



# محاسبه ناحیه جواب‌های ممکن در روش تحلیل عاملی انتقال

## هدف تکراری

پایان نامه کارشناسی ارشد

کیوان پوررضا

اساتید راهنما: دکتر حمید عبداللهی

دکتر آزاده گلشن

مهر ۱۳۹۱

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

خوشا به حال گیاهان که عاشق نورند

دوست دارم تلاش‌هایم را در قالب هدیه‌ای هرچند کوچک  
تقدیم کنم به:

نگاه نگران مادرم

دستان مهربان پدرم

عزم راسخ اساتید گرانقدرم، آقای دکتر عبداللهی و خانم دکتر

آزاده گلشن

## تشکر و قدردانی

به سبک مرسوم نگارش پایان نامه، ناچارم با واژه ها بنویسم: متشکرم، قدردانی می کنم، سپاسگزارم و ... ولی خوب واقفم که آنچه تحریر می شود کجا و مکنونات دلم کجا. چه کنم جز اینکه امید ببندم به این که همگان آگاهند "ورای حد تقریر است ...".

در ابتدا از دوستان عزیزم خانم ها سمیرا بیرامی، مهسا اکبری، بتول حسن زاده، عاطفه ملک خطابی و آقایان عادل پرتوی، رحیم محمدی، صابر حسینی، عثمان اسکندری و محمد رضا سرایی که هر یک به نوعی با همکاری یاریم دادند، تشکر می کنم. از خانم دکتر فریبا نظری به خاطر راهنمایی های مهربانانه و بی دریغشان کمال تشکر و قدردانی را دارم

از اساتید بزرگوامر جناب آقای دکتر حمید عبدالهی و خانم دکتر آزاده گلشن که با دلگرمی ها و تشویق هایشان همواره راهنما و چراغ راه من بوده اند، سپاسگزار و متشکرم (هر چند تشکر کلمه گویایی در اینجا به نظر نمی رسد).

از اساتید گرانقدرم جنابان آقای دکتر محسن کمپانی زارع، دک تر حبیب الله کاظمی، دک تر محمد رفیعی و دک تر بهزاد حقیقی که در محضرشان کسب علم و تجربه کردم نیز کمال تشکر و قدردانی را دارم.

از برادر و خواهر مهربانم رضا و آیلا عزیز که همواره امید و گرما بخش زندگی من بوده اند نیز متشکرم و برایشان از خداوند آرزوی بهترین ها را دارم

در نهایت، حاصل این کار را به پدر و مادرم، پشتیبانان بی قید و شرط و همیشگی ام، خورشیدهایی که هرگاه دنیا برایم به تاریکی می گرایید روشنای راهم بودند و تنها ماوای امن من، تقدیم می کنم. امید که سپاس کوچک مرا پذیرا باشند.

از دوستان و همکاران گرامی که شاید نامشان در این فهرست از قلم افتاده است ولی در پیشبرد این پژوهش یاری ام داده اند، از صمیم قلب سپاسگزارم و برایشان بهترین آرزوها را دارم. و در آخر کو مجالی که سراسر همه تقریر کنم؟

## چکیده

تجزیه و تحلیل با روش‌های نرم به عنوان یکی از مهمترین روش‌های کمومتریکس در بررسی و تفکیک ماتریس داده به دو ماتریس غلظت و طیف خالص گونه‌های فعال در سیستم به کار برده می‌شوند. عدم نیاز به داشتن اطلاعات درباره مدل شیمیایی سیستم، کاربرد این روش‌ها در حضور گونه‌های فعال که در فرآیند مورد نظر شرکت ندارند گسترده می‌سازد. اما تفکیک ریاضی یک ماتریس داده با روش‌های آنالیز نرم، به طور ذاتی با دو نوع ابهام روبرو است که این ابهام‌ها باعث می‌شوند که تعداد زیادی جفت ماتریس داده به وجود آید که با غلظت‌ها و طیف‌های واقعی تفاوت دارند. این تفاوت‌ها ممکن است شدتی یا چرخشی باش‌د. در نتیجه به جای جواب منحصر به فرد دامنه‌ای از جواب به دست می‌آید. با اعمال محدودیت‌های مناسب از جمله غیرمنفی بودن طیف خالص و غلظت گونه‌ها، بسته بودن، تک‌قله‌ای بودن و یا سرهم‌زدن داده‌ها می‌توان دامنه جواب را در این روش‌ها محدودتر کرد.

همه روش‌های نرم بسته به تخمین اولیه مورد نیاز به یک حل ویژه داخل ناحیه ابهام چرخشی همگرا می‌شوند. روش تجزیه و تحلیل حل عوامل (RFA)، به عنوان یکی از روش‌های نرم از این قاعده مستثنی نیست.

