادله (۷) را می‌توان به صورت مجموع تنسورهای رتبه یک که از ضرب خارجی بردارهای مربوط به ستون‌های ماتریس‌های عامل به‌دست می‌آید نوشت. در معادله (۱۲) تنسور ضرایب به صورت مجموع تنسورهای رتبه یک تجزیه شده است.

$$\mathcal{A} \approx \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K q_{d_1 d_2 p} \mathbf{s}_{d_1} \circ \mathbf{t}_{d_2} \circ \mathbf{u}_p \quad (12)$$

که در آن، علامت \circ ضرب خارجی بین بردارها و $\mathbf{s}_{d_1} \in \mathbb{R}^{D_1}$ ؛ $d_1 = 1, \dots, D_1$ ، $\mathbf{t}_{d_2} \in \mathbb{R}^{D_2}$ ؛ $d_2 = 1, \dots, D_2$ ، $\mathbf{u}_p \in \mathbb{R}^P$ ؛ $p = 1, \dots, P$ است. تجزیه توکر مربوط به تنسور ضرایب به صورت محاسبه عنصری در معادله (۱۳) تعریف می‌شود.

$$a_{d_1 d_2 p} \approx \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K q_{d_1 d_2 p} s_{d_1 i} t_{d_2 j} u_{pk}; \quad d_1 = 1, \dots, D_1, d_2 = 1, \dots, D_2, p = 1, \dots, P \quad (13)$$

برای حل معادله (۷)، از روش تجزیه مقدار مفرد مرتبه بالاتر یا به‌اختصار HOSVD استفاده شده است [۲۲]. تابع هدفی که برای این منظور در نظر گرفته شده است به صورت معادله (۱۴) تعریف می‌شود.

$$\mathbf{Y}_i = (\mathbf{A}_{i1} x_{i1} + \mathbf{E}_{i1}) + (\mathbf{A}_{i2} x_{i2} + \mathbf{E}_{i2}) + \dots + (\mathbf{A}_{iP} x_{iP} + \mathbf{E}_{iP}), \quad i = 1, \dots, N \quad (1)$$

که در آن $\mathbf{E}_{ij} \in \mathbb{R}^{D_1 \times D_2}$ ؛ $i = 1, \dots, N, j = 1, \dots, P$ خطای تصادفی عنصر λ م در نمونه λ م که غیرقابل کنترل است، در نظر گرفته می‌شود. حالت ساده شده معادله (۱) به صورت معادله (۲) بازنویسی می‌شود.

$$\mathbf{Y}_i = \sum_{j=1}^P \mathbf{A}_{ij} x_{ij} + \sum_{j=1}^P \mathbf{E}_{ij}, \quad i = 1, \dots, N \quad (2)$$

مجموعه ضرایب تنسوری $\mathbf{A}_{ij} \in \mathbb{R}^{D_1 \times D_2}$ ؛ $i = 1, \dots, N, j = 1, \dots, P$ را با یک تنسور مرتبه بالاتر و به صورت $\mathcal{A} \in \mathbb{R}^{D_1 \times D_2 \times P}$ در نظر می‌گیریم.

در حالت کلی، تنسور مرتبه R به صورت $\mathcal{T} \in \mathbb{R}^{F_1 \times F_2 \times \dots \times F_R}$ نمایش داده می‌شود که در آن F_r نشان دهنده‌ی بُعد حالت r تنسور \mathcal{T} است. معادله (۲) را می‌توان به صورت مدل رگرسیون تنسوری (۳) در نظر گرفت.

$$\mathbf{Y}_i = \mathcal{A} \times_3 \mathbf{X}_i + \mathbf{E}_i, \quad i = 1, \dots, N \quad (3)$$

که در آن عبارت $\mathcal{A} \times_3 \mathbf{X}_i$ ضرب تنسور $\mathcal{A} \in \mathbb{R}^{D_1 \times D_2 \times P}$ از طرف بُعد سوم با ماتریس $\mathbf{X}_i \in \mathbb{R}^{P \times 1}$ ؛ $i = 1, \dots, N$ است؛ همچنین، عبارت $\mathbf{E}_i = \sum_{j=1}^P \mathbf{E}_{ij}$ یعنی مجموع خطاهای تصادفی موجود می‌باشد.

در حالت کلی، ضرب تنسور $\mathcal{T} \in \mathbb{R}^{F_1 \times F_2 \times \dots \times F_R}$ از طرف بُعد r با یک ماتریس $\mathbf{W} \in \mathbb{R}^{D \times F_r}$ به صورت $\mathcal{T} \times_r \mathbf{W}$ نمایش داده می‌شود که خروجی آن تنسوری با بُعد $F_1 \times \dots \times F_{r-1} \times D \times F_{r+1} \times \dots \times F_R$ به‌دست می‌آید. چگونگی محاسبه ضرب \mathcal{T} حالت مربوط به تنسور و ماتریس ذکر شده به صورت عنصری در معادله (۴) تعریف می‌شود.

$$(\mathcal{T} \times_r \mathbf{W})_{f_1, \dots, f_{r-1}, d, f_{r+1}, \dots, f_R} = \sum_{f_r=1}^{F_r} \mathcal{T}_{f_1, \dots, f_{r-1}, f_r, \dots, f_R} W_{df_r} \quad (4)$$

به‌منظور سادگی مفهوم چگونگی محاسبه معادله (۴)، حالت تنسوری و چگونگی حل آن در معادله (۵) ارائه شده است.

$$\mathcal{G} = \mathcal{T} \times_r \mathbf{W} \Leftrightarrow \mathbf{G}_{(f)} = \mathbf{W} \mathbf{T}_{(f)} \quad (5)$$

که در آن $\mathbf{G}_{(f)}$ و $\mathbf{T}_{(f)}$ به ترتیب مدل ماتریس شده حالت f تنسورهای \mathcal{G} و \mathcal{T} هستند که در آن‌ها عنصر تنسوری (f_1, f_2, \dots, f_R) به عنصر ماتریسی (f_r, d) تبدیل می‌شود.

