



دانشگاه حکیم بسزوری

دانشکده ریاضی و علوم کامپیوتر

پایان نامه برای دریافت درجه کارشناسی ارشد در رشته ریاضی کاربردی
گرایش تحقیق در عملیات

خوشه بندی طیفی برای قطعه بندی تصویر

استاد راهنما

دکتر محمود امین طوسی

استاد مشاور

دکتر یاسر علیزاده

پژوهشگر:

طیبه فیاض

شهریور ماه ۱۳۹۴



شماره:

باسمه تعالی

تاریخ: ۱۳۹۴ / ۶ / ۳

صورتجلسه دفاع از پایان نامه کارشناسی ارشد

با تلاوت آیاتی چند از کلام الله مجید جلسه دفاع از پایان نامه آقای / خانم طیبه فیاض دانشجوی رشته ریاضی کاربردی به شماره دانشجویی ۹۲۱۳۱۳۳۰۱۹ با عنوان:

خوشه بندی طیفی برای قطعه بندی تصویر

در ساعت مورخه ۱۳۹۴ / ۶ / ۳ در محل دانشکده ریاضی و علوم کامپیوتر تشکیل گردید. پس از استماع گزارش ارائه شده توسط دانشجو و استاد راهنما، هیات داوران و حاضران سنوالاتی را مطرح و آقای / خانم طیبه فیاض به دفاع از موضوع پرداخت و به سئوالات آنها پاسخ گفت. سپس پایان نامه توسط هیات داوران مورد ارزشیابی قرار گرفت و نمره ۱۹/۲۵ برابر درجه عالی برای آن تعیین گردید. به این ترتیب ضمن تصویب پایان نامه مزبور از این تاریخ آقای / خانم طیبه فیاض به عنوان کارشناس ارشد در رشته ریاضی کاربردی شناخته می شود.

ردیف	نام و نام خانوادگی	سمت	امضاء
۱	دکتر محمود امین طوسی	استاد راهنما	
۲	دکتر یاسر علیزاده	استاد مشاور	
۳	دکتر محمد علی پرتانیان	استاد داور	
۴	دکتر رحیمه پورخاندانی	نماینده تحصیلات تکمیلی	

مدیر گروه: دکتر عبدالله قلی زاده

عبدالله علی زاده

رونوشت:

۱. معاون محترم آموزشی و تحصیلات تکمیلی دانشگاه جهت اطلاع
۲. معاون محترم پژوهشی دانشگاه جهت اطلاع
۳. آموزش دانشکده ریاضی و علوم کامپیوتر جهت درج در پرونده دانشجو



سوگند نامه دانش آموختگان دانشگاه حکیم سبزواری

به نام خداوند جان و خرد کزین برتر اندیشه بر نگذرد

اینک که به خواست آفریدگار پاک، کوشش خویش و بهره گیری از دانش استادان و سرمایه های مادی و معنوی این مرز و بوم، توشه ای از دانش و خرد گردآورده ام، در پیشگاه خداوند بزرگ سوگند یاد می کنم که در به کارگیری دانش خویش، همواره بر راه راست و درست گام بردارم. خداوند بزرگ، شما شاهدان، دانشجویان و دیگر حاضران را به عنوان داورانی امین گواه می گیرم که از همه دانش و توان خود برای گسترش مرزهای دانش بهره گیرم و از هیچ کوششی برای تبدیل جهان به جایی بهتر برای زیستن، دریغ نورزم. پیمان می بندم که همواره کرامت انسانی را در نظر داشته باشم و ممنوعان خود را در هر زمان و مکان تا سر حد امکان یاری دهم. سوگند می خورم که در به کارگیری دانش خویش به کاری که باراه و رسم انسانی، آیین پرهیزگاری، شرافت و اصول اخلاقی برخاسته از ادیان بزرگ الهی، به ویژه دین مبین اسلام، مبادنت دارد دست نیازم. همچنین در سایه اصول جهان شمول انسانی و اسلامی، پیمان می بندم از هیچ کوششی برای آبادانی و سرافرازی میهن و هم میهنانم فروگذاری نکنم و خداوند بزرگ را به یاری طلبم تا همواره در پیشگاه او و در برابر وجدان بیدار خویش و ملت سرافراز، بر این پیمان تا ابد استوار بمانم.

نام و نام خانوادگی: طیبه فیاض

تاریخ و امضا:

تأییدیه‌ی صحت و اصالت نتایج

باسمه تعالی

اینجانب طیبه فیاض به شماره دانشجویی ۹۲۱۳۱۳۳۰۱۹ دانشجوی رشته ریاضی کاربردی مقطع تحصیلی کارشناسی ارشد تأیید می‌نمایم که کلیه‌ی نتایج این پایان‌نامه حاصل کار اینجانب و بدون هرگونه دخل و تصرف است و موارد نسخه برداری شده از آثار دیگران را با ذکر کامل مشخصات منبع ذکر کرده‌ام. در صورت اثبات خلاف مندرجات فوق، به تشخیص دانشگاه مطابق با ضوابط و مقررات حاکم (قانون حمایت از حقوق مؤلفان و مصنفان و قانون ترجمه و تکثیر کتب و نشریات و آثار صوتی، ضوابط و مقررات آموزشی، پژوهشی و انضباطی ...) با اینجانب رفتار خواهد شد و حق هرگونه اعتراض در خصوص احقاق حقوق مکتسب و تشخیص و تعیین تخلف و مجازات را از خویش سلب می‌نمایم. در ضمن، مسئولیت هرگونه پاسخگویی به اشخاص اعم از حقیقی و حقوقی و مراجع ذی صلاح (اعم از اداری و قضایی) به عهده‌ی اینجانب خواهد بود و دانشگاه هیچ‌گونه مسئولیتی در این خصوص نخواهد داشت.

نام و نام خانوادگی: طیبه فیاض

تاریخ و امضا:

مجوز بهره برداری از پایان نامه

بهره برداری از این پایان نامه در چهارچوب مقررات کتابخانه و با توجه به محدودیتی که توسط استاد راهنما

به شرح زیر تعیین می شود، بلامانع است:

بهره برداری از این پایان نامه برای همگان بلامانع است.

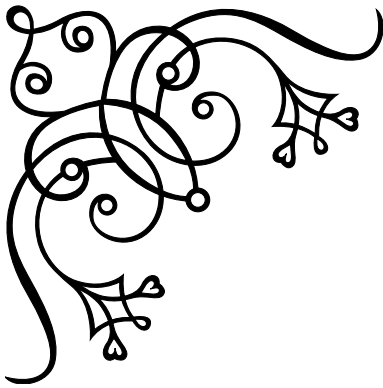
بهره برداری از این پایان نامه با اخذ مجوز از استاد راهنما، بلامانع است.

بهره برداری از این پایان نامه تا تاریخ ممنوع است.

استاد راهنما: دکتر محمود امین طوسی

تاریخ و امضا:

تقدیم به:



پدر و مادرم



سپاس خداوندگار حکیم را که با لطف بی کران خود، آدمی را زیور عقل آراست. در آغاز وظیفه خود می دانم از زحمات بی دریغ استاد راهنمای خود، جناب آقای دکتر محمود امین طوسی، صمیمانه تشکر و قدردانی کنم که قطعاً بدون راهنمایی های ارزنده ایشان، این مجموعه به انجام نمی رسید. از جناب آقای دکتر یاسر علیزاده که زحمت مطالعه و مشاوره این رساله را تقبل فرمودند و در آماده سازی این رساله، به نحو احسن اینجانب را مورد راهنمایی قرار دادند، کمال امتنان را دارم. و نیز از جناب آقای دکتر هادی صدوقی یزدی به خاطر استفاده از مطالب گردآوری شده توسط ایشان در زمینه خوشه بندی طبفی و کمک به پیشرفت علمی اینجانب سپاسگزارم. همچنین لازم می دانم از گروه پارسی لاتک در پاسخگویی به مشکلات کاربران کمال قدردانی را داشته باشم. در پایان، بوسه می زنم بر دستان خداوندگاران مهر و مهربانی، پدر و مادر عزیزم و بعد از خدا، ستایش می کنم وجود مقدس شان را و تشکر می کنم از خانواده عزیزم به پاس عاطفه سرشار و گرمای امیدبخش وجودشان، که بهترین پشتیبان من بودند.

طیبه فیاض

شهریورماه ۱۳۹۴

فهرست مطالب

ب	فهرست تصاویر
ج	فهرست جداول
د	فهرست الگوریتم‌ها
۱	چکیده
۲	پیش‌گفتار
۳	فصل ۱: مقدمات و پیش‌نیازها
۵	۱-۱ تاریخچه‌ای از مسأله خوشه‌بندی
۵	۱-۱-۱ کاربرد خوشه‌بندی
۶	۲-۱ گراف
۸	۳-۱ ماتریس
۱۳	۴-۱ ماتریس لاپلاسیان گراف
۱۶	۵-۱ افراز گراف
۱۸	۶-۱ ماتریس و تابع کرنل
۲۰	فصل ۲: خوشه‌بندی طیفی
۲۰	۱-۲ گراف شباهت
۲۱	۱-۱-۲ انواع گراف‌های شباهت
۲۳	۲-۲ روشهای افرازبندی و برش گراف
۲۴	۱-۲-۲ مسأله برش کمینه
۲۶	۲-۲-۲ مسأله نسبت برش

۲۶	مسأله برش نرمال	۳-۲-۲
۲۷	مسأله نسبت وابستگی	۴-۲-۲
۲۸	مسأله برش $Min - Max$	۵-۲-۲
۲۸	نسبت وابستگی و نسبت برش وزن دار	۶-۲-۲
۲۹	تقریب مسئله برش گراف	۳-۲
۲۹	تقریب مسئله برش نرمال	۱-۳-۲
۴۴	برش k تایی همزمان با چند بردار ویژه	۴-۲
۴۴	الگوریتم های خوشه بندی طیفی	۵-۲
۴۷	الگوریتم برش نرمال	۱-۵-۲
۴۷	خوشه بندی کرنلی	۶-۲
۴۹	فصل ۳: نگاه متحد به K-means کرنلی، خوشه بندی طیفی و برش گراف	
۴۹	مقدمه	۱-۳
۵۱	الگوریتم K-means	۱-۱-۳
۵۲	K-means کرنلی	۲-۳
۵۳	K-means کرنلی وزن دار	۳-۳
۵۴	هم ارزی تابع هدفهای طیفی با K-means کرنلی وزن دار	۴-۳
۵۴	K-means کرنلی وزن دار بصورت مسأله ماکزیم سازی اثر	۱-۴-۳
۵۷	هم ارزی تابع هدف های افراز گراف با K-means کرنلی وزن دار	۵-۳
۵۷	نسبت وابستگی	۱-۵-۳
۵۹	نسبت برش	۲-۵-۳
۶۰	برش نرمال	۳-۵-۳
۶۱	نسبت وابستگی وزن دار	۴-۵-۳
۶۲	نسبت برش وزن دار	۵-۵-۳
۶۴	فصل ۴: دست آوردهای پژوهش	
۶۵	روش نایستروم	۱-۴
۶۶	روش پیشنهادی	۲-۴
۶۹	فهرست منابع	

۷۲

پیوست آ: برنامه های نوشته شده

۷۷

واژه نامه فارسی به انگلیسی

۷۸

واژه نامه انگلیسی به فارسی

فهرست تصاویر

- ۱-۱ یک $v_1 v_6$ برش با $S = \{v_1, v_2, v_4\}$ و $\bar{S} = \{v_3, v_5, v_6\}$ ۷
- ۲-۱ گراف با ۵ رأس ۱۶
- ۳-۱ یک گراف افراز شده ۱۷
- ۱-۲ تغییرات تابع شباهت گوسی به ازای مقادیر مختلف σ ۲۲
- ۲-۲ نمایش اثر σ ثابت و σ محلی در تابع شباهت ۲۳
- ۳-۲ گراف با ۶ نود ۲۴
- ۴-۲ یک افراز بد توسط برش کمینه ۲۵
- ۵-۲ یک تصویر سطح خاکستری از بازی بیس بال ۴۲
- ۶-۲ استفاده از دومین تا نهمین کوچکترین بردار ویژه ماتریس لاپلاسیان ۴۳
- ۱-۳ یک نمونه داده جداپذیر خطی ۵۰
- ۲-۳ یک نمونه داده جداناپذیر خطی ۵۰
- ۱-۴ یک نمونه تصویر ام آر آی و خروجی قطعه بندی شده آن ۶۷
- ۲-۴ یک تصویر دیگر از مجموعه تصاویر ام آر آی و خروجی قطعه بندی شده آن ۶۸

فهرست جداول

۱-۱ مثال‌هایی از توابع کرنل ۱۹

فهرست الگوریتم‌ها

۳۷	الگوریتم گروه بندی	۱-۲
۴۴	الگوریتم $Ncut$ دوتایی بازگشتی	۲-۲
۴۶	دوبخشی طیفی	۳-۲
۴۶	دوبخشی طیفی بازگشتی	۴-۲
۴۷	دوبخشی طیفی اصلاح شده	۵-۲
۴۸	الگوریتم برش نرمال با استفاده از بردارهای ویژه دستگاه ویژه تعمیم یافته	۶-۲
۴۸	الگوریتم برش نرمال با استفاده از بردارهای ویژه دستگاه ویژه استاندارد	۷-۲
۵۲	الگوریتم K-means [۱]	۱-۳
۵۵	الگوریتم کرنل kmeans وزندار	۲-۳