در این روش بردارهای ویژه چرخیده شده به طور متوالی اصلاح می‌شوند تا از محدودیت‌ها تبعیت کنند و در نهایت تابع مجموع مربعات را در فرآیند بهینه‌سازی غیرخطی کمینه کنند.

یکی دیگر از روش‌های مدل‌سازی نرم که در تجزیه و تحلیل ماتریس داده‌ها کاربرد روزافزون پیدا کرده است روش تحلیل عاملی انتقال هدف تکراری (ITTFA) می‌باشد.

روش ITTFA نیز مانند سایر روش‌های مدل‌سازی نرم دارای ابهام چرخشی می‌باشد. در این تحقیق ما بر این باور هستیم که با توجه به الگوریتم روش ITTFA، چون در این روش محدودیت‌ها تنها در یک فضا اعمال می‌شود و

فضای دیگر فارغ از محدودیت است ابهام چرخشی در این روش بیشتر از سایر روش‌های نرم است که در آنها به هر دو فضا توجه می‌شود و در هر دو فضا اعمال محدودیت می‌شود.

با توجه به این موضوع که ابهام چرخشی در روش ITTFA شدیدتر از سایر روش‌های نرم می‌باشد بنابراین برخی از روش‌هایی که در متون علمی از آن‌ها برای محاسبه تخمین اولیه برای روش ITTFA استفاده می‌شود حل قابل قبولی توسط روش ITTFA ارائه نمی‌دهد. از جمله این روش‌ها تجزیه و تحلیل تصویرسازی متعامد (OPA) و روش استفاده‌ی آسان فعل و انفعالات خود مدل ساز برای تجزیه و تحلیل مخلوط (SIMPILISMA) می‌باشد. هر دو این روش‌ها تخمین اولیه ای که ارائه می‌دهند خالص ترین ردیف یا ستون ماتریس داده را ارائه می‌دهند. ردیف‌ها یا ستون‌های ماتریس داده حاصل ترکیب خطی اجزای واقعی سیستم می‌باشد و جزو ناحیه جواب ITTFA می‌باشد اما پروفایل غلظتی یا طیفی مربوطه اش که بوسیله حداقل مربعات محاسبه می‌شود ترکیب خطی اجزاء واقعی سیستم نمی‌باشد و جزو ناحیه جواب روش‌های دیگر مانند RFA نمی‌باشد بنابراین روش‌های مانند OPA و SIMPILISMA نمی‌تواند روش مناسبی برای محاسبه تخمین اولیه برای روش ITTFA باشد.

و در نهایت با توجه به این که پروفایل‌های غلظتی حاصل از داده‌های کروماتوگرافی تک‌قله‌ای می‌باشد این واقعیت را می‌توان به صورت یک محدودیت دیگر علاوه بر محدودیت غیر منفی بودن پروفایل‌ها در روش ITTFA می‌توان به کار برد و در نهایت می‌توان نتیجه گرفت که اگرچه ناحیه جواب‌های ITTFA محدودتر می‌شود اما باز ابهام چرخشی در این روش شدیدتر از سایر روش‌های نرم می‌باشد.

# فهرست

..... ۱	فصل اول
..... ۱	مقدمه
..... ۱	۱.۱ مقدمه
..... ۲	۱.۲ کمومتریکس
..... ۲	۱.۳ داده ها
..... ۳	۱.۴ تجزیه و تحلیلی داده ها
..... ۳	۱.۵ روش های مدل سازی نرم
..... ۴	۱.۵.۱ تحلیلی عاملی
..... ۵	۱.۵.۲ تجزیه ی مقادیری منفرد (SVD)
..... ۶	۱.۵.۳ روش های حل غیری تکراری:
..... ۷	۱.۵.۳.۱ تجزیه و تحلیلی تکاملی عوامل EFA
..... ۷	۱.۵.۴ روش های تکراری
..... ۸	۱.۵.۴.۱ تجزیه و تحلیلی حل منحرفی های چند متغیره - حداقل مربعات متناوب (MCR-ALS)
..... ۱۰	۱.۵.۴.۲ تجزیه و تحلیلی فاکتوری تبدیلی هدف
..... ۱۰	۱.۵.۴.۳ تحلیلی عاملی انتقال هدف تکراری
..... ۱۳	۱.۵.۴.۴ تجزیه و تحلیلی عامل تفکیکی
..... ۱۴	۱.۶ جستجوی سوزنی
..... ۱۶	۱.۷ تجزیه و تحلیلی تصویری سازی متعامد