حالت تنسوری شده معادله (۳) در معادله (۶) تعریف می‌شود.

$$\mathbf{y} = \mathcal{A} \times_3 \mathbf{X} + \boldsymbol{\varepsilon} \quad (6)$$

که در آن $\mathbf{y} \in \mathbb{R}^{D_1 \times D_2 \times N}$ و $\boldsymbol{\varepsilon} \in \mathbb{R}^{D_1 \times D_2 \times N}$ به ترتیب داده‌های متغیر پاسخ و خطای تصادفی را نشان می‌دهند. همچنین، عبارت $\mathbf{X} \in \mathbb{R}^{P \times N}$ ماتریس متغیرهای مستقل است که همه بردارهای \mathbf{X}_i را در یک ماتریس ترکیب کرده است.

در معادله (۶)، هدف پیدا کردن مقدار \mathcal{A} از روی داده‌های جمع‌آوری شده است که حل مستقیم آن با استفاده از روش کمینه‌سازی تابع احتمال منفی مخصوصاً در مسائلی که ابعاد تنسوری

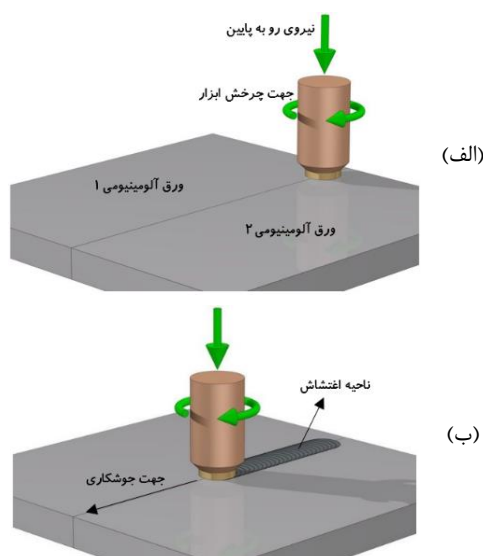
بررسی شرط همگرایی: در هر تکرار k ، متوقف می‌شود اگر برای تمام $n = 1, \dots, N$ ، مقدار $\|V_k^{(n)} - V_{k-1}^{(n)}\|$ کوچکتر از یک عدد کوچک مانند ϵ باشد. برگرداندن مقادیر مربوط به تانسور هسته‌ای (Q) و ماتریس‌های عامل ($V^{(1)}, V^{(2)}, \dots, V^{(N)}$)

شکل (۱): شبه‌کد مربوط به الگوریتم HOSVD

بنابراین، با استفاده از شبه‌کد ارائه شده در شکل (۱) و محاسبه تانسور ضرایب \mathcal{A} ، به راحتی می‌توان مقدار متغیر پاسخ تانسوری \mathcal{Y} را با استفاده از مقادیر عددی متغیر مستقل ورودی \mathbf{X} محاسبه کرد.

۴. مثال انگیزشی و مطالعات شبیه‌سازی

جوشکاری اصطکاکی اغتشاشی یک روش نوین جوشکاری در حالت جامد است که برای اتصال ورق‌های آلومینیومی و فولادی پرکاربرد در صنایع مهمی همچون ساخت تجهیزات هوافضا، خودروسازی، هواپیمایی، راه‌آهن، مترو، کشتی‌سازی، سکوهای دریایی، کامپیوترها و رباتیک استفاده می‌شود. در این روش یک ابزار استوانه‌ای شکل غیرمصرفی به یک دستگاه فرز مانند متصل شده و به کمک آن سرعت دورانی پیدا می‌کند. ابزار در حال چرخش با یک نیرویی به پایین فشار داده می‌شود تا وارد درز اتصال بین دو قطعه‌ای که قرار است بهم جوش داده شوند، بشود. لبه قطعات بر اثر حرارت ایجاد شده از اصطکاک بین ابزار و قطعه، خمیری شده و ابزار در حال چرخش، مواد خمیری را به هم زده و با سرعت پیشروی مناسب جلو می‌رود. ناحیه‌ای که ابزار در آنجا در حال حرکت است را اصطلاحاً ناحیه اغتشاش می‌گویند. پس از عبور ابزار، این ناحیه خنک شده و دو قطعه به هم متصل می‌شوند که در این صورت عمل جوشکاری بدون ذوب و بدون ماده پرکننده انجام می‌شود. شکل (۲) مراحل جوشکاری اصطکاکی اغتشاشی را نشان می‌دهد.



$$\text{Minimize } \|\mathcal{A} - \widehat{\mathcal{A}}\|^2 \Rightarrow \text{Minimize } \|\mathcal{A} - Q \times_1 \mathbf{S} \times_2 \mathbf{T} \times_3 \mathbf{U}\|^2 \quad (14)$$