دانشگاه گیلان

فرم چکیده ی پایان نامه ی دوره ی تحصیلات تکمیلی

مدیریت تحصیلات تکمیلی

نام خانوادگی دانشجو: فیاض	نام: طویه	ش. دانشجویی: ۹۲۱۳۱۳۳۰۱۹
استاد راهنما: دکتر محمود امین طوسی		
استاد مشاور: دکتر یاسر علیزاده		
دانشکده ریاضی و علوم کامپیوتر	رشته: ریاضی کاربردی	گرایش: تحقیق در عملیات
مقطع: کارشناسی ارشد	تاریخ دفاع: شهریورماه ۱۳۹۴	تعداد صفحات: ۷۹
عنوان پایان نامه: خوشه بندی طیفی برای قطعه بندی تصویر		
کلید واژه ها: خوشه بندی، قطعه بندی تصویر، افراز گراف، برش نرمال، الگوریتم ژنتیک		
<p>چکیده: خوشه بندی یکی از مهم ترین مسائل حوزه هوش مصنوعی است که یکی از کاربردهای آن، قطعه بندی تصویر در مبحث پردازش تصویر است. هدف در قطعه بندی تصویر، جداسازی پیکسل های تصویر به دو یا چند گروه مرتبط می باشد به عنوان مثال جداسازی بافت چربی از غیر آن در تصاویر ام آر آی، که در این پایان نامه به آن پرداخته شده است. تبدیل مسئله پردازش تصویر به یک مسئله گراف یکی از روش های قطعه بندی تصویر است که برای انجام آن از روش های مختلف خوشه بندی گراف می توان استفاده کرد. در خوشه بندی گراف، هدف افراز بهینه رئوس گراف به نحوی است که فاصله رئوس هر خوشه کمینه و فاصله بین خوشه ای بیشینه گردد.</p> <p>روش های مختلفی در خوشه بندی گراف مورد استفاده قرار گرفته و می گیرد. در این پایان نامه به روش های خوشه بندی طیفی پرداخته شده و هم ارزی تابع هدف نسخه های مختلف آن با یک روش خوشه بندی معمول دیگر نشان داده شده است.</p> <p>برای قطعه بندی تصاویر ام آر آی از روش خوشه بندی طیفی موسوم به «برش نرمال» استفاده شده و راهکاری برای تعیین یک پارامتر آن با استفاده از الگوریتم ژنتیک ارائه شده و کارایی آن نشان داده شده است.</p>		

پیش‌گفتار

خوشه‌بندی یکی از ابتدایی‌ترین فعالیت‌های ذهنی بشر است که برای استخراج مفهوم یا به بیانی دیگر خلاصه‌سازی اطلاعات استفاده می‌شود. خوشه‌بندی را می‌توان به عنوان مهم‌ترین روش در یادگیری بدون نظارت در نظر گرفت که در تشخیص الگوی بدون ناظر، بینایی ماشینی، تقطیع تصویر کاربرد فراوان دارد. در خوشه‌بندی هدف تقسیم کردن داده‌ها به خوشه‌ها به نحوی است که شباهت بین داده‌های هر خوشه حداکثر و تفاوت بین داده‌ها در خوشه‌های متفاوت حداقل باشد. تقطیع تصاویر ام آر آی^۱ یا همان خوشه‌بندی آنها مرحله مهمی است که با استفاده از بینایی ماشینی و علم خوشه‌بندی قابل انجام است. هدف در قطعه‌بندی تصویر، جداسازی پیکسل‌های تصویر به دو یا چند گروه مرتبط می‌باشد.

یکی از روشهای یادگیری زیرفضا در خوشه‌بندی، خوشه‌بندی طیفی است که با استفاده از آن می‌توان مسئله پردازش تصویر را به یک مسئله گراف تبدیل کرد و سپس با استفاده از روشهای مختلف خوشه‌بندی طیفی به قطعه‌بندی تصویر پرداخت. در خوشه‌بندی طیفی، هدف افراز بهینه رئوس گراف به نحوی است که فاصله رئوس هر خوشه کمینه و فاصله بین خوشه‌های بیشینه گردد.

این پایان‌نامه شامل ۴ فصل است:

در فصل اول تعاریف و مفاهیم موردنیاز بیان خواهند شد.

در فصل دوم به بیان روشها و الگوریتم‌های خوشه‌بندی طیفی پرداخته خواهد شد.

فصل سه به نشان دادن هم‌ارزی تابع هدف نسخه‌های مختلف خوشه‌بندی طیفی با تابع هدف خوشه‌بندی *kmeans* کرنلی اختصاص یافته است.

در فصل چهارم به قطعه‌بندی تصاویر ام آر آی با استفاده از روش خوشه‌بندی طیفی موسوم به «برش نرمال» پرداخته خواهد شد و راهکاری برای تعیین یک پارامتر آن با استفاده از الگوریتم ژنتیک ارائه و کارایی آن نشان داده خواهد شد.

^۱Magnetic Resonance imaging (MRI)

فصل ۱

مقدمات و پیش‌نیازها

یادگیری زیرفضا یکی از اساسی‌ترین مسائل یادگیری ماشین است و به‌عنوان مرحله پیش‌پردازش، برای یافتن بازنمایی با ابعاد کوچکتر برای مجموعه داده به‌کار می‌رود. هدف این فرآیند حذف اطلاعات نامربوط^۱ و تکراری^۲ از مجموعه داده خام اولیه و تولید و بازنمایی مجموعه‌هایی با اطلاعات مفید می‌باشد. یکی از اساسی‌ترین خواستگاه‌های این تکنیک، خوشه‌بندی داده‌ها است. تعریف خوشه‌بندی ما را به تعریف خوشه سوق می‌دهد یعنی باید ابتدا درک درستی از خوشه داشته باشیم.

فرض کنید X مجموعه‌ای از داده‌ها باشد که قصد داریم روی آنها خوشه‌بندی را انجام دهیم:

$$X = \{x_1, x_2, \dots, x_n\}$$

یک خوشه‌بندی m تایی از X ، افزار X به m مجموعه c_1, c_2, \dots, c_m است که سه شرط زیر را دارا است:

$$c_i \neq \emptyset \quad i = 1, 2, \dots, m \quad ۱.$$

$$\bigcup_{i=1}^m c_i = X \quad ۲.$$

$$c_i \cap c_j = \emptyset \quad , i \neq j \quad , i, j = 1, 2, \dots, m \quad ۳.$$

به علاوه بردارهایی که در خوشه c_j هستند به هم دیگر شبیه‌اند و از خوشه‌های دیگر متفاوتند و هر بردار ویژگی تنها به یک خوشه تعلق دارد. بنابراین خوشه به مجموعه‌ای از داده‌ها گفته می‌شود که به هم شباهت داشته باشند. خوشه‌بندی را می‌توان به عنوان مهم‌ترین روش در یادگیری بدون نظارت در نظر گرفت. در خوشه‌بندی هدف تقسیم کردن داده‌ها به خوشه‌ها به نحوی است که شباهت بین داده‌های هر خوشه حداکثر و تفاوت بین داده‌ها در

^۱Irrelevant ^۲Redundant

خوشه‌های متفاوت حداقل باشد. در واقع این هدف، به‌عنوان یک معیار در انجام یک خوشه‌بندی خوب مدنظر است، یعنی نقاطی که در یک خوشه هستند مشابه با یکدیگر، و متفاوت از دیگر خوشه‌ها باشند.

گام‌های اساسی که یک متخصص به منظور توسعه خوشه‌بندی باید انجام دهد به شرح زیر است:

۱. انتخاب ویژگی^۱: ویژگی‌ها باید به‌طور صحیح انتخاب شوند تا توصیف درستی از الگو به‌دست دهند، یعنی اطلاعات بیشتری در خصوص آن فعالیت به‌دست آوریم و اطلاعات زائد را از میان ویژگی‌هایی که هدف اصلی ماست، حذف کنیم.

۲. میزان نزدیکی^۲ بین الگوها: این اندازه کمی است که مشابه یا غیر مشابه بودن دو بردار ویژگی را نشان می‌دهد.

۳. محک خوشه‌بندی^۳: محک خوشه‌بندی یا تابع هزینه به تعبیری میزان معقول^۴ بودن خوشه‌ها است و از یک نگاه تابعی است که با بهینه کردن آن روی داده‌ها، داده‌ها در خوشه‌های معقول قرار می‌گیرند. اشاره به این نکته ضروری است، نمونه‌هایی با d ویژگی طبق یک محک ممکن است در خوشه‌هایی معقول قرار بگیرند در حالی که از نگاه دیگر با یک معیار یا محک دیگر آن چنان معقول نباشد. محک خوشه‌بندی می‌تواند به‌وسیله تابع هدف یا برخی قواعد دیگر بیان گردد.

۴. الگوریتم خوشه‌بندی: الگوریتم خوشه‌بندی به اندازه نزدیکی و محکی که انتخاب می‌کنیم، بستگی دارد. در این گام یک الگوریتم که موفق به خوشه‌بندی داده‌ها شود و مبتنی بر تابع هزینه روش می‌باشد، ارائه می‌گردد.

۵. اعتبار نتایج^۵: باید صحت و درستی نتایجی که از یک الگوریتم خوشه‌بندی به‌دست آمده است، بررسی شود. به عبارتی معیارهایی برای بررسی عملکرد خوشه‌بندی وجود دارد که نشان می‌دهد در روش خوشه‌بندی مورد نظر به چند جنبه خوشه‌بندی توجه شده است. مثلاً ممکن است پس از اجرای الگوریتم، یکنواختی خوشه به عنوان معیار به‌کار رفته اما نامتجانس بودن خوشه‌ها مورد توجه قرار نگرفته باشد.

۶. تفسیر نتایج^۶: در اکثر موارد، متخصص باید نتایج خوشه‌بندی را با دیگر شواهد تجربی به منظور به‌دست آوردن نتایج درست، مقایسه و تحلیل کند.

^۱Feature selection ^۲Proximity measure ^۳Clustering criterion ^۴sensible ^۵Validation
of the results ^۶Interpretation of the results

۱-۱ تاریخچه‌ای از مسأله خوشه‌بندی

خوشه‌بندی^۱ یکی از ابتدایی‌ترین فعالیت‌های ذهنی بشر است که برای استخراج مفهوم یا به بیانی دیگر خلاصه‌سازی اطلاعات استفاده می‌شود. این مفهوم در علوم مختلفی مانند علوم زیست‌شناسی و جانورشناسی، جغرافیا و زمین‌شناسی، جامعه‌شناسی و باستان‌شناسی، علوم روانپزشکی و آسیب‌شناسی، تجارت، سیستم‌های مهندسی و پردازش تصویر به‌کار برده شده است. خوشه‌بندی با توجه به زمینه‌ای که در آن استفاده می‌شود، نام‌های متفاوتی پیدا کرده است مانند یادگیری بدون ناظر در تشخیص الگو^۲، رده‌بندی جانوران در زیست‌شناسی، گونه‌شناسی در علوم اجتماعی و افراز در نظریه گراف.

۱-۱-۱ کاربرد خوشه‌بندی

خوشه‌بندی در تشخیص الگوی بدون ناظر، بینایی ماشین، تقطیع تصویر^۳ کاربرد فراوان دارد. یکی از اهداف شناسایی الگو خوشه‌بندی داده‌هاست. از جمله کاربردهای مهم تشخیص الگو در علوم پزشکی تحت عنوان تشخیص به کمک ماشین^۴ می‌باشد. هدف این سیستم‌ها، در حقیقت کمک به پزشک برای تشخیص درد و نیز بالا بردن کیفیت درمان است. لازم به ذکر است در چنین سیستم‌هایی تشخیص نهایی با پزشک است. تقطیع تصاویر ام آر آی یا همان خوشه‌بندی آنها مرحله مهمی است که با استفاده از بینایی ماشین^۵ و علم خوشه‌بندی قابل انجام است. عکس‌های پزشکی شامل نویز زیادی هستند که یا توسط متصدی اخذ تصویر ام آر آی با تکان خوردن وسایل، یا در هنگام اخذ ام آر آی به وسیله‌ی تجهیزات، به علت کیفیت پایین حسگرها ایجاد می‌شوند. روش‌های تقطیع بدون ناظر به پزشک کمک می‌کند تا تاثیر نویز را بر این عکس کاهش دهد. همچنین این روش برای بررسی ناحیه ضایعه یا تومور هم کاربرد دارد. بعبارتی در تقطیع ام آر آی سعی در برجسب‌زنی بافت‌های نرمال و غیرنرمال برای کمک به پزشک است.

دو دسته از مهم‌ترین روش‌های خوشه‌بندی عبارتند از:

۱. خوشه‌بندی طیفی

۲. خوشه‌بندی کرنلی

در این تحقیق به بیان جزئیات مرتبط با خوشه‌بندی طیفی و ارتباط آن با خوشه‌بندی کرنلی پرداخته می‌شود.

نمادگذاری: در این تحقیق تمام جا بردار $[1, 1, \dots, 1]^T$ را با نماد $\vec{1}$ نشان می‌دهیم.

^۱clustering ^۲Pattern recognition ^۳Image segmentation ^۴Computer aided diagnosis ^۵Computer vision

۲-۱ گراف

برای تعاریف این بخش از مرجع [۲] استفاده شده است.

تعریف ۱-۲-۱. مجموعه چندگانه^۱ مجموعه‌ای است که اعضای آن بیش از یک بار ظاهر می‌شوند، پس هر عضو در یک مجموعه چندگانه یک نرخ تکرار^۲ دارد.

تعریف ۲-۲-۱. زوج $G = (V, E)$ که در آن V مجموعه‌ای متناهی و E مجموعه‌ای از زیرمجموعه‌های دوعضوی مجموعه‌ی V است را گراف G گویند.

مجموعه‌ی V را مجموعه‌ی رئوس و اعضای آن را رئوس گراف G گویند و مجموعه‌ی E را مجموعه‌ی یال‌ها و اعضای آن را یال‌های گراف G گویند.