.....۱۷	۱.۸ روش استفاده‌ی آسان فعل و انفعالات خود مدل سازی برای تجزیه و تحلیلی مخلوط
.....۱۷	۱.۹ ابهام در روش های نرم:
.....۱۹	۱.۹.۱ قی منفی نبودن
.....۱۹	۱.۹.۲ قی تک قله‌ای بودن
.....۱۹	۱.۹.۳ قی بسته بودن
.....۱۹	۱.۹.۴ محدودیت پروفای شناخته شده
.....۱۹	۱.۹.۵ محدودیت مرتبه‌ی موضعی
.....۲۰	۱.۹.۶ محدودیت گزینش پذیری
.....۲۰	۱.۱۰ محاسبه ابهام چرخشی در سیستم‌های دوجزئی
.....۲۳	۱.۱۱ محاسبه ابهام چرخشی در سیستم‌های سه‌جزئی
.....۲۷	فصل دوم
.....۲۷	مروری بر تحقیقات گذشته
.....۲۷	۲.۱ تاریخچه‌ای از روش‌های محاسبه محدودده جواب‌های ممکن
.....۳۰	۲.۲ تاریخچه‌ای از روش تحلیلی عامی انتقال هدف تکراری (ITTFA)
.....۳۶	فصل سوم
.....۳۶	داده های شی سازی شده
.....۳۶	۳.۱ آزمایش‌های عددی
.....۳۶	۳.۲ کروماتوگرافی منجری گوسین
.....۳۷	۳.۲.۱ شی سازی مخلوط دوجزئی بک های کروماتوگرافی با هم‌پوشانی:



.....۳۸..... ۳.۲.۲ شیپ سازی مخلوط سه تایی، بیک های کروماتوگرافی با هم پوشانی:

.....۴۰..... ۳.۳ نرم افزار:

.....۴۱..... فصل چهارم

.....۴۱..... بحث و نتیجه گیری

.....۴۳..... ۴.۱ محاسبه ابهام چرخشی در روش تجزیه و تحلیلی حل عوامل با نرمال کردن قطری ماتریس چرخشی T

.....۴۳..... ۴.۱.۱ محاسبه ابهام چرخشی برای سیستم دوجزئی در روش تجزیه و تحلیلی حل عوامل

.....۴۳..... ۴.۱.۲ محاسبه ابهام چرخشی در روش تجزیه و تحلیلی حل عوامل با نرمال کردن ستونی ماتریس چرخشی T برای سیستم

.....۴۵..... دو جزئی

.....۴۹..... ۴.۱.۳ ابهام چرخشی در روش تحلیلی عاملی انتقال هدف تکراری در سیستم دوجزئی کروماتوگرافی

.....۴۹..... ۴.۱.۴ مقایسه ابهام چرخشی در دو روش تحلیلی عاملی انتقال هدف تکراری و تحلیلی عاملی در سیستم دو جزئی

.....۵۲..... کروماتوگرافی

.....۵۵..... ۴.۱.۵ محاسبه ابهام چرخشی در روش تجزیه و تحلیلی حل عوامل در سیستم های سه جزئی کروماتوگرافی

.....۵۹..... ۴.۱.۶ محاسبه ابهام چرخشی در روش تحلیلی عاملی انتقال هدف تکراری در سیستم سه جزئی کروماتوگرافی

.....۶۶..... ۴.۲ تاثیر روش های مختلف به عنوان تخمین اولی در جواب های حل شده توسط روش تحلیلی عاملی انتقال هدف

.....۶۶..... تکراری

.....۶۷..... ۴.۲.۱ روش تجزیه و تحلیلی تکاملی عوامل به عنوان تخمین اولی

.....۷۵..... ۴.۲.۲ اثر تخمین اولی جستجوی سوزنی در روش تحلیلی عاملی انتقال هدف تکراری

.....۸۲..... ۴.۲.۳ اثر تخمین اولی تجزیه و تحلیلی تصوی سازی عمود بر هم در روش تحلیلی عاملی انتقال هدف تکراری

.....۸۷..... ۴.۲.۴ اثر تخمین اولی روش استفاده ای آسان فعل و انفعالات خود مدل ساز برای تجزیه و تحلیلی مخلوط در روش تحلیلی

.....۸۷..... عاملی انتقال هدف تکراری

.....۸۹.

۴.۳ تاخیر محدودیت تک‌قله‌ای بودن روی ابهام چرخشی در روش تحلیلی عاملی انتقال هدف تکراری

.....۹۶.....

