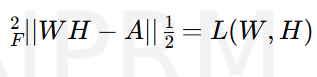
برای رسیدن به قوانین به‌روزرسانی در فاکتورگیری ماتریس غیرمنفی (NMF)، باید از روش‌هایی مانند \*\*کاهش گرادیان\*\* یا \*\*بهینه‌سازی با محدودیت‌های غیرمنفی\*\* استفاده کنیم. ابتدا مفهوم کاهش گرادیان را توضیح می‌دهیم و سپس به چگونگی رسیدن به قوانین ضربی می‌پردازیم.

**### کاهش گرادیان در NMF**

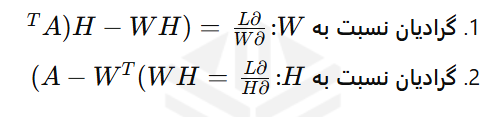
فرض کنید ماتریس اصلی داده‌ها \(A\) داریم که می‌خواهیم آن را به صورت حاصل‌ضرب \(A \approx W H\) بازسازی کنیم، جایی که \(W\) و \(H\) هر دو غیرمنفی هستند. تابع خطای ما معمولاً به صورت زیر تعریف می‌شود:



هدف کمینه‌کردن این تابع خطاست.

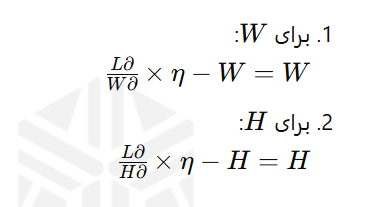
#### محاسبه گرادیان‌ها

برای کاهش تابع خطا، باید گرادیان‌های \(L(W, H)\) نسبت به \(W\) و \(H\) را پیدا کنیم:



#### کاهش گرادیان

با استفاده از کاهش گرادیان، به‌روزرسانی \(W\) و \(H\) به صورت زیر انجام می‌شود:



در اینجا \(\eta\) نرخ یادگیری است که مقدار کوچکی برای تنظیم سرعت یادگیری می‌باشد.

### مشکل در کاهش گرادیان

یک چالش بزرگ در این روش این است که ممکن است مقادیر \(W\) و \(H\) منفی شوند، در حالی که باید غیرمنفی باشند. برای حل این مشکل، دو رویکرد استفاده می‌شود:

1. \*\*تحمیل محدودیت غیرمنفی:\*\* پس از هر مرحله به‌روزرسانی، مقادیر منفی را به صفر تبدیل می‌کنیم.

2. \*\*قوانین ضربی (Multiplicative Update Rules):\*\* روشی که مستقیماً غیرمنفی‌بودن را تضمین می‌کند.

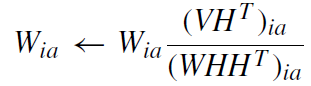
**### قوانین ضربی (Multiplicative Update Rules)**

قوانین ضربی با استفاده از تکنیک‌های بهینه‌سازی مقید (Constrained Optimization) به دست می‌آیند. ایده اصلی این است که به جای کاهش خطی گرادیان، وزن‌ها را به صورت ضربی به‌روزرسانی کنیم تا غیرمنفی‌بودن تضمین شود

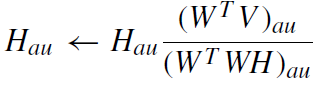
#### رسیدن به قوانین ضربی

برای این کار، از تکنیک \*\*ضریب لاگرانژ\*\* و روش‌هایی مانند الگوریتم KL (Kullback-Leibler) یا قاعده افزایش-کاهش استفاده می‌کنیم. نتیجه نهایی برای به‌روزرسانی \(W\) و \(H\) به این صورت است:

1. به‌روزرسانی \(W\):



2. به‌روزرسانی \(H\):



در این روش، مقادیر جدید \(W\) و \(H\) تنها به مقادیر مثبت وابسته‌اند، بنابراین غیرمنفی‌بودن آنها همیشه حفظ می‌شود.

### مقایسه کاهش گرادیان و قوانین ضربی



Overlapping Community Detection Using Non-Negative Matrix Factorization With Orthogonal and Sparseness Constraints

**### \*\*خلاصه مقاله: شناسایی جوامع هم‌پوشان با استفاده از فاکتورگیری ماتریس غیرمنفی با قیود متعامد و کمینه‌سازی پراکندگی\*\***

این مقاله به معرفی روشی جدید برای شناسایی جوامع هم‌پوشان در شبکه‌های پیچیده می‌پردازد که از \*\*فاکتورگیری ماتریس غیرمنفی (NMF)\*\* با قیود متعامد و کمینه‌سازی پراکندگی استفاده می‌کند. هدف اصلی روش ارائه‌شده، بهبود دقت و کارایی شناسایی جوامع هم‌پوشان و ارائه نتایجی دقیق‌تر در شبکه‌هایی است که ساختار جوامع در آن‌ها از پیش مشخص نیست.