که در آن علامت $\|\mathcal{T}\|$ معرف مفهومی به نام نرم تانسور $\mathcal{T} \in \mathbb{R}^{F_1 \times F_2 \times \dots \times F_R}$ است که به صورت معادله (۱۵) محاسبه می‌شود.

$$\|\mathcal{T}\| = \sqrt{\sum_{f_1=1}^{F_1} \sum_{f_2=1}^{F_2} \dots \sum_{f_R=1}^{F_R} t_{f_1 f_2 \dots f_R}^2} \quad (15)$$

در معادله (۱۴) تابع هدف به گونه‌ای تعریف شده است که اختلاف بین تانسور واقعی \mathcal{A} و مقدار تخمین زده شده آن $\widehat{\mathcal{A}}$ حداقل شود. معادله (۱۴) با معادله (۱۶) هم‌ارز است و می‌توان به جای همدیگر استفاده نمود.

$$\text{Maximize } Q = \mathcal{A} \times_1 \mathbf{S}' \times_2 \mathbf{T}' \times_3 \mathbf{U}' \quad (16)$$

تابع هدف معادله (۱۴) یا معادله (۱۶) در کنار محدودیت‌های معادله‌های (۱۷)، (۱۸) و (۱۹) در نظر گرفته می‌شود که در این معادلات، I نشان‌دهنده‌ی ماتریس یکه است.

$$\mathbf{S}'\mathbf{S} = I \quad (17)$$

$$\mathbf{T}'\mathbf{T} = I \quad (18)$$

$$\mathbf{U}'\mathbf{U} = I \quad (19)$$

برخلاف روش تجزیه مقدار منفرد مرتبه پایین (روش SVD معمولی)، روش حل به صورت فرم بسته برای حل مسائل HOSVD وجود ندارد. برای حل HOSVD در واقعیت، مسأله به صورت تکراری حل می‌شود. به این صورت که تمامی ماتریس‌های عامل به جز یک ماتریس عامل ثابت در نظر گرفته شده و مقدار جواب ماتریس عاملی که ثابت در نظر گرفته نشده است محاسبه می‌شود. این روش برای تمام ماتریس‌های عامل دیگر تکرار می‌شود. شکل (۱) شبه‌کد ارائه شده توسط [۲۷] را نشان می‌دهد که برای حل مسأله HOSVD به صورت تکراری استفاده شده است.

مقداردهی اولیه برای تمامی ماتریس‌های عامل

$$V^{(n)} \in \mathbb{R}^{I_n \times k_n}; n = 1, \dots, N$$

تا رسیدن به همگرایی یا حداکثر تکرار، تکرار کنید

برای $n = 1, \dots, N$ انجام دهید

• محاسبه تانسور G

$$G = \mathcal{A} \times_1 V^{(1)'} \dots \times_{n-1} V^{(n-1)'} \times_{n+1} V^{(n+1)'} \dots \times_N V^{(N)'}$$

بعد مربوط به تانسور ایجاد شده به صورت

$$\mathbb{R}^{k_1 \times \dots \times k_{n-1} \times I_n \times k_{n+1} \times \dots \times k_N} \text{ است}$$

• ماتریسی کردن تانسور G

• اعمال SVD بر روی ماتریس به دست آمده

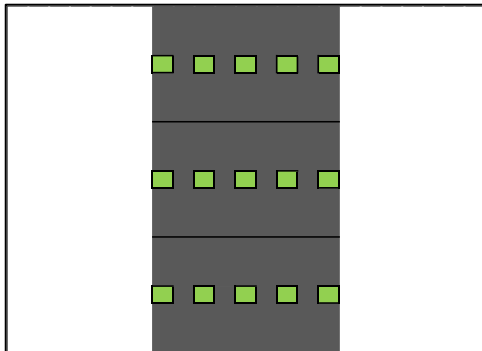
• انتخاب ماتریس عامل $V^{(n)}$ در صورتیکه

جولتر از k_n بردار منفرد چپ باشد

$$Q = \mathcal{A} \times_1 V^{(1)'} \times_2 V^{(2)'} \dots \times_N V^{(N)'}$$

پاسخ مدل در نظر گرفته می‌شود. هر قسمت ایجاد شده دارای دو متغیر مستقل سرعت دورانی و سرعت جوشکاری است که تعداد متغیر مستقل برابر با $P=6$ می‌شود. بنابراین، متغیرهای پاسخ تنسوری با $Y_i \in \mathbb{R}^{3 \times 5}; i = 1, \dots, 100$ و بردار متغیرهای مستقل مربوطه با $X_i \in \mathbb{R}^6; i = 1, \dots, 100$ نشان داده می‌شود.

خطای باقیمانده در برآورد متغیر پاسخ برای یکی از ۱۰۰ نمونه در نظر گرفته شده که تعداد ۱۵ عدد است (ضرب دو عدد ۳ در ۵ که توسط تنسور مرتبه دوم با ابعاد $D_1 = 3$ در $D_2 = 5$ ایجاد شده است)، در جدول (۱) نشان داده شده است.