تعریف ۳-۲-۱. زوج $G = (V, A)$ که در آن V مجموعه‌ای متناهی است و A مجموعه‌ای چندگانه از زوج‌های مرتب از اعضای مجموعه V است را گراف جهت‌دار G گویند.

مجموعه‌ی V را مجموعه‌ی رئوس و اعضای آن را رئوس گراف جهت‌دار G گویند و مجموعه‌ی A را مجموعه‌ی کمان‌ها و اعضای آن را کمان‌های گراف جهت‌دار G گویند.

از نماد n برای نمایش تعداد رئوس گراف استفاده می‌کنیم. یعنی داریم: $|V| = n$

تعریف ۴-۲-۱. راس مجاور رأس v در گراف G رأسی است که با یالی به v وصل شده باشد.

تعریف ۵-۲-۱. رأسی که از آن هیچ یالی عبور نکند، یک رأس منفرد (تنها) نامیده می‌شود.

تعریف ۶-۲-۱. هر یالی که از یک رأس به همان رأس متصل شود، حلقه نامیده می‌شود.

تعریف ۷-۲-۱. گراف بدون جهت، بدون حلقه و بدون یال‌های چندگانه، گراف ساده نامیده می‌شود.

تعریف ۸-۲-۱. تعداد یال‌های متصل به یک رأس، درجه آن رأس نامیده می‌شود.

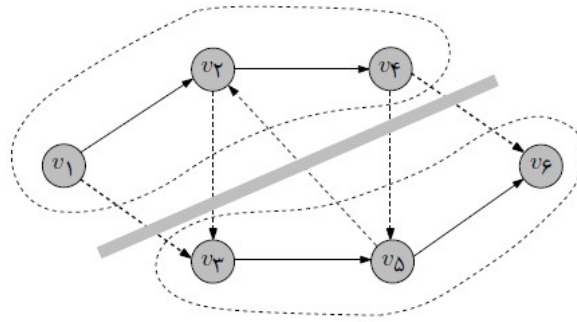
تعریف ۹-۲-۱. یک گردش بین رأس‌ها که در آن، رأس تکراری مجاز نیست، مسیر نامیده می‌شود.

تعریف ۱۰-۲-۱. گراف G کامل نامیده می‌شود اگر از هر رأس به تمامی رئوس آن، یالی متصل باشد.

تعریف ۱۱-۲-۱. گراف H را یک زیرگراف از G گوئیم اگر مجموعه‌ی رئوس H زیرمجموعه‌ای از مجموعه رئوس G باشد و همچنین هر یال uv از H یالی از G باشد.

^۱multi-set

^۲multiplicity



شکل ۱-۱: یک برش v_1v_6 با $S = \{v_1, v_2, v_4\}$ و $\bar{S} = \{v_3, v_5, v_6\}$

تعریف ۱-۲-۱۲. یک گراف را همبند^۱ گوئیم اگر به ازای هر جفت رأس u و v یک مسیر بدون جهت بین u و v وجود داشته باشد.

تعریف ۱-۲-۱۳. زیرگراف H از G را یک مؤلفه از G گوئیم هرگاه H همبند باشد و زیرگراف همبندی از G به غیر از H وجود نداشته باشد که H زیرگراف آن باشد.

تعریف ۱-۲-۱۴. افزاز مجموعه رئوس یک گراف به دو زیرمجموعه ناتهی S و $\bar{S} = V - S$ را یک برش گوئیم و آن را با $[S, \bar{S}]$ نمایش می‌دهیم.

تعریف ۱-۲-۱۵. برای دو رأس s و t از گراف، برش $[S, \bar{S}]$ را یک st -برش گوئیم هرگاه $s \in S$ و $t \in \bar{S}$ باشد.

در شکل ۱-۱ مفهوم برش نشان داده شده است.

تعریف ۱-۲-۱۶. یک برش از یک گراف همبند را یک برش ساده گوئیم اگر گراف حاصل از حذف کمان‌های متناظر با آن برش دقیقاً دو مؤلفه داشته باشد.

تعریف ۱-۲-۱۷. ماتریس قطری یک ماتریس مربعی است که تمام درایه‌های خارج از قطر اصلی آن صفر هستند.

تعریف ۱-۲-۱۸. اگر رئوس گراف G را با v_1, v_2, \dots, v_n نمایش دهیم، ماتریس مجاورت $W_{n \times n}$ ماتریسی می‌باشد که درایه ij آن برابر است با تعداد یال‌هایی که v_i را به v_j متصل می‌کند.

^۱connected

بدیهی است در گراف‌های بدون جهت

$$\forall i, j : \begin{cases} W_{ij} = 1 \\ \text{یا} \\ W_{ij} = 0 \end{cases}$$

و $W_{ij} = W_{ji}$ ، یعنی ماتریس مجاورت یک گراف بدون جهت، متقارن است.

اگر گراف بدون حلقه باشد $\forall i : W_{ii} = 0$ ، یعنی درایه‌های روی قطر اصلی ماتریس مجاورت در این حالت همگی صفر می‌باشند.

لذا به راحتی دیده می‌شود ماتریس مجاورت یک گراف ساده، یک ماتریس متقارن با درایه‌های صفر و یک می‌باشد که تمام درایه‌های قطر اصلی آن الزاماً صفر است.

تعریف ۱-۲-۱۹. برای گراف بدون جهت G ماتریس درجه D ، ماتریسی قطری است که عناصر روی قطر اصلی آن نشان‌دهنده مجموع یال‌های متصل به هر رأس گراف G است یا به عبارتی d_{ii} درجه رأس i ام است.

۳-۱ ماتریس

تعاریف و قضایای این بخش تا تعریف ۱-۳-۱۵ از مرجع [۳] اقتباس شده اند.

تعریف ۱-۳-۱. $f : \mathbb{R}^n \rightarrow \mathbb{R}$ یک نرم ماتریسی نامیده می‌شود اگر سه شرط زیر برقرار باشند [۳]:

$$1. \quad A = 0 \quad \text{اگر و تنها اگر} \quad (f(A) = 0) \quad \forall A \in \mathbb{R}^{m \times n} \quad ; \quad f(A) \geq 0$$

$$2. \quad f(A+B) \leq f(A) + f(B) \quad ; \quad \forall A, B \in \mathbb{R}^{m \times n}$$

$$3. \quad f(\alpha A) = |\alpha| f(A) \quad ; \quad \forall \alpha \in \mathbb{R}, A \in \mathbb{R}^{m \times n}$$

تعریف ۱-۳-۲. نرم فروبینیوس یک ماتریس $A_{m \times n}$ که نرم اقلیدسی نیز نامیده می‌شود برابر است با ریشه دوم مجموع توان دوم تمام درایه‌های ماتریس A و به صورت زیر نوشته می‌شود [۳]:

$$\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^m |a_{ij}|^2} \quad (1-1)$$

تعریف ۱-۳-۳. ماتریس A را قطری شدنی گوئیم اگر ماتریس وارون پذیر P وجود داشته باشد بطوریکه
 $P^{-1}AP = B$ که در آن B یک ماتریس قطری است [۳].

تعریف ۱-۳-۴. ماتریس $A \in M_n$ نرمال نامیده می‌شود [۳] اگر $AA^T = A^T A$.

تعریف ۱-۳-۵. ماتریس $B \in M_n$ متقارن نامیده می‌شود [۳] اگر $B = B^T$.

تعریف ۱-۳-۶. ماتریس $A \in M_n$ معین مثبت نامیده می‌شود هرگاه به ازای هر بردار $x \in \mathbb{R}^n$ داشته باشیم
 [۳]:

$$x^T Ax > 0$$

تعریف ۱-۳-۷. ماتریس $A \in M_n$ نیمه معین مثبت نامیده می‌شود هرگاه به ازای هر بردار $x \in \mathbb{R}^n$ داشته
 باشیم [۳]:

$$x^T Ax \geq 0$$

تعریف ۱-۳-۸. افزایش سطرها و ستونهای ماتریس $A_{m \times n}$ بلوکی کردن ماتریس A نامیده می‌شود و هر یک از
 افزایشهای ایجاد شده یک بلاک یا زیر ماتریس برای ماتریس A نامیده می‌شود [۳].

تعریف ۱-۳-۹. اگر V یک فضای برداری روی میدان F باشد [۳] مجموعه $S = \{x_1, \dots, x_m\} \subseteq V$ را
 مستقل خطی می‌نامیم اگر

$$\{\forall \lambda_i \in F : \lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_m x_m = 0 \quad \ni \quad \lambda_1 = \lambda_2 = \dots = \lambda_m = 0\}$$

و S را وابسته خطی می‌نامیم اگر

$$\{\exists (\lambda_1, \dots, \lambda_m) \neq (0, \dots, 0) \ni \lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_m x_m = 0\}$$

تعریف ۱-۳-۱۰. ماتریس دلخواه $M_{n \times n}$ را در نظر بگیرید. بردار غیرصفر v یک بردار ویژه با مقدار ویژه λ
 برای ماتریس M خواهد بود اگر رابطه $Mv = \lambda v$ برقرار باشد. در مجموع ماتریس $n \times n$ بعدی M دارای n
 بردار ویژه و n مقدار ویژه خواهد بود [۳].

تعریف ۱-۳-۱۱. تعداد دفعات تکرار یک مقدار ویژه به عنوان ریشه معادله مشخصه ماتریس را تکرار جبری^۱
 آن مقدار ویژه می‌نامیم [۳].

^۱Algebraic multiplicity

تعریف ۱-۳-۱۲. برای ماتریس $M_{n \times n}$ با مقادیر ویژه $\lambda_1, \lambda_2, \dots, \lambda_n$ اثر ماتریس به صورت زیر محاسبه می شود [۳]:

$$\text{trace}(M) = \lambda_1 + \lambda_2 + \dots + \lambda_n$$

قضیه ۱-۳-۱۳. (قضیه ریلی ریتز^۱)

فرض کنید A یک ماتریس متقارن حقیقی و x یک بردار غیر صفر باشد که بر $1 - j$ تا از کوچکترین بردارهای ویژه A ، یعنی v_1, \dots, v_{j-1} عمود است در این صورت کسر $\frac{x^T A x}{x^T x}$ با استفاده از j امین کوچکترین بردار ویژه A ، یعنی v_j مینیمم می شود و مقدار این مینیمم متناظر با مقدار ویژه متناظر با v_j یعنی λ_j است [۳].

برهان. اگر $\{v_1, \dots, v_{j-1}, \dots, v_n\}$ بردارهای ویژه ماتریس A متناظر با مقادیر ویژه $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ باشند، رابطه

$$A v_i = \lambda_i v_i \quad (2-1)$$

برقرار است. طبق فرض قضیه داریم:

$$x \perp v_i \quad (x \cdot v_i = 0) \quad \forall \quad i = 1, \dots, j-1$$

چون $\{v_1, v_2, \dots, v_n\}$ بردارهای ویژه ماتریس A هستند، مستقل خطی می باشند بنابراین تشکیل یک پایه برای فضا می دهند پس x را می توان به صورت ترکیب خطی این بردارها نوشت، در نتیجه اسکالرهای c_1, \dots, c_n موجودند طوریکه

$$x = \sum_{i=1}^n c_i v_i \quad (3-1)$$

لذا با توجه به روابط ۱-۲ و ۱-۳ می توان نوشت:

^۱Rayleigh Ritz

$$\begin{aligned}
x^T Ax &= x^T A \left(\sum_{i=1}^n c_i v_i \right) = x^T \left(\sum_{i=1}^n c_i A v_i \right) = x^T \left(\sum_{i=1}^n c_i \lambda_i v_i \right) \\
&= \sum_{i=1}^n c_i \lambda_i x^T v_i = \sum_{i=j}^n c_i \lambda_i x^T v_i \\
&= \sum_{i=j}^n c_i \lambda_i c_i v_i^T v_i = \sum_{i=j}^n c_i^2 \lambda_i \|v_i\|^2
\end{aligned}$$

همچنین بنا به رابطه ۳-۱ داریم:

$$x^T x = \left(\sum_{i=1}^n c_i v_i^T \right) \left(\sum_{i=1}^n c_i v_i \right) = \sum_{i=1}^n c_i^2 v_i^T v_i = \sum_{i=1}^n c_i^2 \|v_i\|^2$$

بنابراین می توان نوشت:

$$\frac{x^T Ax}{x^T x} = \frac{\sum_{i=j}^n c_i^2 \lambda_i \|v_i\|^2}{\sum_{i=1}^n c_i^2 \|v_i\|^2}$$

چون

$$\lambda_j \leq \lambda_i \quad \forall \quad i = j, \dots, n$$

نتیجه می گیریم که:

$$\frac{(\sum_{i=j}^n c_i^2 \|v_i\|^2) \lambda_j}{\sum_{i=1}^n c_i^2 \|v_i\|^2} \leq \frac{x^T Ax}{x^T x}$$

با فرض

$$x = v_j = \sum_{i=1}^n c_i v_i \quad \ni \quad c_j = 1, c_i = 0 \quad \forall \quad i \neq j$$

داریم:

$$\frac{x^T Ax}{x^T x} = \frac{c_j^2 \|v_j\|^2 \lambda_j}{c_j^2 \|v_j\|^2} = \lambda_j$$

□

بنابراین کسر $\frac{x^T Ax}{x^T x}$ در $x = v_j$ مینیمم می شود.

تعریف ۱-۳-۱۴. اگر X ماتریس شامل بردارهای ستونی $\{x_1, \dots, x_m\}$ باشد، ماتریس $G = X^T X$ ماتریس

گرام نامیده می شود و درایه های آن برابر است با حاصلضرب داخلی بردارهای $\{x_1, \dots, x_m\}$ یعنی $G_{ij} = x_i^T x_j$.

ماتریس گرام متقارن و نیمه معین مثبت است.