مراجع

# فهرست شکل‌ها

- شکل (۱.۱): تجزیه دو خطی ماتریس داده به پروفایل‌های غلظتی و طیفی ..... ۴
- شکل (۱.۲) دیاگرام روش تجزیه و تحلیل حل منحنی‌های چند متغیره - حداقل مربعات متناوب ..... ۹
- شکل (۱.۳): سوزن‌های مختلف با عناصر صفر و یک ..... ۱۵
- شکل (۱.۴) نقشه خطا،  $\log(\text{ssq})$  بر حسب  $t_{12}$  و  $t_{21}$  ..... ۲۳
- شکل (۱.۵) ناحیه جواب‌های ممکن برای سیستم دوجزئی بر اساس الگوریتم جستجوی شبکه‌ای ..... ۲۴
- شکل (۱.۶) دیاگرام محاسبه ابهام چرخشی در سیستم‌های سه‌جزئی بر اساس الگوریتم جستجوی شبکه‌ای ..... ۲۷
- شکل (۳.۱) پروفایل‌های غلظتی مربوط به داده‌های دوجزئی ..... ۳۸
- شکل (۳.۲) پروفایل‌های طیفی مربوط به داده‌های دوجزئی ..... ۳۹
- شکل (۳.۳) ماتریس داده‌ی جذبی سیستم دو جزئی ..... ۳۹
- شکل (۳.۴) پروفایل‌های غلظتی سیستم سه جزئی کروماتوگرافی ..... ۴۰
- شکل (۳.۵) پروفایل‌های طیفی سیستم سه جزئی کروماتوگرافی ..... ۴۰
- شکل (۳.۶) ماتریس داده‌ی جذبی سیستم سه جزئی ..... ۴۱
- شکل (۴.۱) ناحیه جواب‌های ممکن برای سیستم دوجزئی بر اساس الگوریتم جستجوی شبکه‌ای ..... ۴۵
- شکل (۴.۲) پروفایل‌های طیفی و غلظتی ناحیه جواب‌های ممکن برای سیستم دوجزئی بر اساس الگوریتم جستجوی شبکه‌ای ..... ۴۶
- شکل (۴.۳) ابهام چرخشی در روش تجزیه و تحلیل حل عوامل با نرمال کردن ستونی ماتریس چرخشی  $T$  ..... ۴۸
- شکل (۴.۴) ناحیه جواب‌های ممکن در روش تجزیه و تحلیل حل عوامل در فضای ردیف‌ها ..... ۴۹
- شکل (۴.۵) پروفایل‌های طیفی و غلظتی چرخیده شده در روش تجزیه و تحلیل حل عوامل ..... ۴۹
- شکل (۴.۶) محاسبه ابهام چرخشی و ناحیه جواب‌های ممکن در روش تحلیل عاملی انتقال هدف تکراری ..... ۵۲

- شکل (۴.۷) مقایسه ابهام چرخشی در دو روش تحلیل عاملی انتقال هدف تکراری و تجزیه و تحلیل حل عوامل ..... ۵۳
- شکل (۴.۸) مقایسه‌ی پروفایل‌های طیفی و غلظتی چرخیده شده در دو روش تحلیل عاملی انتقال هدف تکراری و تجزیه و تحلیل حل عوامل ..... ۵۵
- شکل (۴.۹) محاسبه ناحیه جواب‌های ممکن در سیستم‌های سه‌جزئی براساس الگوریتم جستجوی شبکه‌ای ..... ۵۷
- شکل (۴.۱۰) ناحیه جواب‌های ممکن و پروفایل‌های طیفی مربوطه برای جزء اول ..... ۵۸
- شکل (۴.۱۱) ناحیه جواب‌های ممکن و پروفایل‌های طیفی مربوطه برای جزء دوم ..... ۵۹
- شکل (۴.۱۲) ناحیه جواب‌های ممکن و پروفایل‌های طیفی مربوطه برای جزء سوم ..... ۶۰
- شکل (۴.۱۳) مقایسه ابهام چرخشی در دو روش تحلیل عاملی انتقال هدف تکراری و تجزیه و تحلیل حل عوامل در سیستم سه جزئی در فضای ردیف‌ها ..... ۶۳
- شکل (۴.۱۴) پروفایل‌های طیفی چرخیده شده در روش تحلیل عاملی انتقال هدف تکراری ..... ۶۵
- شکل (۴.۱۵) مقایسه ابهام چرخشی در بردارهای پایه فضای ستون‌ها (U-space) در دو روش تحلیل عاملی انتقال هدف تکراری و تجزیه و تحلیل حل عوامل (دایره‌های قرمز رنگ پروفایل‌های غلظتی واقعی در بردار پایه‌ی فضای ستون‌ها می‌باشد) ..... ۶۶
- شکل (۴.۱۶) پروفایل‌های غلظتی چرخیده شده در روش تحلیل عاملی انتقال هدف تکراری ..... ۶۷
- شکل (۴.۱۷) پروفایل‌های غلظتی تکاملی ..... ۶۸
- شکل (۴.۱۸) تخمین‌های اولیه غلظتی توسط روش تجزیه و تحلیل تکاملی عوامل در بردار پایه فضای ستون‌ها ..... ۷۰
- شکل (۴.۱۹) اولین پروفایل غلظتی بدست آمده توسط روش تجزیه و تحلیل تکاملی عوامل به عنوان تخمین اولیه در روش تحلیل عاملی انتقال هدف تکراری ..... ۷۱
- شکل (۴.۲۰) نحوه برداشتن روش تحلیل عاملی انتقال هدف تکراری در بردار پایه فضای ستون‌ها برای جزء اول ..... ۷۱