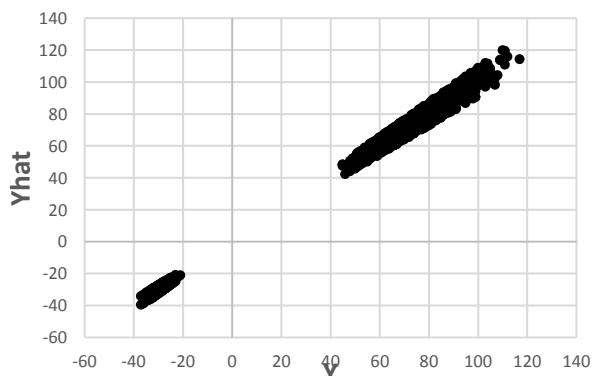


شکل (۳): نمای قطعه جوش داده شده به روش جوشکاری اصطکاکی اغتشاشی و سوراخ‌های ایجاد شده برای محاسبه تنش پسماند

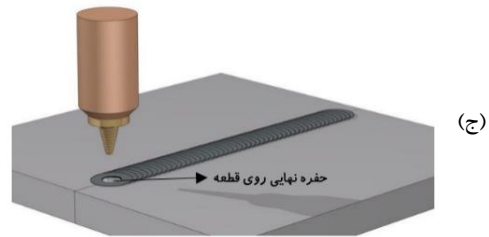
جدول (۱)

	-۲۸	۸۵	۵۲	۸۱	-۲۶
Y_i	-۳۰	۸۴	۵۷	۸۷	-۲۹
	-۲۴	۸۲	۵۱	۷۸	-۲۶
	-۲۹/۵	۸۱/۰	۵۳/۸	۸۴/۲	-۲۶/۱
\hat{Y}_i	-۳۱/۳	۸۸/۵	۶۰/۵	۹۴/۲	-۲۷/۵
	-۲۴/۷	۸۱/۳	۴۶/۸	۷۱/۱	-۲۶/۶
	۱/۵	۴/۰	-۱/۸	-۳/۲	۰/۱
Error	۱/۳	-۴/۵	-۳/۵	-۷/۲	-۱/۵
	۰/۷	۰/۷	۴/۲	۶/۹	۰/۶

نمودار پاسخ و برآورد آن برای ۱۰۰ نمونه که هر کدام دارای ۱۵ نقطه است در شکل (۴) نشان داده شده است.



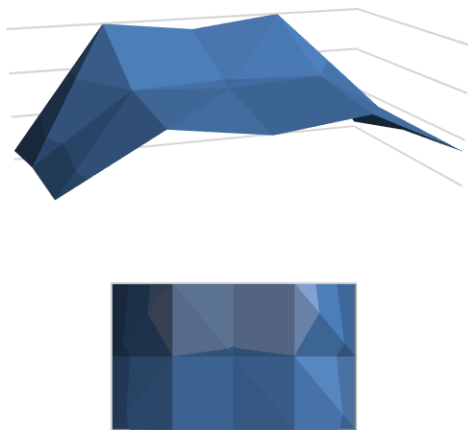
شکل (۴): مقدار متغیر پاسخ و برآورد آن برای ۱۰۰ نمونه



شکل (۲): (الف) آغاز جوشکاری و ورود ابزار جوشکاری بین درز و ورق مورد جوشکاری، (ب) حرکت ابزار در لبه دو قطعه و (ج) پایان جوشکاری و خروج ابزار جوشکاری

تنش پسماند در جوشکاری اصطکاکی اغتشاشی از اهمیت ویژه‌ای برخوردار است که در این مطالعه به‌عنوان متغیر پاسخ مدل رگرسیونی در نظر گرفته شده است. تنش پسماند تنشی است که پس از اعمال تنش خارجی مانند تغییرات حرارتی یا تغییرات سازه‌ای به یک جسم جامد، در آن باقی می‌ماند. مطالعات نسبتاً خوبی از روش مدل‌سازی رگرسیون برای پیدا کردن رابطه بین تنش پسماند و متغیرهای مستقل مختلف در روش جوشکاری اصطکاکی اغتشاشی برای دستیابی به محصولات با کیفیت بالا استفاده شده است. در اکثر این مطالعات متغیرهای پاسخ مربوط به تنش پسماند به‌صورت عددی در نظر گرفته شده‌اند و از مدل‌های رگرسیون ساده استفاده شده است. این درحالی است که تنش پسماند در سرتاسر قطعه وجود دارد و باید به‌صورت تنسور مرتبه دوم از اعداد در نظر گرفته شود. نادیده گرفتن ساختار تنسوری تنش پسماند ممکن است نتایج گمراه‌کننده‌ای را ایجاد نماید. در کتاب ارائه شده توسط لواسر و چن [۴۹] که بررسی جامعی در جوشکاری اصطکاکی اغتشاشی داشته‌اند، متغیرهای مستقل زیادی معرفی شده است که بر ایجاد تنش پسماند تأثیر می‌گذارند. دو متغیر سرعت دورانی (RS) و سرعت جوشکاری (WS) به‌عنوان دو متغیر مهمی که تأثیر زیادی بر ایجاد تنش پسماند دارد معرفی شده است. این دو متغیر در این مطالعه به‌عنوان دو متغیر مستقل مدل رگرسیونی در نظر گرفته شده است.