تعریف ۱-۳-۱۵. فرض کنید A و B دو ماتریس $n \times n$ باشند. λ را یک مقدار ویژه تعمیم یافته زوج (A, B) می نامیم هرگاه بردار $x \neq 0$ یافت شود به طوریکه $Ax = \lambda Bx$. بردار x بردار ویژه تعمیم یافته زوج (A, B) نامیده می شود و معادله مشخصه آن بصورت زیر است [۴]:

$$(A - \lambda B)x = 0, \quad x \neq 0$$

تعریف ۱-۳-۱۶. حاصلضرب داخلی دو بردار $x = (x_1, \dots, x_n)^T$ و $y = (y_1, \dots, y_n)^T$ بصورت زیر تعریف می شود [۵]:

$$x \cdot y = x^T y = \sum_{i=1}^n x_i y_i$$

حاصلضرب داخلی دو بردار x و y را با $\langle x, y \rangle$ نیز نمایش می دهند.

لم ۱-۳-۱۷. اگر $M_{n \times n}$ یک ماتریس متقارن باشد، آنگاه بردارهای ویژه آن متعامدند.

برهان. برای هر ماتریس حقیقی A و هر بردار x و y رابطه زیر برقرار است:

$$\langle Ax, y \rangle = \langle x, A^T y \rangle$$

اکنون فرض می کنیم M یک ماتریس متقارن است و x و y بردارهای ویژه ماتریس M به ترتیب متناظر با مقادیر ویژه متمایز λ و u می باشند، با توجه به رابطه فوق داریم:

$$\lambda \langle x, y \rangle = \langle \lambda x, y \rangle = \langle Mx, y \rangle = \langle x, M^T y \rangle = \langle x, My \rangle = \langle x, uy \rangle = u \langle x, y \rangle$$

بنابراین

$$(\lambda - u) \langle x, y \rangle = 0$$

و چون $\lambda - u \neq 0$ پس داریم:

$$\langle x, y \rangle = 0 \quad (x \perp y)$$

□

پس نتیجه می گیریم که بردارهای ویژه یک ماتریس متقارن متعامدند.

۴-۱ ماتریس لاپلاسین گراف

اصلی‌ترین ابزار خوشه‌بندی طیفی ماتریس‌های لاپلاسین گراف هستند [۶]. الگوریتم‌های خوشه‌بندی طیفی عموماً بر پایه استفاده از این ماتریس‌ها و تجزیه ویژه و استفاده از بردارها و مقادیر ویژه آنها برای تحلیل دانش موجود در ساختار داده‌ها می‌باشند. تئوری طیفی گراف [۷] عنوانی است که به این موضوع می‌پردازد.

تعریف ۱-۴-۱. فرض کنید $G = (V, E)$ یک گراف ساده باشد که n رأس دارد و W ماتریس مجاورت گراف باشد، ماتریس لاپلاسین $L(G)$ متناظر با گراف G یک ماتریس $n \times n$ متقارن است و به صورت زیر تعریف می‌شود:

$$L = D - W \quad (۴-۱)$$

و مقادیر درایه‌های آن بصورت زیر می‌باشند:

$$L_{i,j}(G) = \begin{cases} d_i & \text{if } i = j \\ -1 & \text{if } i \neq j \quad (i, j) \in E \\ 0 & \text{if } i \neq j \quad (i, j) \notin E \end{cases} \quad \forall \quad i, j = 1, \dots, n$$

قضیه ۱-۴-۲. ویژگی‌های ماتریس لاپلاسین عبارتند از [۸]:

۱. برای هر بردار $f \in \mathbb{R}^n$ داریم:

$$f^T L f = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n W_{ij} (f_i - f_j)^2$$

۲. L ماتریس متقارن و نیمه‌معین مثبت است.

۳. کوچکترین مقدار ویژه ماتریس L برابر است با صفر و متناظر است با بردار ویژه \vec{j} .

برهان. ۱. با توجه به تعریف L ، داریم:

$$\begin{aligned}
f^T L f &= f^T D f - f^T W f \\
&= \sum_{i=1}^n d_i f_i^2 - \sum_{i=1}^n \sum_{j=1}^n f_i f_j W_{ij} \\
&= \frac{1}{2} \left(\sum_{i=1}^n d_i f_i^2 - 2 \sum_{i=1}^n \sum_{j=1}^n f_i f_j W_{ij} + \sum_{j=1}^n d_j f_j^2 \right) \\
&= \frac{1}{2} \left(\sum_{i=1}^n \sum_{j=1}^n W_{i,j} f_i^2 - 2 \sum_{i=1}^n \sum_{j=1}^n f_i f_j W_{ij} + \sum_{i=1}^n \sum_{j=1}^n W_{i,j} f_j^2 \right) \\
&= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n W_{i,j} (f_i - f_j)^2
\end{aligned}$$

۲. متقارن بودن ماتریس L ، مستقیماً از متقارن بودن ماتریس‌های D و W نتیجه می‌شود. برای نیمه‌معین مثبت بودن یک ماتریس باید شرط $f^T L f \geq 0$ برای هر بردار دلخواه $f \in \mathbb{R}^n$ برقرار باشد. با توجه به اثبات قسمت (۱) و نامنفی بودن W_{ij} ها، به‌وضوح این شرط برای ماتریس L برقرار است.

۳. با توجه به نیمه معین مثبت بودن ماتریس L واضح است که $f^T L f \geq 0$ $\forall f \in \mathbb{R}^n$. اگر فرض کنیم x یک بردار ویژه با مقدار ویژه λ برای ماتریس L باشد، داریم:

$$Lx = \lambda x \implies x^T L x = \lambda x^T x$$

بنابراین با توجه به اینکه $x^T L x \geq 0$ و $x^T x > 0$ نتیجه می‌گیریم که $\lambda \geq 0$ و $\lambda = 0$ یک مقدار ویژه متناظر با بردار ویژه $\vec{j} = [1, 1, \dots, 1]^T$ است زیرا

$$L\vec{j} = (D - W)\vec{j} = D\vec{j} - W\vec{j} = D - D = 0$$

بنابراین کوچکترین مقدار ویژه ماتریس لاپلاسیان، همیشه صفر است و بردار ویژه متناظر با آن، بردار \vec{j} می‌باشد. \square

یکی از ویژگیهای مهمی که ماتریس لاپلاسیان و بردارها و مقادیر ویژه آن در خوشه‌بندی طیفی دارد، به‌دست آوردن تعداد اجزاء همبند گراف با استفاده از طیف L می‌باشد. قضیه زیر این ویژگی را مطرح می‌کند.

قضیه ۱-۴-۳. فرض کنید G یک گراف وزن دار بدون جهت با اوزان نامنفی باشد، آنگاه تعداد اجزاء همبند گراف G برابر با تکرار جبری مقدار ویژه صفر ماتریس لاپلاسیان گراف G است.

برهان. تعداد اجزاء همبند گراف G را با k نمایش می‌دهیم. ابتدا فرض می‌کنیم $k = 1$ است در این حالت گراف همبند است. همچنین فرض کنید f یک بردار ویژه با مقدار ویژه صفر ماتریس لاپلاسین است. با توجه به بخش اول قضیه ۱-۴-۲ و رابطه ۱-۴ می‌توان نوشت:

$$0 = f^T L f = \sum_{i=1}^n \sum_{j=1}^n W_{i,j} (f_i - f_j)^2$$

با توجه به اینکه وزن‌های $W_{i,j}$ نامنفی هستند، مجموع فوق زمانی می‌تواند صفر باشد که تمام جمله‌های $W_{i,j} (f_i - f_j)^2$ صفر باشند. بنابراین، اگر دو راس v_i و v_j متصل باشند ($W_{i,j} > 0$)، آنگاه باید f_i و f_j برابر باشند. به عبارت دیگر f باید برای تمام رئوس که از طریق مسیری به هم متصلند، ثابت باشد. بنابراین در گرافی که تنها از یک جزء متصل تشکیل شده است، فقط بردار ویژه $\vec{1}$ متناظر با مقدار ویژه صفر را خواهیم داشت که بردار شاخص این جزء متصل خواهد بود.

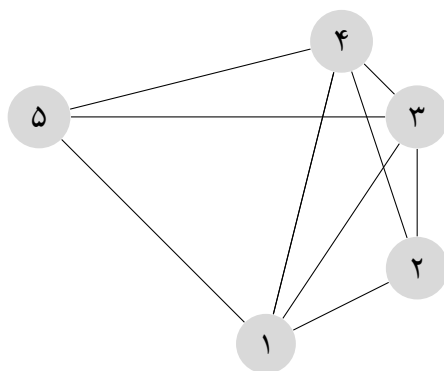
حال فرض کنید گراف از $k \leq n$ جزء همبند تشکیل شده باشد. بدون از دست دادن کلیت مسئله فرض کنید که رئوس گراف به ترتیب درون اجزاء همبند مرتب شده قرار دارند. در این حالت ماتریس L و W بلاک قطری می‌باشند.

$$\begin{pmatrix} L_1 & & & \\ & L_2 & & \\ & & \ddots & \\ & & & L_k \end{pmatrix}$$

هر یک از بلاک‌های L_i خود یک ماتریس لاپلاسین گراف است که مرتبط با زیرگراف i امین جزء همبند می‌باشد. برای ماتریس‌های بلاک-قطری مانند L در اینجا، طیف گراف از اجتماع بردارهای ویژه بلاک‌های L_i در قسمتی که مرتبط با بلاک دیگری است صفر قرار داده شده باشد، به دست می‌آید. با توجه به اینکه هر L_i خود یک لاپلاسین مرتبط با یک جزء همبند در گراف است و با توجه به بخش اول اثبات، هر L_i یک مقدار ویژه صفر را خواهد داشت و بردار ویژه متناظر با آن بردار $\vec{1}$ خواهد بود.

بنابراین ماتریس L به تعداد k جزء همبندی که دارد، دارای مقدار ویژه صفر و بردار ویژه ثابت $\vec{1}$ خواهد بود. \square

مثال ۱-۴-۴. برای گراف شکل ۱-۲ ماتریسهای شباهت (W)، درجه (D)، و لاپلاسین (L) محاسبه شده است. همانطور که ملاحظه می‌شود، ماتریس لاپلاسین متقارن است.



شکل ۱-۲: گراف با ۵ رأس

$$W = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix}, D = \begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}, L = \begin{pmatrix} 4 & -1 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 & 0 \\ -1 & -1 & 4 & -1 & -1 \\ -1 & -1 & -1 & 4 & -1 \\ -1 & 0 & -1 & -1 & 3 \end{pmatrix}$$

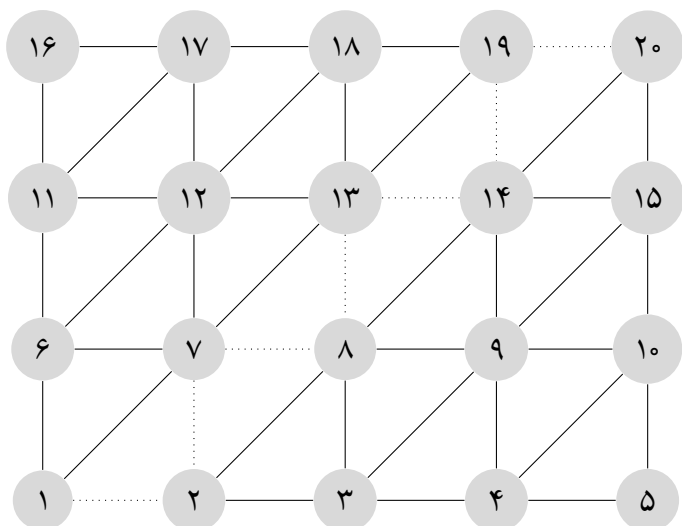
مثال ۱-۴-۵. با تجزیه ویژه ماتریس L گراف مثال قبل، مقادیر ویژه آن برابر است با:

$$\lambda = \{0, 3, 5, 5, 5\}$$

ماتریس L دارای یک مقدار ویژه صفر است و گراف مربوطه نیز همبند است بنابراین همانطور که در قضیه ۱-۴-۳ بیان شد تعداد اجزاء همبند گراف برابر است با تکرر مقدار ویژه صفر ماتریس لاپلاسیان گراف.

۵-۱ افزایش گراف

هدف مساله افزایش k تایی گراف، حذف یال‌های گراف به نحوی است که باعث ایجاد k زیرگراف مجزا شود [۹]. گراف $\tilde{G} = (V, E, W_V, W_E)$ را در نظر می‌گیریم، رأس v_i در V نماینده یک رأس گراف و یال $e = (i, j)$ در E متصل‌کننده دو رأس i و j و W_V و W_E مجموعه‌هایی هستند که به ترتیب نشان‌دهنده وزنه‌های نامنفی برای هر رأس و هر یال هستند.



شکل ۱-۳: یک گراف افراز شده

در این صورت W_{Vv_i} وزن رأس i و $W_{Ee(i,j)}$ وزن یال e_{ij} را نشان می‌دهد. افراز G به معنی تقسیم کردن V به اجتماعی از n مجموعه مجزا V_1, V_2, \dots, V_n است بطوریکه

$$V = V_1 \cup V_2 \cup \dots \cup V_n \quad (5-1)$$

برای افراز بهینه شرایط زیر باید برقرار باشند [۹]:

۱. مجموع وزن‌های W_V رأس‌ها در هر V_i باید تقریباً برابر باشند.
۲. مجموع وزن‌های W_E از یال‌های اتصال‌دهنده رأس‌ها در V_i و V_j متفاوت باید مینیمم باشد. در ادامه وزن هر رأس و هر یال را یک در نظر می‌گیریم.