**### \*\*۱. مقدمه\*\***

ساختار جوامع یکی از ویژگی‌های مهم شبکه‌های پیچیده است که می‌تواند خصوصیات و عملکردهای آن‌ها را منعکس کند. جوامع هم‌پوشان، به ویژه در شبکه‌های واقعی مانند شبکه‌های اجتماعی و شبکه‌های پروتئینی، اهمیت بیشتری دارند؛ زیرا گره‌ها می‌توانند به چندین جامعه تعلق داشته باشند.

NMF به‌عنوان یک روش کاهش بعد و استخراج ویژگی، در این تحقیق برای کشف جوامع استفاده شده است. مزیت این روش در مقایسه با روش‌های دیگر، توانایی آن در شناسایی جوامع بدون نیاز به مشخص بودن تعداد جوامع از پیش است.

**### \*\*۲. روش پیشنهادی: فاکتورگیری ماتریس غیرمنفی با قیود متعامد و پراکندگی \*\***

روش پیشنهادی مقاله از \*\*فاکتورگیری ماتریس غیرمنفی (NMF)\*\* استفاده می‌کند و آن را با قیود متعامد و پراکندگی بهبود می‌بخشد. این قیود به روش اضافه شده‌اند تا دقت و کیفیت شناسایی جوامع هم‌پوشان در شبکه‌ها بهبود یابد. در ادامه، جزئیات مربوط به این روش و قیود به تفصیل شرح داده می‌شود:

### \*\*۱. پیش‌پردازش ماتریس ویژگی\*\*

1. \*\*استفاده از الگوریتم خوشه‌بندی مارکوف (MCL):\*\*

- الگوریتم MCL یک روش سریع برای خوشه‌بندی گراف است که بر اساس شبیه‌سازی قدم تصادفی (random walk) عمل می‌کند.

- در این مقاله، MCL برای تکمیل اطلاعات ماتریس پراکنده استفاده می‌شود. فرآیندهای انبساط (Expansion) و تورم (Inflation) در این الگوریتم، احتمال اتصال بین گره‌ها را تغییر می‌دهند و روابط قوی‌تر میان گره‌های مرتبط را برجسته می‌کنند.

- نتیجه این فرآیند، تولید ماتریس اولیه‌ای است که تعداد تقریبی جوامع را نشان می‌دهد. این تعداد به عنوان ورودی برای NMF استفاده می‌شود.

### \*\*۲. فاکتورگیری ماتریس غیرمنفی\*\*

NMF ماتریس ویژگی \( V \) را به دو ماتریس غیرمنفی \( W \) و \( H \) تجزیه می‌کند:



- \*\*ماتریس \( W \):\*\*

- این ماتریس نشان‌دهنده عضویت گره‌ها در جوامع است. هر سطر ماتریس \( W \) وزن تعلق یک گره به جوامع مختلف را نشان می‌دهد.

- قیود متعامد روی این ماتریس اعمال می‌شود تا عضویت هر گره در جامعه‌ها به صورت واضح‌تری تعریف شود.

- \*\*ماتریس \( H \):\*\*

- این ماتریس نشان‌دهنده وزن ارتباط گره‌ها با جوامع مختلف است.

- قیود پراکندگی روی این ماتریس اعمال می‌شود تا تنها مقادیر کلیدی باقی بمانند و نویز حذف شود.

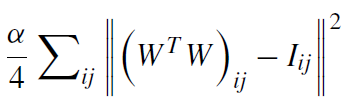
### \*\*۳. قیود اضافه‌شده برای بهبود دقت\*\*

#### \*\*الف) قید متعامد (Orthogonal Constraint)\*\*

- قید متعامد تضمین می‌کند که ستون‌های ماتریس \( W \) به صورت متعامد هستند.

- این قید باعث می‌شود که ویژگی‌های شبکه بهتر بازتاب داده شوند و عضویت گره‌ها در جوامع به طور واضح‌تر مشخص شود.

- این قید با اضافه کردن یک ترم به فرمول هدف مدل‌سازی می‌شود:



که در آن:

- \( \alpha \): وزن قید متعامد.

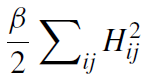
- \( I\_{ij} \): ماتریس هویت.

#### \*\*ب) قید پراکندگی (Sparsity Constraint)\*\*

- قید پراکندگی تضمین می‌کند که تنها تعداد محدودی از مقادیر ماتریس \( H \) غیرصفر باشند.

- این قید با کاهش نویز و کاهش زمان محاسباتی، کارایی روش را افزایش می‌دهد.

- ترم مرتبط با پراکندگی به فرمول هدف اضافه می‌شود:



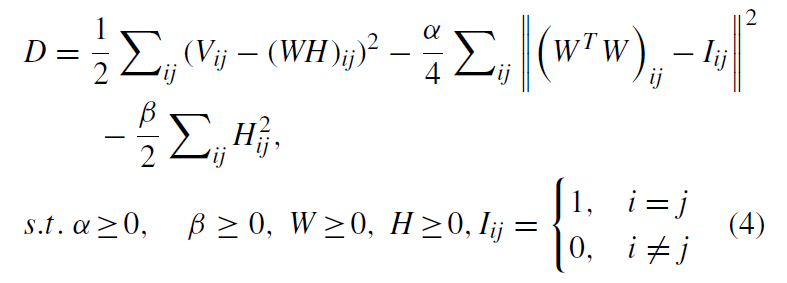
که در آن:

- \( \beta \): وزن قید پراکندگی.