شبیه‌سازی برای ارزیابی عملکرد روش پیشنهادی برای پیش‌بینی کیفیت مسائلی که داده‌های آن به‌صورت تنسوری هستند انجام شده است. یک مجموعه آموزشی با اندازه نمونه ۱۰۰ عدد در نظر گرفته شده است که با روش جوشکاری اصطکاکی اغتشاشی جوش داده می‌شوند. تمام نمونه‌ها با استفاده از روش و پارامترهای فرآیندی یکسانی پردازش شده‌اند. قطر بین ۱۰ میلی‌متر، سرعت جوشکاری ۱۰ میلی‌متر بر ثانیه، سرعت چرخش ابزار ۸۰۰ دور در دقیقه، زاویه شیب ابزار و قطعه ۳ درجه و شکل پین ابزار استوانه‌ای مخروطی است. هریک از این نمونه‌ها به سه قسمت فرضی تقسیم می‌شوند. نمایی از یک نمونه جوش داده شده در شکل (۳) نشان داده شده است که در هریک از قسمت‌های ایجاد شده نقاطی برای محاسبه تنش پسماند به‌روش سوراخکاری در نظر گرفته شده است. اعداد به‌دست آمده تنش پسماند برای هریک از این مجموعه نقاط به‌صورت یک تنسور مرتبه دوم با ابعاد $D_1 = 3$ در $D_2 = 5$ یا به‌عبارتی $\mathbb{R}^{3 \times 5}$ ایجاد می‌شود که به‌عنوان متغیر



شکل (۶): تصویر سه‌بعدی تنسور ضرایب تخمین زده شده

تنسور فوق از ترکیب تنسور هسته‌ای (Q) و ماتریس‌های عامل $V(1)$ ، $V(2)$ و $V(3)$ به دست آمده است. مقدار تنسور هسته‌ای که به صورت یک تنسور است در ماتریس‌های زیر ارائه شده است.

$$Q_1 = \begin{bmatrix} 0.0009 & 0.0002 & -0.0059 & 0.0184 & 0.2802 \\ 0.0009 & 0.0067 & 0.0063 & 0.0115 & -0.0015 \\ 0.0012 & 0.0020 & 0.0041 & -0.0092 & -0.0007 \end{bmatrix}$$

$$Q_2 = \begin{bmatrix} 0.0012 & -0.0048 & 0.0288 & -0.0247 & -0.0015 \\ 0.0015 & 0.0014 & -0.0084 & 0.0427 & -0.1558 \\ -0.0002 & 0.0176 & -0.0007 & -0.0341 & -0.0161 \end{bmatrix}$$

$$Q_3 = \begin{bmatrix} -0.0007 & 0.0046 & 0.0217 & 0.0411 & -0.0037 \\ -0.0010 & -0.0019 & 0.0054 & 0.0265 & 0.0119 \\ -0.0003 & -0.0065 & -0.0077 & -0.0110 & -0.0560 \end{bmatrix}$$

$$Q_4 = \begin{bmatrix} -0.0004 & 0.0034 & 0.0100 & 0.0174 & -0.0009 \\ -0.0001 & -0.0022 & -0.0036 & 0.0106 & -0.0012 \\ -0.0006 & 0.0059 & -0.0032 & 0.0088 & 0.0196 \end{bmatrix}$$

$$Q_5 = \begin{bmatrix} 0.0003 & 0.0063 & -0.009 & 0.0054 & -0.0009 \\ 0.0007 & 0.0011 & -0.0091 & 0.0012 & 0.0004 \\ -0.0005 & 0.0013 & 0.0058 & -0.0108 & 0.0019 \end{bmatrix}$$

$$Q_6 = \begin{bmatrix} 0.0003 & 0.0012 & 0.0027 & -0.0075 & -0.0003 \\ 0.0004 & 0.0032 & 0.0041 & 0.0118 & 0.0052 \\ 0.0004 & -0.0029 & 0.0022 & -0.0054 & 0.0037 \end{bmatrix}$$

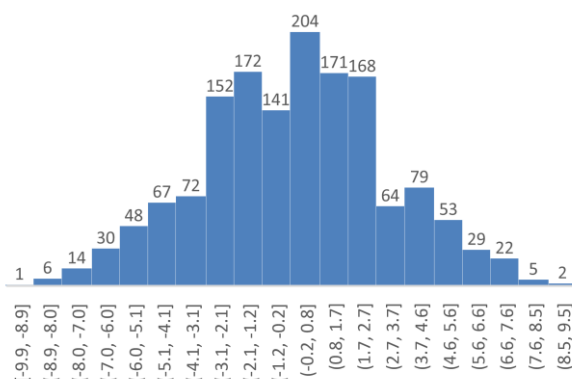
مقادیر مربوط به ماتریس‌های عامل به دست آمده نیز در ماتریس‌های زیر ارائه شده است.

$$V(1) = \begin{bmatrix} -0.551 & 0.452 & -0.701 \\ -0.125 & -0.876 & -0.466 \\ 0.825 & 0.17 & -0.539 \end{bmatrix}$$

$$V(2) = \begin{bmatrix} -0.899 & -0.189 & 0.036 & -0.374 & -0.124 \\ -0.02 & 0.039 & 0.742 & -0.117 & 0.658 \\ -0.434 & 0.484 & -0.147 & 0.703 & 0.248 \\ 0.024 & -0.088 & -0.652 & -0.304 & 0.688 \\ 0.051 & 0.849 & -0.01 & -0.51 & -0.128 \end{bmatrix}$$

$$V(3) = \begin{bmatrix} -0.124 & -0.062 & 0.307 & 0.602 & 0.436 & -0.578 \\ 0.712 & 0.201 & -0.563 & 0.081 & 0.04 & -0.358 \\ -0.191 & 0.564 & 0.045 & -0.578 & 0.516 & -0.207 \\ -0.379 & -0.613 & -0.564 & -0.24 & 0.214 & -0.242 \\ -0.278 & 0.187 & 0.035 & -0.132 & -0.704 & -0.611 \\ 0.469 & -0.476 & 0.517 & -0.471 & 0.02 & -0.251 \end{bmatrix}$$