مثال ۱-۵-۱. شکل ۱-۳ یک گراف افراز شده را نشان می‌دهد. رأس‌های $\{1, 6, 7, 11, 12, 13, 16, 17, 18, 19\}$ متعلق به زیرگراف V_1 و رأس‌های $\{2, 3, 4, 5, 8, 9, 10, 14, 15, 20\}$ متعلق به زیرگراف V_2 و یالهایی که به صورت نقطه‌چین مشخص شده‌اند یال‌های برش هستند. همانطور که مشاهده می‌کنید هر دو خوشه به دست آمده دارای اندازه مشابه اند و مجموع یال‌های اتصال‌دهنده خوشه‌های V_1 و V_2 هفت است.

۶-۱ ماتریس و تابع کرنل

به منظور استخراج و تجزیه تحلیل داده‌ها باید به جستجوی نمایش مناسبی از داده‌ها پرداخت [۱۰]. به عنوان مثال برای داده‌های پیچیده مانند متن، دنباله‌ها، تصاویر و غیره ابتدا مجموعه‌ای از ویژگیها را ساخته یا استخراج می‌کنیم سپس داده‌ها را به صورت بردارهای چندمتغیره نشان می‌دهیم. برای مجموعه داده X نگاشت ϕ را پیدا کنیم طوریکه $\phi(X)$ بردار نمایش X است. در اصطلاح به فضای داده‌های اولیه، فضای ورودی و به فضای بردارهای نگاشت داده شده $\phi(X)$ ، فضای ویژگی گفته می‌شود.

مثال ۱-۶-۱. مجموعه داده توالی DNA را به ترتیب حروف الفبا به عنوان فضای ورودی در نظر بگیرید.

$$\Sigma = \{A, C, G, T\}$$

یک فضای ویژگی ساده نمایش توزیع احتمال نمادهای Σ است. برای این منظور یک دنباله X را که $|X| = m$ در نظر بگیرید. نحوه نگاشت دادن داده‌های ورودی به فضای ویژگی به صورت زیر است:

$$\phi(X) = \{p(A), p(C), p(G), p(T)\}$$

که $p(s) = \frac{n_s}{m}$ برابر است با احتمال مشاهده نماد $s \in \Sigma$ و n_s تعداد دفعات ظاهر شدن s در دنباله X است. مثلا اگر $X = ACAGCAGTA$ با $m = |X| = 9$ چون A ۴ بار ظاهر شده و C و G ۲ بار و T یکبار، پس داریم:

$$\phi(X) = \{0.44444444, 0.22222222, 0.22222222, 0.11111111\}$$

میزان شباهت بین هر جفت از داده‌های ورودی، توسط ماتریس کرنل $n \times n$ که به صورت زیر تعریف می‌شود،

نشان داده می‌شود:

$$K = \begin{pmatrix} k(x_1, x_1) & k(x_1, x_2) & \dots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \dots & k(x_2, x_n) \\ \vdots & \vdots & \vdots & \vdots \\ k(x_n, x_1) & k(x_n, x_2) & \dots & k(x_n, x_n) \end{pmatrix}$$

Kernel Polynomial	$\mathcal{K}(a_i, a_j) = (a_i \cdot a_j + c)^d$
Kernel Gaussian	$\mathcal{K}(a_i, a_j) = \exp\left(\frac{-\ a_i - a_j\ ^2}{2\alpha^2}\right)$
Kernel Sigmoid	$\mathcal{K}(a_i, a_j) = \tanh(c(a_i \cdot a_j) + \theta)$

جدول ۱-۱: مثال‌هایی از توابع کرنل

$\mathcal{K} : \mathcal{T} \times \mathcal{T} \rightarrow \mathbb{R}$ یک تابع کرنل روی هر جفت نقطه در فضای ورودی است که باید در شرط زیر صدق

کند:

$$\mathcal{K}(x_i, x_j) = \phi(x_i)^T \phi(x_j)$$

در بسیاری از روشهای داده‌کاوی می‌توان به‌جای نگاشت فضای ورودی به فضای ویژگی، داده‌ها را با ماتریس

کرنل K نمایش داد و تمام تجزیه تحلیل‌ها را روی K انجام داد.

فصل ۲

خوشه‌بندی طیفی

در این فصل ابتدا مسئله افراز گراف و برخی از روشهای آن مطرح شده و پس از آشنایی با این مفهوم، بحث تقریب برش گراف بررسی می‌شود که پیش نیاز الگوریتم‌های خوشه‌بندی طیفی است، سپس مهمترین الگوریتم‌های خوشه‌بندی طیفی مطرح می‌شود.

۱-۲ گراف شباهت

مجموعه $X = \{x_1, x_2, \dots, x_n\}$ را با n داده d بعدی در نظر بگیرید. می‌توان دانشی که درباره میزان شباهت داده‌ها وجود دارد را به زیبایی در قالب یک گراف، که گراف شباهت^۱ نامیده می‌شود، بیان نمود. برای چنین توصیفی، گراف $G = (V, E)$ با n رأس را در نظر بگیرید که هر رأس v_i متناظر با یکی از داده‌های مجموعه X می‌باشد. چنانچه میزان شباهت بین دو داده عددی مثبت یا بزرگتر از حد آستانه‌ی خاصی باشد، یالی رئوس متناظر این دو داده در گراف را به یکدیگر متصل خواهد نمود و وزنی که این یال خواهد داشت برابر با (W_{ij}) می‌باشد. در صورت در اختیار داشتن چنین بازنمایی از مجموعه داده X می‌توان خوشه‌بندی را به این صورت بیان نمود که هدف یافتن افرازی برای G است به صورتیکه وزن یال‌های اتصال‌دهنده نقاط در هر گروه زیاد (شباهت درون گروهی بالا) و وزن یال‌های بین نقاط گروه‌های مختلف کم (شباهت بین گروهی کم) باشد.

^۱Similarity graphs

۱-۱-۲ انواع گراف‌های شباهت

برای ساختن گراف شباهت راه‌های مختلفی وجود دارد که هدف همه آنها مشخص نمودن رابطه بین هر نقطه با همسایگانش می‌باشد. [۸]

گراف ϵ -همسایگی: در این گراف رئوس متناظر با جفت داده‌ای که فاصله بین آنها از ϵ کمتر باشد به یکدیگر متصل می‌شوند. فاصله بین یال‌های متصل در این گراف تقریباً از یک مقیاس است (حداکثر ϵ) و بنابراین مشخص کردن وزن یال‌ها اطلاعات بیشتری را در اختیار ما قرار نمی‌دهد. بنابراین، گراف‌های ϵ -همسایگی معمولاً به صورت بدون وزن در نظر گرفته می‌شوند.

گراف k -نزدیکترین همسایگی: در این گراف رأس v_i در صورتی به رأس v_j متصل می‌شود که داده متناظر با رأس v_j جزء یکی از k نزدیکترین همسایگان داده متناظر با رأس v_i باشد. اما با توجه به اینکه رابطه همسایگی متقارن نمی‌باشد، چنین تعریفی منجر به گراف جهت‌دار خواهد شد. برای خارج کردن گراف از چنین حالتی دو راه وجود دارد:

راه اول: در این حالت دو رأس v_i و v_j به هم متصل می‌شوند اگر فقط یکی از رئوس جزء یکی از k نزدیکترین همسایگان رأس دیگری باشد. این گراف معمولاً گراف k -نزدیکترین همسایه نامیده می‌شود.

راه دوم متصل نمودن دو رأس v_i و v_j به یکدیگر است در صورتیکه v_i جزء یکی از k نزدیکترین همسایگان v_j باشد و همزمان v_j نیز یکی از k نزدیکترین همسایگان v_i باشد. این گراف، گراف k -نزدیکترین همسایه متقابل نامیده می‌شود.

گراف کاملاً متصل: در این گراف همه رئوس به یکدیگر متصلند و همه یال‌ها بر اساس میزان شباهت رئوس متناظرشان وزنی دارند.

با توجه به اینکه هدف از بازنمایی داده‌ها مدل کردن همسایگی‌های محلی نقاط است، تابع شباهتی که استفاده می‌شود باید بتواند این همسایگی را مدل نماید. یک مثال برای چنین تابعی، تابع شباهت گوسی می‌باشد که بصورت رابطه زیر تعریف می‌شود:

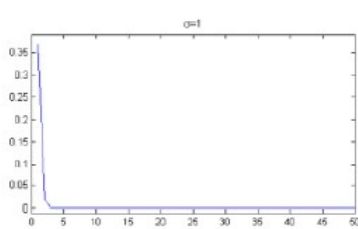
$$W(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right) \quad (1-2)$$

که در آن $x_i, x_j \in X$ بوده و مقادیر متناظر با رئوس v_i و v_j می‌باشند و پارامتر σ عرض همسایگی را کنترل می‌کند یا به بیان دیگر، میزان کاهش تابع شباهت با افزایش فاصله بین دو نمونه را مشخص می‌نماید. برای بررسی

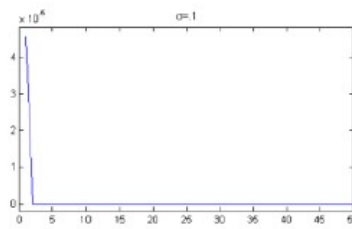
اثر σ ، می‌توان تغییرات تابع شباهت به ازای مقادیر مختلف σ را بررسی نمود. برای این منظور، شکل ۱-۲ را در نظر بگیرید.

محور افقی تمام نمودارها نشان دهنده میزان فاصله بین داده‌های مورد بررسی هستند.

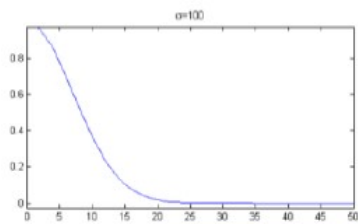
تغییرات تابع شباهت به ازای مقادیر مختلف $\sigma = 0.1$ ، $\sigma = 10$ ، $\sigma = 100$ به ترتیب در شکل‌های ۱-۲(آ) تا ۱-۲(د) نشان داده شده است. همانطور که مشاهده می‌کنید هرچه مقدار σ عدد بزرگتری باشد، تابع شباهت که متأثر از این پارامتر است، مقادیر بزرگتری می‌گیرد و همچنین دامنه شباهت نیز گسترش می‌یابد. به عنوان مثال، دو عدد با فاصله ۱۰ هنگامیکه $\sigma = 0.1$ باشد، مشابه نیستند و مقدار تابع شباهت صفر است (شکل ۱-۲(آ)) اما اگر $\sigma = 100$ باشد، مقدار تابع شباهت افزایش چشمگیری می‌یابد. (شکل ۱-۲(د))



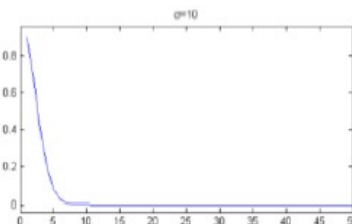
شکل ۱-۲(ب) $\sigma = 1$



شکل ۱-۲(آ) $\sigma = 0.1$



شکل ۱-۲(د) $\sigma = 100$



شکل ۱-۲(ج) $\sigma = 10$

شکل ۱-۲: تغییرات تابع شباهت گوسی به ازای مقادیر مختلف σ

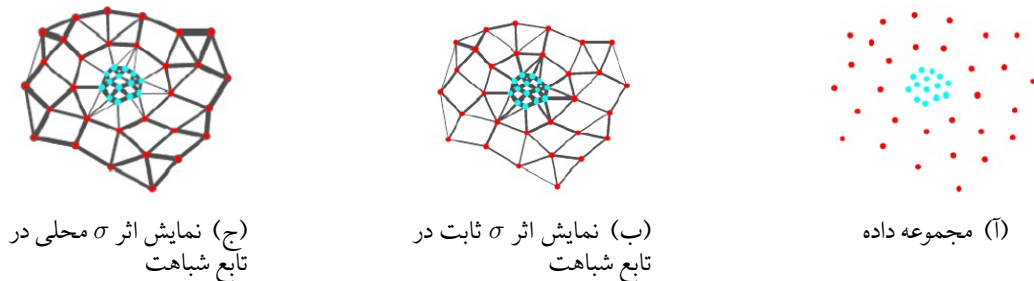
همانطور که در بحث فوق نشان داده شد، تغییرات پارامتر σ در تابع شباهت بسیار اثر می‌گذارد و در نتیجه بر روی الگوریتم‌های خوشه‌بندی نیز اثرگذار می‌باشد.

روشهای مختلفی برای انتخاب یک مقدار مناسب برای σ ارائه شده است. روشی که در این بخش معرفی می‌شود یکی از پرکاربردترین روش‌های تنظیم خودکار این پارامتر می‌باشد. این روش توسط زلنیک^۱ و پرونا^۲ ارائه و پیشنهاد شده است [۱۱] که در آن به جای انتخاب یک مقدار ثابت برای σ ، این پارامتر برای هر نمونه داده به صورت محلی تنظیم شود. بنابراین تعریف، رابطه ۱-۲ بصورت

$$W(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{\sigma_i \sigma_j}\right) \quad (2-2)$$

^۱Zelnik ^۲Perona

نوشته می شود که در آن σ_i فاصله نمونه داده x_i تا نزدیکترین p امین همسایه اش می باشد. این روش به انتخاب p امین همسایه وابسته است و در [۱۱] پیشنهاد شده است که از هفتمین نزدیکترین همسایه هر داده برای انتخاب σ استفاده شود.