- شکل (۴.۲۱) پروفایل نهایی بدست آمده توسط روش تحلیل عاملی انتقال هدف تکراری برای جزء اول..... ۷۲
- شکل (۴.۲۲) دومین پروفایل غلطی بدست آمده توسط روش تجزیه و تحلیل تکاملی عوامل به عنوان تخمین اولیه در روش تحلیل عاملی انتقال هدف تکراری..... ۷۳
- شکل (۴.۲۳) نحوه گام برداشتن روش تحلیل عاملی انتقال هدف تکراری در بردار پایه فضای ستون‌ها برای جزء دوم..... ۷۳
- شکل (۴.۲۴) پروفایل نهایی بدست آمده توسط روش تحلیل عاملی انتقال هدف تکراری برای جزء دوم..... ۷۴
- شکل (۴.۲۵) سومین پروفایل غلطی بدست آمده توسط روش تجزیه و تحلیل تکاملی عوامل به عنوان تخمین اولیه در روش تحلیل عاملی انتقال هدف تکراری..... ۷۴
- شکل (۴.۲۶) نحوه گام برداشتن روش تحلیل عاملی انتقال هدف تکراری در بردار پایه فضای ستون‌ها برای جزء سوم..... ۷۵
- شکل (۴.۲۷) پروفایل نهایی بدست آمده توسط روش تحلیل عاملی انتقال هدف تکراری برای جزء سوم..... ۷۵
- شکل (۴.۲۸) اولین پروفایل غلطی بدست آمده توسط روش جستجوی سوزنی به عنوان تخمین اولیه در روش تحلیل عاملی انتقال هدف تکراری..... ۷۶
- شکل (۴.۲۹) نحوه گام برداشتن روش تحلیل عاملی انتقال هدف تکراری در بردار پایه فضای ستون‌ها برای جزء اول..... ۷۷
- شکل (۴.۳۰) پروفایل نهایی بدست آمده توسط روش تحلیل عاملی انتقال هدف تکراری برای جزء اول..... ۷۸
- شکل (۴.۳۱) دومین پروفایل غلطی بدست آمده توسط روش جستجوی سوزنی به عنوان تخمین اولیه در روش تحلیل عاملی انتقال هدف تکراری..... ۷۹
- شکل (۴.۳۲) نحوه گام برداشتن روش تحلیل عاملی انتقال هدف تکراری در بردار پایه فضای ستون‌ها برای جزء دوم..... ۸۰
- شکل (۴.۳۳) پروفایل نهایی بدست آمده توسط روش تحلیل عاملی انتقال هدف تکراری برای جزء دوم..... ۸۱