با استفاده از تنسور ضرایب تخمین زده شده از داده‌های آموزشی به راحتی می‌توان دقت پیش‌بینی آن را بر روی داده‌های تست بررسی کرد. برای این منظور تعداد ۳۰ نمونه قطعه جوش داده شده جهت آزمون دقت پیش‌بینی در نظر گرفته شده است. مقدار خطای به دست



شکل (۵): هیستوگرام و شکل توزیع خطای برآورد متغیر پاسخ برای ۱۰۰ نمونه

هیستوگرام و شکل توزیع مربوط به خطای باقیمانده در برآورد متغیر پاسخ در شکل (۵) نشان داده شده است. داده‌های مربوط به خطاها نشان از این است که خطا دارای توزیع نرمال با میانگین صفر و واریانس ۱۰ است.

با استفاده از داده‌های آموزش و مدل ارائه شده در این مطالعه، تنسور ضرایب به صورت یک تنسور مرتبه سوم $\hat{A} \in \mathbb{R}^{3 \times 5 \times 6}$ تخمین زده می‌شود. تنسور ضرایب تخمین زده شده در ماتریس‌های زیر که برش‌هایی از صفحات تنسور ضرایب است نشان داده می‌شود.

$$\hat{A}_1 = \begin{bmatrix} -0.025 & 0.083 & 0.036 & 0.054 & -0.019 \\ 0.001 & 0.107 & 0.016 & 0.085 & -0.008 \\ 0.001 & 0.054 & 0.011 & 0.015 & -0.003 \end{bmatrix}$$

$$\hat{A}_2 = \begin{bmatrix} -0.017 & 0.048 & 0.035 & 0.048 & -0.017 \\ -0.011 & 0.032 & 0.017 & 0.031 & -0.009 \\ -0.011 & 0.028 & 0.010 & 0.030 & -0.005 \end{bmatrix}$$

$$\hat{A}_3 = \begin{bmatrix} -0.021 & -0.013 & 0.018 & -0.020 & -0.010 \\ -0.022 & 0.056 & 0.034 & 0.074 & -0.016 \\ 0.002 & 0.020 & 0.021 & 0.026 & -0.009 \end{bmatrix}$$

$$\hat{A}_4 = \begin{bmatrix} -0.008 & 0.016 & 0.021 & 0.020 & -0.010 \\ -0.016 & 0.037 & 0.035 & 0.038 & -0.018 \\ -0.006 & 0.014 & 0.020 & 0.014 & -0.010 \end{bmatrix}$$

$$\hat{A}_5 = \begin{bmatrix} 0.001 & 0.110 & 0.009 & 0.119 & -0.005 \\ -0.010 & -0.03 & 0.018 & -0.006 & -0.012 \\ -0.035 & 0.076 & 0.035 & 0.088 & -0.019 \end{bmatrix}$$

$$\hat{A}_6 = \begin{bmatrix} -0.006 & 0.016 & 0.010 & 0.015 & -0.005 \\ -0.009 & 0.020 & 0.018 & 0.020 & -0.009 \\ -0.016 & 0.040 & 0.035 & 0.040 & -0.018 \end{bmatrix}$$

باتوجه به این که تابع پاسخ ارائه شده در شبیه‌سازی این مطالعه به صورت تنسور مرتبه دوم یا همان ماتریس در نظر گرفته شده است، امکان حل دقیق مسأله وجود دارد که مقدار ضرایب تنسور به دست آمده با حل دقیق بادقت بسیار بالایی (2.022356×10^{-16}) شبیه به ضرایب تنسوری تخمین زده شده با مدل ارائه شده در این مطالعه می‌باشد. تصویر سه‌بعدی تنسور ضرایب تخمین زده شده در شکل (۶) ارائه شده است.