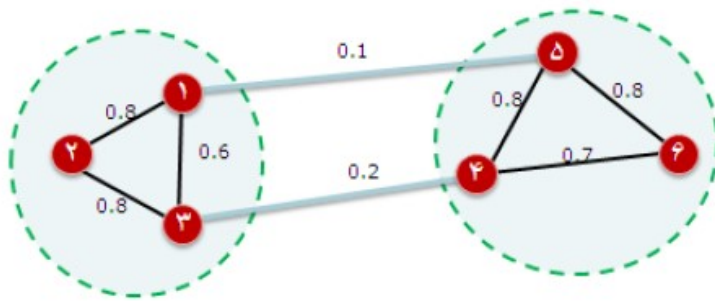
شکل ۲-۲: نمایش اثر σ ثابت و σ محلی در تابع شباهت

تأثیر این روش در تنظیم σ را می توان در شکل ۲-۲ مشاهده نمود. در قسمت ۲-۲ (آ) مجموعه داده، به صورت یک خوشه متراکم (خوشه داخلی) و خوشه پراکنده دیگری که خوشه اولیه را در بر گرفته است نشان داده شده است. در ۲-۲ (ب) اثر انتخاب یک مقدار ثابت برای σ نمایش داده شده است که ضخامت خطوط نشان دهنده بیشتر بودن میزان شباهت دو داده متناظر در ماتریس شباهت می باشد. همچنان که مشاهده می شود، در این حالت شباهت بین خوشه ای در مقایسه با شباهت های درون خوشه ای زیاد است و بنابراین ساختار داده ها به خوبی نشان داده نمی شود. در شکل ۲-۲ (ج) اثر انتخاب σ به صورت محلی نمایش داده شده است. همانطور که در این شکل مشاهده می شود، شباهت درون خوشه ای در این حالت برای هر دو خوشه زیاد (ضخامت خطوط بیشتر) و شباهت بین خوشه ای کم (ضخامت خطوط کم) می باشد.

در بین گراف های شباهت موجود در خوشه بندی طیفی، معمولاً از گراف کامل برای نشان دادن شباهت داده های ورودی استفاده می شود.

۲-۲ روشهای افرازبندی و برش گراف

الگوریتم های خوشه بندی طیفی بر پایه جواب مسأله افراز گراف می باشند. این الگوریتم ها معمولاً از یک یا تعداد بیشتری از بردارهای ویژه ماتریس لاپلاسیان برای خوشه بندی استفاده می نمایند و این بردارها جواب مسأله افراز گراف هستند. روش های مختلفی برای افراز گراف وجود دارند که هر کدام مزایا و معایب خود را دارند [۱۲] در این بخش تعدادی از این روشها بیان و بررسی خواهد شد.



شکل ۲-۳: گراف با ۶ نود

۱-۲-۲ مسأله برش کمینه

هدف مسأله برش کمینه^۱ [۱۳]، افراز گراف به صورتی است که داده‌های هر خوشه بیشترین میزان شباهت با یکدیگر را داشته باشند. در این حالت یال‌هایی برای حذف شدن انتخاب می‌شوند که میزان شباهت رأس‌های متناظرشان مینیمم باشد. بنابراین تابع برش در حالتی که مجموعه داده به دو خوشه A و B افراز می‌شود به صورت زیر تعریف می‌گردد:

$$cut(A, B) = \sum_{i \in A, j \in B} W_{i,j}$$

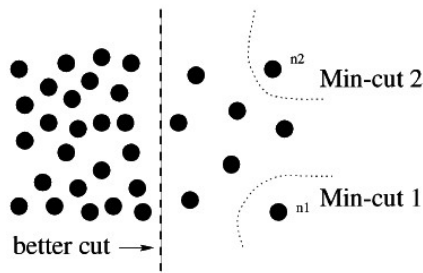
در آن $W_{i,j}$ بیانگر وزن یال اتصال دهنده رئوس i و j می‌باشد.

مثال ۱-۲-۲. شکل ۲-۳ نشان‌دهنده گراف متناظر با مجموعه داده‌ای با ۶ نمونه می‌باشد. بنا بر روش برش کمینه، باید به دنبال انتخاب یال‌هایی از گراف برای برش بود که مجموع وزن آنها (که بیانگر شباهت بین خوشه‌ای می‌باشد) مینیمم باشد. بنابراین یال متناظر رئوس ۱ و ۵ و نیز یال متناظر رئوس ۳ و ۴ حذف خواهند شد و خواهیم داشت:

$$mincut(A, B) = 0.1 + 0.2 = 0.3$$

مسئله برش کمینه برای افراز گراف به k خوشه (افراز k -تایی) به صورت کلی زیر نوشته می‌شود:

^۱minimum cut problem



شکل ۲-۴: یک افراز بد توسط برش کمینه

$$\min_{\pi^k \in \Pi^k} cut(\pi^k) \quad (۳-۲)$$

که در آن

$$cut(\pi^k) = \frac{1}{2} \sum_{i=1}^k W(C_i, \bar{C}_i)$$

و در آن \bar{C}_i هر خوشه‌ای مانند C_j است به طوری که $j \neq i$ باشد و Π^k مجموعه تمام افرازهای k تایی ممکن برای گراف G می‌باشد. به دلیل اینکه هر یال در محاسبه cut دوبار در نظر گرفته می‌شود ضریب $\frac{1}{2}$ را برای رابطه در نظر می‌گیریم.

تابع هدف برش کمینه (۳-۲) با افزایش تعداد یال‌هایی که بین دو خوشه افراز شده وجود دارند افزایش می‌یابد، بدیهی است که نقاط تنها می‌توانند افراز خوبی را مطابق با معیار برش کمینه ایجاد کنند، لذا معیار برش کمینه نقاط تنها را در گراف به عنوان یک خوشه در نظر می‌گیرد و باعث می‌شود خوشه‌ها اندازه معقولی نداشته باشند [۱۴]. این مشکل در شکل ۲-۴ نمایش داده شده است. همانطور که در شکل ۲-۴ مشاهده می‌کنید برشی که رأس‌ها را به دو رأس n_1 (یا n_2) و سایر نقاط افراز می‌کند نسبت به برشی که رأس‌ها را به دو نیمه چپ و راست افراز می‌کند، مقدار خیلی کوچکتري دارد.

دوروشی که از مرسوم‌ترین راه‌حلهای این مشکل شناخته می‌شوند، مسئله نسبت برش و برش نرمال می‌باشند که در ادامه معرفی خواهد شد. هدف این روشها دست یافتن به افرازی می‌باشد که خوشه‌ها در آن متعادل باشند.

۲-۲-۲ مسأله نسبت برش

مسأله نسبت برش^۱ که برای اجتناب از ایجاد خوشه‌هایی با رأس‌های تنها پیشنهاد شده است تقسیم رابطه ۳-۲ بر تعداد اعضای هر خوشه می‌باشد. این مسئله ابتدا در [۱۵] برای افراز دو تایی گراف پیشنهاد شد و سپس در [۱۶] برای حالت افراز k تایی نیز تعمیم داده شد. این مسئله به صورت رابطه زیر نوشته می‌شود:

$$\min_{\pi^k \in \Pi^k} \text{ratiocut}(\pi^k) \quad (۴-۲)$$

درحالی‌که

$$\text{ratiocut}(\pi^k) = \frac{1}{2} \sum_{i=1}^k \frac{W(C_i, \bar{C}_i)}{|C_i|}$$

که در آن، هر خوشه‌ای مانند C_j است بصورتیکه $i \neq j$ باشد و Π^k مجموعه تمام افرازهای k تایی ممکن برای گراف G می‌باشد.

۳-۲-۲ مسأله برش نرمال

یکی از روش‌های پرکاربرد برای مقابله با مشکل رأس‌های تنها که در مسأله برش کمینه به آن اشاره شد، مسأله برش نرمال^۲ می‌باشد. این مسئله توسط شی و مالیک^۳ [۱۷] مطرح شده که در آن همزمان که به مینیمم نمودن وزن اتصالات بین خوشه‌ها پرداخته می‌شود، وزن اتصالات درون هر خوشه نیز افزایش یابد. بنابراین برای افراز مجموعه داده به دو خوشه A و B تابع هدف به صورت زیر تعریف می‌شود:

$$\text{ncut}(A, B) = \frac{\text{cut}(A, B)}{\text{vol}(A)} + \frac{\text{cut}(A, B)}{\text{vol}(B)} \quad (۵-۲)$$

که در آن $\text{vol}(A)$ و $\text{vol}(B)$ به ترتیب، مجموع وزن‌های خوشه A و B می‌باشد و به صورت زیر محاسبه می‌شود:

$$\text{vol}(A) = \sum_{u \in A, t \in V} W(u, t)$$

$$V = A \cup B$$

^۱ ratio cut problem

^۲ normalized cut problem

^۳ Shi and Malik

همچنان که در رابطه ۲-۵ مشاهده می‌شود، به جای اینکه فقط وزن یال‌های اتصال‌دهنده خوشه‌های به دست آمده در نظر گرفته شود، نسبت وزن یال‌های برش به مجموع درجه رئوس هر خوشه محاسبه شود. در این صورت، برشی که باعث ایجاد نقاط تنها می‌گردد مقدار کوچکی نخواهد داشت و چنین برشی پذیرفته نخواهد شد. مسئله را در حالت کلی می‌توان به صورت زیر تعریف نمود:

$$\min_{\pi^k \in \Pi^k} ncut(\pi^k) \quad (۶-۲)$$

درحالی‌که

$$ncut(\pi^k) = \frac{1}{2} \sum_{i=1}^k \frac{W(C_i, \bar{C}_i)}{vol(C_i)}$$

که در آن، \bar{C}_i هر خوشه‌ای مانند C_j است بصورتیکه $j \neq i$ باشد و Π^k مجموعه تمام افرازهای k تایی ممکن برای گراف G می‌باشد.

۴-۲-۲ مسأله نسبت وابستگی

هدف روش نسبت وابستگی^۱ (که میانگین وابستگی نیز نامیده می‌شود)، ماکزیمم کردن تراکم درون خوشه‌ای نسبت به سایر خوشه است.

$$RAssoc(G) = \max_{c_1, \dots, c_k} \sum_{i=1}^k \frac{W(C_i, C_i)}{|C_i|} \quad (۷-۲)$$

^۱Ratio Association Problem

۵-۲-۲ مسأله برش $Min - Max$

این مسأله توسط دینگ^۱ و همکارانش [۱۸] ارائه شده است و هدف آن حداقل سازی شباهت بین خوشه‌ای و حداکثر سازی شباهت‌های درون خوشه‌ای به صورت همزمان می‌باشد که به صورت زیر تعریف می‌شود:

$$\min_{\pi^k \in \Pi^k} MinMaxCut(\pi^k) \quad (۸-۲)$$

در حالیکه

$$MinMaxCut(\pi^k) = \sum_{i=1}^k \frac{W(C_i, \bar{C}_i)}{W(C_i, C_i)}$$

در اینجا نیز \bar{C}_i هر خوشه‌ای مانند C_j است بصورتیکه $i \neq j$ باشد و Π^k مجموعه تمام افرازهای k تایی ممکن برای گراف G می‌باشد.

۶-۲-۲ نسبت وابستگی و نسبت برش وزن دار

فرض می‌کنیم که علاوه بر یال‌های گراف، به هر یک از رئوس نیز وزنی اختصاص داده شود، بنابراین تابع هدف مسأله نسبت برش وزن دار بصورت زیر تعریف می‌شود:

$$wcut(G) = \min_{C_1, \dots, C_k} \sum_{i=1}^k \frac{W(C_i, V \setminus C_i)}{w(C_i)} \quad (۹-۲)$$

و تابع هدف مسأله نسبت وابستگی وزن دار بصورت زیر فرموله می‌شود:

$$wAssoc(G) = \max_{C_1, \dots, C_k} \sum_{i=1}^k \frac{W(C_i, C_i)}{w(C_i)}$$

$w(C_i)$ مجموع وزن رئوس زیرگراف C_i ام گراف است.

روشهای دیگری نیز برای افراز گراف پیشنهاد شده است که برای مطالعه بیشتر در این زمینه می‌توان به [۱۲]

مراجعه نمود.

^۱Ding

۳-۲ تقریب مسئله برش گراف

همچنان که در بخش قبل اشاره شد، الگوریتم‌های خوشه‌بندی طیفی از یک یا چند بردار ویژه ماتریس لاپلاسیان گراف استفاده می‌نمایند که جواب ساده شده بعضی از مسائل برش گراف می‌باشد. استفاده از یک ساده‌سازی^۱ مناسب برای مسئله افراز گراف، امکان کاوش در خصوصیات بردارها و مقادیر ویژه ماتریس لاپلاسیان گراف را فراهم می‌آورد [۱۲]. در این بخش با توجه به اهمیت مسئله برش نرمال، تقریب این مسئله به‌طور مفصل مورد بررسی قرار می‌گیرد.

۱-۳-۲ تقریب مسئله برش نرمال

تابع هدف مسأله برش نرمال (۵-۲) را در نظر می‌گیریم. میزان وابستگی (شباهت) بین گروهی برای یک افراز داده شده بصورت زیر فرموله می‌شود [۱۷]:

$$N_{assoc}(A, B) = \frac{assoc(A, A)}{assoc(A, V)} + \frac{assoc(B, B)}{assoc(B, V)} \quad (۱۰-۲)$$

که در آن

$$assoc(A, V) = \sum_{u \in A, t \in V} W(u, t)$$

$assoc(A, A)$ و $assoc(B, B)$ به ترتیب مجموع وزن یال‌های اتصال‌دهنده رأس‌ها در خوشه‌ی A و B می‌باشند. رابطه فوق میزان تراکم خوشه‌ها یا به عبارت دیگر میزان شباهت بین خوشه‌ای را نشان می‌دهد. بدیهی است که با کم شدن میزان شباهت بین خوشه‌ای تراکم خوشه‌ها بالا می‌رود. رابطه ۵-۲ را می‌توان به صورت زیر نوشت:

^۱Relaxation

$$\begin{aligned}
Ncut(A, B) &= \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)} \\
&= \frac{assoc(A, V) - assoc(A, A)}{assoc(A, V)} + \frac{assoc(B, V) - assoc(B, B)}{assoc(B, V)} \\
&= ۲ - \left(\frac{assoc(A, A)}{assoc(A, V)} + \frac{assoc(B, B)}{assoc(B, V)} \right) \\
&= ۲ - Nassoc(A, B)
\end{aligned}$$

از تساوی فوق و اینکه $Ncut(A, B) \geq ۰$ نتیجه می‌گیریم که:

$$۰ \leq Nassoc(A, B) \leq ۲$$

پس با زیاد شدن شباهت بین گروهی یا بالا رفتن تراکم خوشه‌ها تابع هدف برش نرمال کمتر و بهینه‌تر می‌شود. با توجه به اینکه دو معیاری که برای افراز گراف در الگوریتم خوشه‌بندی مد نظر قرار می‌گیرد، مینیمم کردن شباهت بین گروهی و ماکزیمم کردن شباهت درون گروهی است می‌توان برش نرمال شده را به عنوان یک معیار خوب برای افراز در نظر گرفت.