- شکل (۴.۳۴) سومین پروفایل غلظتی بدست آمده توسط روش جستجوی سوزنی به عنوان تخمین اولیه در روش تحلیل عاملی انتقال هدف تکراری..... ۸۲
- شکل (۴.۳۵) نحوه برداشتن روش تحلیل عاملی انتقال هدف تکراری در بردار پایه فضای ستون‌ها برای جزء سوم..... ۸۲
- شکل (۴.۳۶) پروفایل نهایی بدست آمده توسط روش تحلیل عاملی انتقال هدف تکراری برای جزء سوم..... ۸۳
- شکل (۴.۳۷) تخمین اولیه ، نحوه گام برداشتن روش تحلیل عاملی انتقال هدف تکراری در بردار پایه فضای ستون‌ها و جواب نهایی بدست آمده توسط الگوریتم برای جزء اول..... ۸۵
- شکل (۴.۳۸) تخمین اولیه ، نحوه گام برداشتن روش تحلیل عاملی انتقال هدف تکراری در بردار پایه فضای ستون‌ها و جواب نهایی بدست آمده توسط الگوریتم برای جزء دوم..... ۸۶
- شکل (۴.۳۹) تخمین اولیه ، نحوه گام برداشتن روش تحلیل عاملی انتقال هدف تکراری در بردار پایه فضای ستون‌ها و جواب نهایی بدست آمده توسط الگوریتم برای جزء سوم..... ۸۷
- شکل (۴.۴۰) تخمین های اولیه، نحوه گام برداشتن روش تحلیل عاملی انتقال هدف تکراری و جواب نهایی ارائه شده توسط روش تجزیه و تحلیل تکراری تبدیل هدف عوامل در بردار پایه فضای ستون‌ها..... ۸۹
- شکل (۴.۴۱) کروماتوگرام دو قله‌ای و مکان آن در بردار پایه فضای ستون‌ها..... ۹۰
- شکل (۴.۴۲) مقایسه‌ی ناحیه جواب‌های ممکن با اعمال قید منفی نبودن و ناحیه جواب‌های ممکن با اعمال قید منفی نپیدن و تک قله‌ای بودن در روش تجزیه و تحلیل حل عوامل..... ۹۱
- شکل (۴.۴۳) ناحیه جواب‌های ممکن در روش تحلیل عاملی انتقال هدف تکراری و تجزیه و تحلیل حل عوامل با اعمال محدودیت غیرمنفی بودن و تک‌قله‌ای بودن..... ۹۲

شکل (۴.۴۴) مقایسه‌ی ناحیه جواب‌های ممکن روش تحلیل عاملی انتقال هدف تکراری هنگامی که تنها قید منفی نبودن اعمال شده است و

هنگامی که هم قید منفی نبودن و هم قید تک‌قله‌ای بودن اعمال شده است ..... ۹۳

شکل (۴.۴۵) ناحیه جواب‌های الگوریتم **ITTFA** و **RFA** در بردار پایه‌ی فضای طیفی با اعمال محدودیت غیرمنفی بودنی تک‌قله‌ای

بودن ..... ۹۴

# فصل اول

## مقدمه

### ۱.۱ مقدمه

ما در دنیای مواد زندگی می‌کنیم. از مواد غذایی که زندگی ما را احاطه کرده است تا اشیاء مختلفی که در اطراف ما قرار دارند همگی از ماده ساخته شده اند. بنابراین تحقیقات بنیادی در مورد فعل و انفعالات مواد موضوعی بوده است که در طول تاریخ مورد بررسی قرار گرفته است. این تحقیقات در نهایت باعث بهبود یافتن کیفیت زندگی بشر و همچنین عمیق تر شدن دانسته‌های ما در مورد جهانی که در آن زندگی می‌کنیم و هر روزه با آن در تماس هستیم شده است [۱].

پیشرفت علم شیمی مطالعه‌ی مواد را ممکن ساخته است. نقش شیمی در زمینه‌هایی مثل کشاورزی، صنعت، بهداشت، صنعت مواد غذایی و حمل و نقل بسیار ارزشمند است.

توسعه‌ی سریع شیمی تجزیه هم زمان با پیشرفت علوم کامپیوتری بوده است. برای شیمیدان‌ها و به ویژه شیمیدان‌های تجزیه آسان تر شده است تا رایانه‌ها و روش‌های آماری و ریاضیات پیشرفته را در زمینه‌ی کارهای خود به کار برند. این چنین توسعه‌ها در تکنولوژی، به ویژه در زمینه دستگاه‌وری<sup>۱</sup> و قابلیت رایانه‌ها تحقیقات و نتایج را در شیمی تجزیه‌ی پیشرفته شتاب داده است. در نتیجه پیشرفت خیره کننده در دو زمینه‌ی شیمی و علم رایانه باعث پیدایش زمینه جدیدی به نام کمومتریکس<sup>۲</sup> شده است [۲-۴].

---

<sup>1</sup> Instrumentation

<sup>2</sup> Chemometrics



## ۱.۲ کمومتریکس

نام کموتریکس اولین بار به وسیله ی یک دانشمند جوان سوئدی به نام سونت ولد<sup>۱</sup> در اوایل دهه ی ۱۹۷۰ به کار برده شد [۵]. همکاری وی با شیمیدان تجزیه آمریکایی به نام بروس کوالسکی<sup>۲</sup> باعث پیدایش جامعه ی بین‌المللی کمومتریکس شد [۶].