---

### \*\*۴. فرمول هدف روش پیشنهادی\*\*

فرمول هدف برای NMF با قیود متعامد و پراکندگی به صورت زیر تعریف شده است:

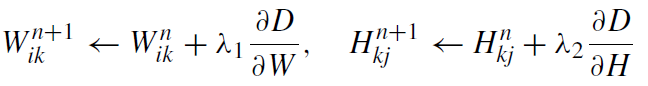


- هدف این فرمول، بهینه‌سازی ماتریس‌ها \( W \) و \( H \) برای بازسازی ماتریس \( V \) با دقت بالا است.

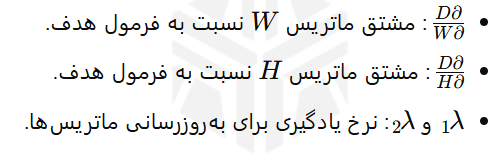
- قیدها تضمین می‌کنند که مقادیر ماتریس‌ها غیرمنفی بوده و ویژگی‌های شبکه به صورت بهینه بازتاب داده شوند.

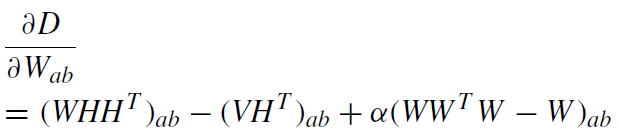
### \*\*۵. روش بهینه‌سازی و به‌روزرسانی ماتریس‌ها\*\*

ماتریس‌ها \( W \) و \( H \) به صورت تکراری به‌روزرسانی می‌شوند:

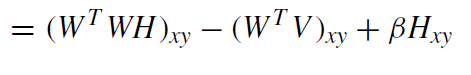


که در آن:





و اون یکی میشه:



---

### \*\*۶. مزایای روش پیشنهادی\*\*

1. \*\*شناسایی خودکار تعداد جوامع:\*\*

- روش پیشنهادی بدون نیاز به تعداد جوامع از پیش مشخص‌شده عمل می‌کند.

2. \*\*کاهش نویز و بهبود دقت:\*\*

- قیود متعامد و پراکندگی به بهبود دقت و کاهش اثر نویز کمک می‌کنند.

3. \*\*سازگاری با شبکه‌های هم‌پوشان:\*\*

- امکان شناسایی گره‌هایی که به چندین جامعه تعلق دارند.

### \*\*۷. ارتباط با جوامع هم‌پوشان\*\*

- پس از فاکتورگیری، هر گره بر اساس ماتریس‌های \( W \) و \( H \) به یک یا چند جامعه اختصاص داده می‌شود.

- اگر احتمال عضویت یک گره در بیش از یک جامعه بالا باشد، آن گره به عنوان گره هم‌پوشان شناسایی می‌شود.

### \*\*۳. ارزیابی و آزمایش‌ها\*\*

#### \*\*الف) معیارهای ارزیابی\*\*

1. \*\*رسانایی (Conductance):\*\* نسبت یال‌های داخل جامعه به یال‌های خروجی.

2. \*\*اطلاعات متقابل نرمال‌شده (NMI):\*\* دقت شناسایی جوامع در مقایسه با جوامع واقعی.

3. \*\*ماژولاریتی (Modularity):\*\* کیفیت تفکیک جوامع.

#### \*\*ب) شبکه‌های مصنوعی\*\*

این روش روی شبکه‌های مصنوعی با ۱۲۸ گره و ۱۰۲۴ یال آزمایش شد. نتایج نشان داد که با افزایش هم‌پوشانی، دقت روش NMFOSC نسبت به سایر الگوریتم‌ها کاهش کمتری داشت و در معیار NMI عملکرد بهتری نشان داد.

#### \*\*ج) شبکه‌های واقعی\*\*

آزمایش‌ها روی شبکه‌های واقعی (مانند شبکه Southern Women) نیز نشان داد که روش پیشنهادی عملکرد بهتری در شناسایی جوامع هم‌پوشان دارد و دقت آن از سایر الگوریتم‌های مقایسه‌شده بالاتر بود.

### \*\*۴. نتیجه‌گیری\*\*

روش NMFOSC با ترکیب خوشه‌بندی MCL و بهینه‌سازی رتبه‌بندی، امکان شناسایی جوامع هم‌پوشان را بدون نیاز به تعداد اولیه جوامع فراهم می‌کند. این روش به دلیل قیود متعامد و پراکندگی، قادر به استخراج ویژگی‌های اصلی شبکه و کاهش نویز است. نتایج آزمایش‌ها نشان می‌دهد که این الگوریتم در شبکه‌های مصنوعی و واقعی از دقت و کیفیت بالایی برخوردار است.