فرض کنید گراف $G = (V, E)$ را به دو خوشه A و B افراز کرده‌ایم و $d_i = \sum_j w(i, j)$ وزن رأس i ام گراف است. بردار شاخص $x_{n \times 1}$ را به صورت زیر در نظر می‌گیریم:

$$x_i = \begin{cases} 1 & i \in A \\ -1 & i \notin A \end{cases}$$

با این فرضیات $Ncut(A, B)$ را به صورت زیر بازنویسی می‌کنیم:

$$\begin{aligned}
Ncut(A, B) &= \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)} \\
&= \frac{\sum_{(x_i > ۰, x_j < ۰)} -W_{i,j} x_i x_j}{\sum_{x_i > ۰} d_i} + \frac{\sum_{(x_i < ۰, x_j > ۰)} -W_{i,j} x_i x_j}{\sum_{x_i < ۰} d_i}
\end{aligned}$$

که در آن D ماتریس درجه و W ماتریس مجاورت گراف است. فرض می‌کنیم k نسبت مجموع وزن‌های

رئوس خوشه A به مجموع کل رئوس گراف باشد:

$$k = \frac{\sum_{x_i > \circ} d_i}{\sum_i d_i}$$

از $\frac{\vec{j}+x}{\gamma}$ و $\frac{\vec{j}-x}{\gamma}$ که به ترتیب بردارهای شاخص برای $\circ < x_i < \circ$ هستند استفاده می‌کنیم و $Ncut(A, B)$ را به صورت زیر می‌نویسیم:

$$\begin{aligned} \gamma Ncut(A, B) &= \frac{(\vec{j}+x)^T(D-W)(\vec{j}+x)}{kj^T D \vec{j}} + \frac{(\vec{j}-x)^T(D-W)(\vec{j}-x)}{(\gamma-k)\vec{j}^T D \vec{j}} \\ &= \frac{(\vec{j}+x)^T(D-W)(\vec{j}+x)(\gamma-k) + (\vec{j}-x)^T(D-W)(\vec{j}-x)k}{k(\gamma-k)\vec{j}^T D \vec{j}} \\ &= \frac{[(\vec{j}+x)^T(D-W)\vec{j} + (\vec{j}+x)^T(D-W)x](\gamma-k)}{k(\gamma-k)\vec{j}^T D \vec{j}} \\ &\quad + \frac{[(\vec{j}-x)^T(D-W)\vec{j} - (\vec{j}-x)^T(D-W)x]k}{k(\gamma-k)\vec{j}^T D \vec{j}} \\ &= \frac{(\vec{j}+x)^T(D-W)\vec{j} + (\vec{j}+x)^T(D-W)x}{k(\gamma-k)\vec{j}^T D \vec{j}} \\ &\quad - \frac{(\vec{j}+x)^T(D-W)\vec{j}k + (\vec{j}+x)^T(D-W)xk}{k(\gamma-k)\vec{j}^T D \vec{j}} \\ &\quad + \frac{(\vec{j}-x)^T(D-W)\vec{j}k - (\vec{j}-x)^T(D-W)xk}{k(\gamma-k)\vec{j}^T D \vec{j}} \\ &= \frac{\vec{j}^T(D-W)\vec{j} + x^T(D-W)\vec{j} + \vec{j}^T(D-W)x + x^T(D-W)x}{k(\gamma-k)\vec{j}^T D \vec{j}} \\ &\quad - \frac{\gamma x^T(D-W)\vec{j}k + \gamma \vec{j}^T(D-W)xk}{k(\gamma-k)\vec{j}^T D \vec{j}} \\ &= \frac{x^T(D-W)x + \vec{j}^T(D-W)x(\gamma-\gamma k) + \vec{j}^T(D-W)\vec{j} + x^T(D-W)\vec{j}(\gamma-\gamma k)}{k(\gamma-k)\vec{j}^T D \vec{j}} \\ &= \frac{x^T(D-W)x + \vec{j}^T(D-W)\vec{j}}{k(\gamma-k)\vec{j}^T D \vec{j}} + \frac{\gamma(\gamma-\gamma k)\vec{j}^T(D-W)x}{k(\gamma-k)\vec{j}^T D \vec{j}} \end{aligned}$$

فرض می‌کنیم

$$\gamma = \vec{j}^T(D-W)\vec{j}, \quad M = \vec{j}^T D \vec{j}, \quad \alpha(x) = x^T(D-W)x, \quad \beta(x) = \vec{j}^T(D-W)x$$

طبق فرضیات فوق ادامه روند اثبات به صورت زیر ادامه می‌یابد:

$$\begin{aligned}\mathfrak{V}ncut(A, B) &= \frac{(\alpha(x) + \gamma) + \mathfrak{V}(\mathfrak{V} - \mathfrak{V}k)\beta(x)}{k(\mathfrak{V} - k)M} \\ &= \frac{(\alpha(x) + \gamma) + \mathfrak{V}(\mathfrak{V} - \mathfrak{V}k)\beta(x)}{k(\mathfrak{V} - k)M} - \frac{\mathfrak{V}(\alpha(x) + \gamma)}{M} + \frac{\mathfrak{V}\alpha(x)}{M} + \frac{\mathfrak{V}\gamma}{M}\end{aligned}$$

به دلیل صفر بودن γ آخرین عبارت یعنی $\frac{\mathfrak{V}\gamma}{M}$ را از تساوی بالا حذف می کنیم در نتیجه:

$$\begin{aligned}\mathfrak{V}ncut(A, B) &= \frac{(\alpha(x) + \gamma) + \mathfrak{V}(\mathfrak{V} - \mathfrak{V}k)\beta(x) - \mathfrak{V}(\alpha(x) + \gamma)k(\mathfrak{V} - k)}{k(\mathfrak{V} - k)M} + \frac{\mathfrak{V}\alpha(x)}{M} \\ &= \frac{(\alpha(x) + \gamma) + \mathfrak{V}(\mathfrak{V} - \mathfrak{V}k)\beta(x) + \mathfrak{V}(k^\mathfrak{V} - k)(\alpha(x) + \gamma)}{k(\mathfrak{V} - k)M} + \frac{\mathfrak{V}\alpha(x)}{M} \\ &= \frac{(\mathfrak{V} - \mathfrak{V}k + \mathfrak{V}k^\mathfrak{V})(\alpha(x) + \gamma) + \mathfrak{V}(\mathfrak{V} - \mathfrak{V}k)\beta(x)}{k(\mathfrak{V} - k)M} + \frac{\mathfrak{V}\alpha(x)}{M} \\ &= \frac{\frac{(\mathfrak{V} - \mathfrak{V}k + \mathfrak{V}k^\mathfrak{V})}{(\mathfrak{V} - k)^\mathfrak{V}}(\alpha(x) + \gamma) + \frac{\mathfrak{V}(\mathfrak{V} - \mathfrak{V}k)}{(\mathfrak{V} - k)^\mathfrak{V}}\beta(x)}{\frac{k}{\mathfrak{V} - k}M} + \frac{\mathfrak{V}\alpha(x)}{M}\end{aligned}$$

فرض می کنیم $b = \frac{k}{\mathfrak{V} - k}$ بنابراین داریم:

$$\begin{aligned}\mathfrak{V}ncut(A, B) &= \frac{(\mathfrak{V} + b^\mathfrak{V})(\alpha(x) + \gamma) + \mathfrak{V}(\mathfrak{V} - b^\mathfrak{V})\beta(x)}{bM} + \frac{\mathfrak{V}b\alpha(x)}{bM} \\ &= \frac{(\mathfrak{V} + b^\mathfrak{V})(\alpha(x) + \gamma)}{bM} + \frac{\mathfrak{V}(\mathfrak{V} - b^\mathfrak{V})\beta(x)}{bM} + \frac{\mathfrak{V}b\alpha(x)}{bM} - \frac{\mathfrak{V}b\gamma}{bM} \\ &= \frac{(\mathfrak{V} + b^\mathfrak{V})(x^T(D - W)x + \vec{j}^T(D - W)\vec{j})}{b\vec{j}^T D \vec{j}} + \frac{\mathfrak{V}(\mathfrak{V} - b^\mathfrak{V})\vec{j}^T(D - W)x}{b\vec{j}^T D \vec{j}} \\ &\quad + \frac{\mathfrak{V}bx^T(D - W)x}{b\vec{j}^T D \vec{j}} - \frac{\mathfrak{V}b\vec{j}^T(D - W)\vec{j}}{b\vec{j}^T D \vec{j}} \\ &= \frac{x^T(D - W)x + \vec{j}^T(D - W)\vec{j} + b^\mathfrak{V}x^T(D - W)x + b^\mathfrak{V}\vec{j}^T(D - W)\vec{j}}{b\vec{j}^T D \vec{j}} \\ &\quad + \frac{\mathfrak{V}\vec{j}^T(D - W)x - \mathfrak{V}b^\mathfrak{V}\vec{j}^T(D - W)x + \mathfrak{V}bx^T(D - W)x - \mathfrak{V}b\vec{j}^T(D - W)\vec{j}}{b\vec{j}^T D \vec{j}} \\ &= \frac{(\vec{j} + x)^T(D - W)(\vec{j} + x)}{b\vec{j}^T D \vec{j}} + \frac{b^\mathfrak{V}(\vec{j} - x)^T(D - W)(\vec{j} - x)}{b\vec{j}^T D \vec{j}} - \frac{\mathfrak{V}b(\vec{j} - x)^T(D - W)(\vec{j} + x)}{b\vec{j}^T D \vec{j}} \\ &= \frac{[(\vec{j} + x) - b(\vec{j} - x)]^T(D - W)[(\vec{j} + x) - b(\vec{j} - x)]}{b\vec{j}^T D \vec{j}} \tag{11-2}\end{aligned}$$

فرض می‌کنیم $y = (\vec{j} + x) - b(\vec{j} - x)$ با توجه به اینکه $b = \frac{k}{k-1} = \frac{\sum_{x_i > \circ} d_i}{\sum_{x_i < \circ} d_i}$ نتیجه می‌گیریم که:

$$\begin{aligned} y^T D \vec{j} &= [(\vec{j} + x) - b(\vec{j} - x)]^T D \vec{j} & (۱۲-۲) \\ &= (\vec{j} + x)^T D \vec{j} - b(\vec{j} - x)^T D \vec{j} \\ &= \sum_{x_i > \circ} d_i - b \sum_{x_i < \circ} d_i = \circ \end{aligned}$$

از طرفی:

$$\begin{aligned} y^T D y &= [(\vec{j} + x) - b(\vec{j} - x)]^T D [(\vec{j} + x) - b(\vec{j} - x)] & (۱۳-۲) \\ &= (\vec{j} + x)^T D (\vec{j} + x) - b(\vec{j} + x)^T D (\vec{j} - x) - b(\vec{j} - x)^T D (\vec{j} + x) + b^2 (\vec{j} - x)^T D (\vec{j} - x) \\ &= \sum_{x_i > \circ} d_i + b^2 \sum_{x_i < \circ} d_i = b \sum_{x_i < \circ} d_i + b^2 \sum_{x_i < \circ} d_i \\ &= b \left(\sum_{x_i < \circ} d_i + b \sum_{x_i < \circ} d_i \right) = b \vec{j}^T D \vec{j} \end{aligned}$$

باکنار هم قرار دادن روابط ۱۱-۲ و ۱۲-۲ و ۱۳-۲ به مسأله بهینه‌سازی زیر می‌رسیم:

$$\min_x \frac{y^T L y}{y^T D y} \quad (۱۴-۲)$$

where

$$y^T D \vec{j} = \circ$$

$$y(i) \in \{1, -b\}$$

فرض می‌کنیم

$$r(y) = \frac{y^T L y}{y^T D y} = \lambda$$

اگر فرض کنیم y هر مقدار حقیقی‌ای را بگیرد یا به عبارتی مسأله مینیمم سازی فوق را از قید $y(i) \in \{1, -b\}$ آزاد کنیم، برای حل مسأله ۱۵-۲ کفایت مشتق $r(y)$ نسبت به y را مساوی صفر قرار می‌دهیم:

$$\begin{aligned}\nabla r(y) &= \frac{\nabla Ly(y^T Dy) - \nabla(y^T Ly)Dy}{(y^T Dy)^2} \\ &= \frac{\nabla Ly - \nabla r(y)Dy}{y^T Dy} \\ &= 0\end{aligned}$$

که به دستگاه ویژه تعمیم یافته زیر می‌رسیم:

$$Ly = \lambda Dy \quad (15-2)$$

برای حل دستگاه تعمیم یافته ۱۵-۲ ابتدا این دستگاه را به یک دستگاه ویژه استاندارد تبدیل می‌کنیم، لذا داریم:

$$Ly = \lambda Dy \implies D^{-\frac{1}{2}}LD^{-\frac{1}{2}}D^{\frac{1}{2}}y = \lambda D^{-\frac{1}{2}}Dy$$