کمومتریکس کاربرد ریاضیات و روش های آماری برای تجزیه و تحلیل داده های شیمیایی است. جامعه ی بین‌المللی کمومتریکس<sup>۳</sup> (ICS) تعریف زیر را برای کمومتریکس ارائه داده است:

کمومتریکس علمی است که در رابطه با اندازه گیری‌هایی است که روی سیستم های شیمیایی از طریق کاربرد ریاضیات و روش های آماری انجام شده است [۷]. توسعه ی روش های کمومتریکس بسیار وابسته به توسعه و استفاده از رایانه ها در علم شیمی است. فوآیندهای تجزیه ای شامل ریاضیات، آمار و روش های گرافیکی که قواعد منطقی را به کار می‌گیرند تا دانسته‌ها را راجع به پدیده های شیمیایی افزایش دهد.

اکثر روش های دستگاهی نیاز به تجزیه و تحلیل های ریاضی دارند تا داده های خام آزمایشگاهی را به نتایج معنی دار برای شیمیدان ها تبدیل کنند. روش های کمومتریکس در توسعه ی روش های مختلف برای جمع آوری اطلاعات مفید (بهبود دادن پارامترهای آزمایشگاهی<sup>۴</sup>، طراحی آزمایش<sup>۵</sup> و کالیبراسیون<sup>۶</sup>) کاربرد دارند [۸].

## ۱.۳ داده ها

داده های آزمایشگاهی می توان به سه دسته مختلف تقسیم بندی شود ، داده های تک متغیره<sup>۷</sup>، دو متغیره<sup>۸</sup> و داده‌های چند متغیره<sup>۹</sup> [۸].

---

<sup>1</sup>SVANTE WOLD

<sup>2</sup>BRUCE R.KOWALSKI

<sup>3</sup>International Chemometrics Society

<sup>4</sup>Optimization of lab parameter

<sup>5</sup>Experimental design

<sup>6</sup>Calibration

<sup>7</sup>Univariate

<sup>8</sup>Bivariate

<sup>9</sup>Multivariate

داده های تک متغیره یک سری اندازه گیری ساده است که به صورت تابعی از یک متغیر مستقل است. مثال این نوع داده ها اندازه گیری جذب در یک طول موج خاص در یک فرآیند سینتیکی<sup>۱</sup> است. داده های دو متغیره اندازه گیری است که به صورت تابعی از دو متغیر وابسته است مثال این نوع داده نیز مثل طیف جذبی بدست آمده بوسیله ی یک آشکار ساز آرایه دیودی<sup>۲</sup> در کروماتوگرافی مایع<sup>۳</sup> است که متغیرهای وابسته در آن زمان کروماتوگرافی و طول موج می باشد. داده های چندمتغیره شامل اندازه گیری های منفردی است که به صورت تابعی از چندین م متغیر (بیش از دو متغیر) است. برای مثال سینتیک اندازه گیری شده در طول موج های مختلف به صورت تابعی از دما، pH و...

#### ۱.۴ تجزیه و تحلیل داده ها

اسپکتروسکوپی جذبی یک روش قوی و در دسترس است، کمتر آزمایشگاهی وجود دارد که مجهز به دستگاه اسپکتروفوتومتر<sup>۴</sup> نباشد [۹].

داده های اسپکتروسکوپی می توانند به صورت ریاضی با استفاده از دو روش متفاوت زیر مدل شوند:

۱- آنالیز داده ها بر پایه ی مدل های سخت<sup>۵</sup>: این روش ها نیاز به مدل شیمیایی دارند تا بتوانند فرایندهای شیمیایی مورد نظر را شرح بدهند.

۲- آنالیز داده ها بر پایه مدل های نرم<sup>۶</sup>: در این روش ها نیاز به داشتن مدل شیمیایی نیست [۱۰].

#### ۱.۵ روش های مدل سازی نرم [۹]

یکی از روش های اصلی کمومتریکس که در تفسیر نتایج حاصل از فرایندهای تجزیه ای به کار می رود روش مدل سازی نرم می باشد. در این روش لازم نیست اطلاعاتی در زمینه ی رفتار فیزیکی - شیمیایی سیستم و یا حتی تعداد اجزای موجود در سیستم وجود داشته باشد. تنها فرض اساسی در این سیستم ها وجود یک رابطه

<sup>1</sup> Kinetic

<sup>2</sup> Photo diode array detector

<sup>3</sup> Liquid chromatography

<sup>4</sup> Spectrophotometer

<sup>5</sup> Hard Modeling

<sup>6</sup> Soft Modeling

خطی بین پاسخ مورد اندازه گیری (داده‌های آزمایشگاهی) و متغیرهای مستقل خاص است. عدم نیاز به داشتن اطلاعات درباره ی مدل شیمیایی سیستم، کاربرد این روش‌ها را در حضور گونه های فعال که در فرایند مورد نظر شرکت ندارند را گسترده تر می‌سازد. در تمامی روش‌های آنالیز نرم مدل دو خطی<sup>۱</sup> به کار می‌رود و هدف نهایی تمام این روش‌ها تفکیک ماتریس داده‌ها به پروفایل‌های خالص می‌باشد. در مورد مطالعات جذبی هدف به دست آوردن پروفایل‌های طیفی و غلظتی است [۱۱-۱۳].