- Networks, vol. 15, no. 1, pp. 16–28, 2004.
- [11] H. Zou and T. Hastie, "Regression Shrinkage and Selection via the Elastic Net, with Applications to Microarrays," *J. R. Stat. Soc. Ser. B*, vol. 67, no. 1, pp. 301–320, 2003.
- [12] R. X. Liu, J. Kuang, Q. Gong, and X. L. Hou, "Principal component regression analysis with SPSS," *Comput. Methods Programs Biomed.*, vol. 71, no. 2, pp. 141–147, 2003.
- [13] H. Abdi, "Partial least square regression (PLS regression)," *Encycl. Res. methods Soc. Sci.*, vol. 6, no. 4, pp. 792–795, 2003.
- [14] S. P. and D. C. P. Debasish Basak, "Support vector regression," *Neural Inf. Process.*, vol. 11, no. 10, pp. 203–224, 2007.
- [15] P. A. Gutiérrez, M. Pérez-Ortiz, J. Sánchez-Monedero, F. Fernández-Navarro, and C. Hervás-Martínez, "Ordinal Regression Methods: Survey and Experimental Study," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 1, pp. 127–146, 2016.
- [16] E. L. Frome, "The Analysis of Rates Using Poisson Regression Models," *Biometrics*, vol. 39, no. 3, p. 665, 1983.
- [17] J. F. Lawless, "Negative binomial and mixed poisson regression," *Can. J. Stat.*, vol. 15, no. 3, pp. 209–225, 1987.
- [18] J. M. Ver Hoef and P. L. Boveng, "Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data?," *Ecology*, vol. 88, no. 11, pp. 2766–2772, 2007.
- [19] M. Lunn and D. McNeil, "Applying Cox Regression to Competing Risks," *Biometrics*, vol. 51, no. 2, p. 524, 1995.
- [20] E. B. H. T, J. I, and T. R, "Least angle regression," *Ann. Stat.*, vol. 32, no. 2, pp. 407–499, 2004.
- [21] J. O. Ramsay and B. W. Silvermann, *Functional Data Analysis*, vol. 40, no. 1, 1998.
- [22] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.
- [23] H. Fanae-e-t and J. Gama, "Tensor-based anomaly detection: An interdisciplinary survey," *Knowledge-Based Syst.*, vol. 98, pp. 130–147, 2016.
- [24] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.
- [25] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart-Young' decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, 1970.
- [26] P. Giordani and H. A. L. Kiers, "A review of tensor-based methods and their application to hospital care data," *Stat. Med.*, vol. 37, no. 1, pp. 137–156, 2018.
- [27] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, 2000.
- [28] Y. Zerehsaz, C. Shao, and J. Jin, "Tool wear monitoring in ultrasonic welding using high-order decomposition," *J. Intell. Manuf.*, no. 1999, 2016.
- [29] J. Zhou, C. K. Pang, Z. Zhong, and F. L. Lewis, "Tool Wear Monitoring Using Acoustic Emissions by Dominant-Feature Identification," vol. 60, no. 2, pp. 547–559, 2011.
- [30] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "MPCA: Multilinear Principal Component Analysis of Tensor Objects," *IEEE Trans. Neural Networks*, vol. 19, no. 1, pp. 18–39, 2008.
- [31] H. Yan, K. Paynabar, and J. Shi, "Image-Based Process Monitoring Using Low-Rank Tensor Decomposition,"

آمده برای متغیر پاسخ تنسوری با استفاده از مقدار پارامتر تخمین زده شده و متغیرهای عددی مستقل حدود ۰/۰۱۱ به دست آمد که نشان از دقت بالای روش پیشنهادی است.

۵. نتیجه و جمع‌بندی

در این مقاله، مدل رگرسیون تنسوری به منظور کمی‌سازی و پیش‌بینی کیفیت خروجی فرآیند بدون انجام آزمایش‌های مخرب توسعه یافته است. توسعه یک مدل ریاضی با متغیرهای مستقل عددی و متغیر پاسخ تنسوری و روش حل آن در حالیکه تعداد نمونه‌ها خیلی کم‌تر از تعداد متغیرها باشد به‌عنوان نوآوری اصلی این مقاله نسبت به روش‌های موجود در نظر گرفته می‌شود.

روش جوشکاری اصطکاکی اغتشاشی به‌عنوان مثال انگیزشی برای اعتبارسنجی نتایج در نظر گرفته شده است. مدل پیشنهادی با استفاده از الگوریتم تجزیه تنسور در نرم‌افزار R کدنویسی شده است.

در فرآیندهای مختلف، آزمایش‌های مخرب زیادی مانند آزمایش کشش، آزمایش خستگی، سطح تنش پسماند، آزمون ضربه و آزمون سختی جهت تعیین کیفیت فرآیند انجام می‌شوند. با استفاده از مدل پیشنهادی، خروجی فرآیندها را می‌توان با استفاده از متغیرهای تابعی پیش‌بینی کننده و بدون انجام آزمایش‌های مخرب با دقت خیلی خوب پیش‌بینی کرد.

جهت بهبود و توسعه مدل پیشنهادی برای مطالعات آتی دو موضوع ۱- در نظرگیری همبستگی بین متغیرها و درون متغیرها و ۲- در نظرگیری سایر متغیرهای مهم از جمله متغیرهای تابعی در کنار متغیرهای تنسوری پیشنهاد می‌گردد.

مراجع

- [1] A. Singh, S. Garg, R. Kaur, S. Batra, N. Kumar, and A. Y. Zomaya, "Probabilistic data structures for big data analytics: A comprehensive review," *Knowledge-Based Syst.*, vol. 188, 2020.
- [2] W. Karwowski, *Advances in Intelligent Systems and Computing 971 Advances in Manufacturing*. 2019.
- [3] D. Maulud and A. M. Abdulazeez, "A Review on Linear Regression Comprehensive in Machine Learning," *J. Appl. Sci. Technol. Trends*, vol. 1, no. 4, pp. 140–147, 2020.
- [4] R. C. Morilla, R. D. Omabe, C. J. S. Tolibas, E. E. C. C. Jr., and J. K. D. Treceñe, "Application of machine learning algorithms in predicting the performance of students in mathematics in the modern world," vol. 1, no. 1, pp. 49–57, 2020.
- [5] S. WEISBERG, *Applied Linear Regression*. John Wiley & Sons, 2005.
- [6] D. W. Hosmer and S. Lemeshow, "Applied Logistic Regression." John Wiley & Sons, pp. 161–164, 2000.
- [7] R. Koenker and K. F. Hallock, "Quantile regression," *J. Econ. Perspect.*, vol. 15, no. 4, pp. 143–156, 2001.
- [8] D. W. Marquardt and R. D. Sneec, "Ridge regression in practice," *Am. Stat.*, vol. 29, no. 1, pp. 3–20, 1975.
- [9] S. Asadzadeh1 and V. Mahdavi, "A Ridge-Regression Based Control Chart in the Presence of Multicollinearity," *J. Ind. Eng. Res. Prod. Syst.*, vol. 7, no. 15, pp. 241–253, 2020.
- [10] V. Roth, "The generalized LASSO," *IEEE Trans. Neural*