$$D=C \times A \quad (1.2)$$

شکل (۱.۱): تجزیه دو خطی ماتریس داده به پروفایل‌های غلظتی و طیفی

در شکل (۱.۱) شماتیک ماتریس‌های **D**، **C**، **A**، به ترتیب داده‌ها، پروفایل‌های غلظتی اجزاء، و طیف خالص گونه‌ها می‌باشند.  $ns$  تعداد نمونه‌های آزمایش،  $n\lambda$  تعداد طول موج‌هایی که در آنها جذب اندازه‌گیری شده، و  $nc$  تعداد اجزای شیمیایی فعال در ناحیه‌ی طول موجی می‌باشد.

### ۱.۵.۱ تحلیل عاملی<sup>۲</sup> [۹]

تحلیل عاملی تفسیرهای فراوانی دارد و توافق کلی در مورد مفهوم آن وجود ندارد از نقطه نظر ریاضی تحلیل عاملی به سادگی به این صورت تعریف می‌شود که یک ماتریس به دو یا سه ماتریس تجزیه می‌شود. این چنین تعریفی کلی است و مفید نمی‌باشد ماتریس‌های حاصل از تحلیل عاملی دارای ویژگی‌های خاصی هستند مثلاً دارای ستون‌های عمود برهم هستند.

<sup>1</sup> Bilinear

<sup>2</sup>Factor Analysis

همان‌طور که گفته شد تحلیل عاملی تفسیرهای فراوانی دارد اما همگی این تفسیرهای بر اساس تجزیه ی مقادیر منفرد تعریف می‌شوند.

## ۱.۵.۲ تجزیه‌ی مقادیر منفرد<sup>۱</sup> (SVD)

تجزیه‌ی مقادیر منفرد یکی از پای‌دارترین و قوی‌ترین الگوریتم‌های موجود در جهان است که محاسبات عددی انجام می‌دهد. بنابراین از این الگوریتم برای محاسبات تحلیل عاملی استفاده می‌شود [۱۴-۱۵].

بر اساس این الگوریتم هر ماتریسی مانند  $\mathbf{D}$  به سه ماتریس طبق معادله‌ی زیر تجزیه می‌شود:

$$\mathbf{D} = \mathbf{U}\mathbf{\Lambda}\mathbf{V} \quad (۱.۳)$$

ابعاد ماتریس‌های فوق به ترتیب زیر است :

$\mathbf{D}$  یک ماتریس  $i \times j$  است که  $i \geq j$ ،  $\mathbf{U}$  یک ماتریس  $i \times j$  است در حالی که ماتریس‌های  $\mathbf{A}$  و  $\mathbf{V}$   $j \times j$  هستند

اولین هدف تحلیل عاملی تعیین تعداد ردیف‌ها یا ستون‌های مستقل خطی در ماتریس  $\mathbf{D}$  می‌باشد. اغلب به این تعداد ستون‌ها یا ردیف‌ها تعداد اجزاء یا عوامل (مرتبه‌ی ماتریس<sup>۲</sup>) سیستم تحت مطالعه گفته می‌شود. تجزیه‌ی مقادیر منفرد توسعه پیدا کرده است تا به این سوال آماری به طور خاص پاسخ دهد.

معادله‌ی (۱.۱) را به صورت زیر هم می‌توان نوشت:

$$\mathbf{D} = \sum_i^N \mathbf{U}_{:,i} \mathbf{\Lambda}_{i,i} \mathbf{V}_{i,:} \quad (۱.۴)$$

$\mathbf{U}$  از ویژه بردارهای<sup>۳</sup> معنی‌دار  $\mathbf{D}\mathbf{D}^t$  و  $\mathbf{V}$  از ویژه بردارهای معنی‌دار  $\mathbf{D}^t\mathbf{D}$  ساخته شده است.

ستون‌های ماتریس  $\mathbf{U}$  و ردیف‌های ماتریس  $\mathbf{V}$  برهم عمودند،  $\mathbf{U}^t\mathbf{U} = \mathbf{V}^t\mathbf{V} = \mathbf{I}$  (ماتریس یکسان<sup>۴</sup>).

<sup>1</sup> Singular Value Decomposition

<sup>2</sup> Rank

<sup>3</sup> Eigenvector

<sup>4</sup> Identity Matrix