- Manufacturing Process Quality Control Using Sensor Data,” PhD thesis, Industrial and Operations Engineering, University of Michigan, 2017.
- [42] X. Zhang, L. Li, H. Zhou, and D. Shen, “Tensor Generalized Estimating Equations for Longitudinal Imaging Analysis,” (*Preprint*), 2014.
- [43] H. Zhou, L. Li, and H. Zhu, “Tensor regression with applications in neuroimaging data analysis,” *J. Am. Stat. Assoc.*, vol. 108, no. 502, pp. 540–552, 2013.
- [44] M. Hou and B. Chaib-Draa, “Hierarchical tucker tensor regression: Application to brain imaging data analysis,” *Image Process. (ICIP), 2015 IEEE Int. Conf.*, pp. 1344–1348.
- [45] D. Spencer, R. Guhaniyogi, and R. Prado, “Parsimonious Bayesian sparse tensor regression using the Tucker decomposition,” pp. 1–25, 2022.
- [46] R. Han, R. Willett, and A. R. Zhang, “An optimal statistical and computational framework for generalized tensor estimation,” *Am. Stat.*, vol. 50, no. 1, 2022.
- [47] B. Hao, A. Zhang, and G. Cheng, “Sparse and Low-Rank Tensor Estimation via Cubic Sketchings,” *IEEE Trans. Inf. Theory*, vol. 66, no. 9, pp. 5927–5964, 2020.
- [48] J. Zhou, W. W. Sun, J. Zhang, and L. Li, “Partially Observed Dynamic Tensor Response Regression,” *J. Am. Stat. Assoc.*, pp. 1–35, 2021.
- [49] D. Lohwasser and Z. Chen, *Friction stir welding: From basics to applications*. Woodhead Publishing, Cambridge, 2010.
- EEE Trans. Autom. Sci. Eng.*, vol. 12, no. 1, pp. 216–227, 2015.
- [32] T. G. Kolda, B. W. Bader, and S. N. Laboratories, “Tensor Decompositions and Applications,” *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.
- [33] E. E. Papalexakis, C. Faloutsos, and N. D. Sidiropoulos, “Tensors for Data Mining and Data Fusion: Models, Applications, and Scalable Algorithms,” *ACM Trans. Intell. Syst. Technol.*, vol. 8, no. 2, 2017.
- [34] H. Chun and S. Kele, “Sparse partial least squares regression for simultaneous dimension reduction and variable selection,” *J. R. Stat. Soc. Ser. B (Statistical Methodol.)*, vol. 72, no. 1, pp. 3–25, 2010.
- [35] I. S. Helland, “Partial least squares regression and statistical models,” *Scand. J. Stat.*, pp. 97–114, 1990.
- [36] X. Li, H. Zhou, and L. Li, “Tucker Tensor Regression and Neuroimaging Analysis,” *Stat. Biosci.*, pp. 1–26, 2013.
- [37] W. Guo, I. Kotsia, and I. Patras, “Tensor Learning for Regression,” *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 816–827, 2012.
- [38] M. Hou, “Tensor-based Regression Models and Applications,” PhD thesis, University Laval, Quebec, Canada, 2017.
- [39] B. M. Colosimo, “Modeling and monitoring methods for spatial and image data,” *Qual. Eng.*, vol. 30, no. 1, pp. 94–111, 2018.
- [40] E. F. Lock, “Tensor-on-tensor regression,” *J. Comput. Graph. Stat.*, p. (just-accepted), 2017.
- [41] Y. Zerehsaz, “Tensorial Data Modeling and Analysis for



DOI: 10.22084/ier.2022.25297.2057

Predicting the Quality of Production Processes with High-Dimensional Data via Tensor Regression

R. Noorossana^{1*}, F. Ramezankhani²

¹ Professor, Department of Industries, Faculty of Industrial Engineering, University of Science and Technology, Tehran, Iran

² PhD student, Department of Industries, Faculty of Industrial Engineering, University of Science and Technology, Tehran, Iran

ARTICLE INFO

Article history:

Received 5 October 2021

Accepted 18 December 2021

Keywords:

Tensor regression
Tensor Analysis
Process Modeling
Complex Structured Data
Analysis.

ABSTRACT

Advances in modern computer technologies and measurement systems embedded in various fields, there is a significant growth in the volume, variety and velocity of the data produced, which is a rich source of information. Tensor data is one of the most important types of complex structured data with high dimensions that traditional methods are not able to solve this type of data. Tensor regression models have many applications in optimizing problems in which there are a number of independent numerical variables as input to the problem that can be changed to create a better output. In this study, a regression model is presented in which the independent variables are numerical and the response variable is tensor. A set of linear algebraic methods and tensor approaches are proposed in order to find patterns within a set of points in space and their relationship to process variables. Friction stir welding (FSW) is a motivational example of this study to confirm the results. The proposed model is coded in R software. The results showed that the proposed model has a good performance in predicting the output of processes by dynamically adjusting its parameters.

* Corresponding author. R. Noorossana
E-mail address: rassoul@iust.ac.ir