فرض می‌کنیم $z = D^{\frac{1}{2}}y$

$$D^{-\frac{1}{2}}LD^{-\frac{1}{2}}z = \lambda z \quad (16-2)$$

همانطور که مشاهده می‌کنید رابطه فوق یک دستگاه مقدار ویژه استاندارد است که در آن مقدار ویژه z و

بردار ویژه برای ماتریس $D^{-\frac{1}{2}}LD^{-\frac{1}{2}}$ است. نشان می‌دهیم که $z_0 = D^{\frac{1}{2}}\vec{j}$ یک بردار ویژه با مقدار ویژه $\lambda = 0$

برای دستگاه ۱۶-۲ می‌باشد. برای این منظور کفایت ثابت کنیم که

$$(D^{-\frac{1}{2}}LD^{-\frac{1}{2}})z_0 = 0$$

با جایگذاری مقدار z_0 در رابطه فوق داریم:

$$D^{-\frac{1}{2}}LD^{-\frac{1}{2}}D^{\frac{1}{2}}\vec{j} = D^{-\frac{1}{2}}(D-W)D^{-\frac{1}{2}}D^{\frac{1}{2}}\vec{j} = D^{-\frac{1}{2}}(D-W)\vec{j} = D^{-\frac{1}{2}}(D\vec{j}-W\vec{j}) = D^{-\frac{1}{2}}(\vec{0}) = 0$$

پس $z = D^{-\frac{1}{2}} \vec{j}$ یک بردار ویژه با مقدار ویژه $\lambda = 0$ برای دستگاه ۲-۱۶ می‌باشد.
به ازای هر بردار حقیقی دلخواه f داریم:

$$\begin{aligned}
 f^T (D^{-\frac{1}{2}} L D^{-\frac{1}{2}}) f &= f^T (D^{-\frac{1}{2}} (D - W) D^{-\frac{1}{2}}) f \\
 &= f^T (I - D^{-\frac{1}{2}} W D^{-\frac{1}{2}}) f \\
 &= f^T f - f^T (D^{-\frac{1}{2}} W D^{-\frac{1}{2}}) f \\
 &= \sum_{i=1}^N f_i^2 - \sum_{i,j=1}^N \frac{f_i}{\sqrt{d_i}} W_{i,j} \frac{f_j}{\sqrt{d_j}} \\
 &= \frac{1}{2} \left(\sum_{i=1}^N f_i^2 - 2 \sum_{i,j=1}^N \frac{f_i}{\sqrt{d_i}} W_{i,j} \frac{f_j}{\sqrt{d_j}} + \sum_{j=1}^N f_j^2 \right) \\
 &= \frac{1}{2} \left(\sum_i d_i \frac{f_i^2}{d_i} - 2 \sum_{i,j=1}^N \frac{f_i}{\sqrt{d_i}} W_{i,j} \frac{f_j}{\sqrt{d_j}} + \sum_{j=1}^N d_j \frac{f_j^2}{d_j} \right) \\
 &= \frac{1}{2} \left(\sum_{i,j=1}^N W_{i,j} \frac{f_i^2}{d_i} - 2 \sum_{i,j=1}^N \frac{f_i}{\sqrt{d_i}} W_{i,j} \frac{f_j}{\sqrt{d_j}} + \sum_{i,j=1}^N W_{i,j} \frac{f_j^2}{d_j} \right) \\
 &= \frac{1}{2} \sum_{i,j=1}^N W_{i,j} \left(\frac{f_i}{\sqrt{d_i}} - \frac{f_j}{\sqrt{d_j}} \right)^2
 \end{aligned}$$

چون W_{ij} ها نامنفی هستند نتیجه می‌گیریم که

$$\forall f \in \mathbb{R}^n : f^T (D^{-\frac{1}{2}} L D^{-\frac{1}{2}}) f \geq 0$$

بنابراین ماتریس متقارن $D^{-\frac{1}{2}} L D^{-\frac{1}{2}}$ نیمه معین مثبت است. پس بنا به قضیه ۱-۴-۲، z کوچکترین بردار ویژه دستگاه ۲-۱۶ است و همه بردارهای ویژه این دستگاه بر هم عمودند. به‌ویژه، z_1 که دومین کوچکترین بردار ویژه است بر z عمود است پس شرایط قضیه ریلی (۱-۳-۱۳) برقرار است و نتیجه می‌گیریم که:

$$z_1 = \operatorname{argmin}_z \frac{z^T D^{-\frac{1}{2}} L D^{-\frac{1}{2}} z}{z^T z} \quad (17-2)$$

که در آن z_1 نشان‌دهنده اندیس بردار ویژه مینیمم‌کننده دستگاه ۲-۱۶ می‌باشد و مقدار این مینیمم برابر است با مقدار ویژه متناظر با بردار ویژه z_1 . با توجه به اینکه $z = D^{-\frac{1}{2}} y$ از رابطه ۲-۱۷ نتیجه می‌گیریم که

$$y_1 = \operatorname{argmin}_y \frac{y^T Ly}{y^T Dy} \quad (18-2)$$

که نشان دهنده اندیس بردار ویژه مینیمم کننده دستگاه ویژه ۲-۱۵ است. با معادل سازی مفاهیم بالا برای دستگاه ویژه ۲-۱۵ به نتایج زیر می‌رسیم:

۱. $y_0 = \vec{j}$ کوچکترین بردار ویژه ۲-۱۵ با مقدار ویژه صفر است.

۲. چون

$$0 = z_1^T z_0 = (D^{-\frac{1}{2}} y_1)^T (D^{-\frac{1}{2}} y_0) = y_1^T D^{-\frac{1}{2}} D^{-\frac{1}{2}} y_0 = y_1^T D \vec{j}$$

پس قید $0 = y^T D \vec{j}$ به ازای y_1 که دومین کوچکترین بردار ویژه ۲-۱۵ است، برقرار است. بنابراین دومین کوچکترین بردار ویژه دستگاه ۲-۱۵ که در اصطلاح بردار فیدلر^۱ نامیده می‌شود، یک جواب حقیقی مقدار برای مسأله برش نرمال است.

پس از افراز کردن گراف به دو خوشه مجزا با استفاده از دومین کوچکترین بردار ویژه یا بردار فیدلر، می‌توان با استفاده از سومین کوچکترین بردار ویژه هر کدام از خوشه‌های به دست آمده را به دو قسمت تقسیم کرد. در واقع این خط از استدلال را می‌توان به این صورت تعمیم داد که هر بار با استفاده از بردار ویژه کوچکتر بعدی گراف‌های موجود را به قسمت‌های کوچکتر افراز کنیم که شرایط افراز را دارا باشند.

در ادامه مسأله برش نرمال با روشی دیگر تقریب زده می‌شود.

برای انجام این کار، تابع هدف $Ncut$ رابطه (۲-۵)، برای تقسیم گراف به دو زیر بخش A و B در نظر گرفته ($k = 2$) و چند قضیه را بیان و اثبات می‌کنیم سپس مسأله برش نرمال را تقریب می‌زنیم.

تعریف ۲-۳-۱. بردار f به‌عنوان بردار شاخص، به صورت زیر تعریف می‌شود:

$$f_i = \begin{cases} \sqrt{\frac{\text{vol}(B)}{\text{vol}(A)}} & \text{اگر } v_i \in A \\ -\sqrt{\frac{\text{vol}(A)}{\text{vol}(B)}} & \text{اگر } v_i \in B \end{cases} \quad (19-2)$$

منظور از بردار شاخص، برداری است که پس از مشخص شدن مؤلفه‌های آن بتوان برای خوشه‌بندی داده‌ها از آن بردار استفاده نمود.

ورودی: تصویر

خروجی: گراف شباهت افراز شده

۱: گراف وزن دار $G = (V, E)$ را از روی تصویر ورودی بساز، به این صورت که هر یک از پیکسلهای تصویر را به عنوان یک رأس گراف در نظر گرفته و وزن هر یال گراف را که بیان کننده میزان شباهت بین دو رأس است و می تواند فاصله بین دو پیکسل تصویر در نظر گرفته شود محاسبه کن و اطلاعات را به D و W خلاصه کن.

۲: معادله $Lx = \lambda Dx$ را برای بردارهای ویژه متناظر با کوچکترین مقادیر ویژه ماتریس لاپلاسیان حل کن.

۳: از بردار ویژه متناظر با دومین کوچکترین مقدار ویژه برای افراز گراف استفاده کن.

۴: در مورد تقسیم کردن افراز جاری تصمیم بگیر و اگر لازم هست افراز بازگشتی را روی بخش های قطعه بندی شده انجام بده.

$$f^T Lf = \text{vol}(V) Ncut(A, B)$$

برهان. با استفاده از قضیه ۱-۴-۲ و تعریف بردار شاخص (رابطه ۲-۱۹)، رابطه $f^T Lf$ را بسط می دهیم.

$$\begin{aligned} f^T Lf &= \frac{1}{2} \sum_{i,j=1}^N w_{i,j} (f_i - f_j)^2 \\ &= \frac{1}{2} \sum_{i \in A, j \in B} w_{i,j} \left(\sqrt{\frac{\text{vol}(B)}{\text{vol}(A)}} + \sqrt{\frac{\text{vol}(A)}{\text{vol}(B)}} \right)^2 + \frac{1}{2} \sum_{i \in B, j \in A} w_{i,j} \left(-\sqrt{\frac{\text{vol}(B)}{\text{vol}(A)}} - \sqrt{\frac{\text{vol}(A)}{\text{vol}(B)}} \right)^2 \\ &= \text{cut}(A, B) \left(\frac{\text{vol}(B)}{\text{vol}(A)} + \frac{\text{vol}(A)}{\text{vol}(B)} + 2 \right) \\ &= \text{cut}(A, B) \left(\frac{\text{vol}(A) + \text{vol}(B)}{\text{vol}(B)} + \frac{\text{vol}(A) + \text{vol}(B)}{\text{vol}(A)} \right) \\ &= \text{vol}(V) Ncut(A, B) \end{aligned}$$

□

$$f^T Df = \text{vol}(V) \quad \text{قضیه ۲-۳-۳}$$

برهان.

$$\begin{aligned}
 f^T Df &= [f_1, \dots, f_N] \times \begin{bmatrix} d_{11} & & & \\ & d_{22} & & \\ & & \ddots & \\ & & & d_{NN} \end{bmatrix} \times \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_N \end{bmatrix} \\
 &= \sum_{i \in A} f_i d_{ii} f_i + \sum_{i \in \bar{A}} f_i d_{ii} f_i \\
 &= \sum_{i \in A} d_{ii} \left(\frac{\text{vol}(\bar{A})}{\text{vol}(A)} \right) + \sum_{i \in \bar{A}} d_{ii} \left(\frac{\text{vol}(A)}{\text{vol}(\bar{A})} \right) \\
 &= \text{vol}(A) \left(\frac{\text{vol}(\bar{A})}{\text{vol}(A)} \right) + \text{vol}(\bar{A}) \left(\frac{\text{vol}(A)}{\text{vol}(\bar{A})} \right) \\
 &= \text{vol}(V)
 \end{aligned}$$

□

قضیه ۲-۳-۴. $(Df)^T \mathbf{1} = 0$.

برهان.

$$\begin{aligned}
 (Df)^T \mathbf{1} &= [f_1, \dots, f_N] \times \begin{bmatrix} d_{11} & & & \\ & d_{22} & & \\ & & \ddots & \\ & & & d_{NN} \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \\
 &= \sum_{i \in A} d_{ii} \sqrt{\frac{\text{vol}(\bar{A})}{\text{vol}(A)}} - \sum_{i \in \bar{A}} d_{ii} \sqrt{\frac{\text{vol}(A)}{\text{vol}(\bar{A})}} \\
 &= \text{vol}(A) \sqrt{\frac{\text{vol}(\bar{A})}{\text{vol}(A)}} - \text{vol}(\bar{A}) \sqrt{\frac{\text{vol}(A)}{\text{vol}(\bar{A})}} \\
 &= \sqrt{\text{vol}(A)\text{vol}(\bar{A})} - \sqrt{\text{vol}(\bar{A})\text{vol}(A)} \\
 &= 0
 \end{aligned}$$

□

با توجه به قضایای ذکر شده فوق می‌توان مسأله برش نرمال را به صورت زیر مورد بررسی قرار داد. برای این منظور رابطه ۲-۵ را به صورت زیر بازنویسی می‌کنیم:

$$\begin{aligned}
 ncut(A, B) &= \frac{cut(A, B)}{vol(A)} + \frac{cut(A, B)}{vol(B)} & (20-2) \\
 &= \frac{vol(V)}{vol(V)} \left(\frac{cut(A, B)}{vol(A)} + \frac{cut(A, B)}{vol(B)} \right) \\
 &= \frac{vol(A) + vol(B)}{vol(V)} \left(\frac{cut(A, B)}{vol(A)} + \frac{cut(A, B)}{vol(B)} \right) \\
 &= \frac{cut(A, B)}{vol(V)} \left(2 + \frac{vol(A)}{vol(B)} + \frac{vol(B)}{vol(A)} \right) \\
 &= \frac{cut(A, B)}{vol(V)} \left(\sqrt{\frac{vol(A)}{vol(B)}} + \sqrt{\frac{vol(B)}{vol(A)}} \right)^2
 \end{aligned}$$

با استفاده از قضیه ۲-۳-۲ مشاهده می‌شود

$$cut(A, B) \left(\sqrt{\frac{vol(A)}{vol(B)}} + \sqrt{\frac{vol(B)}{vol(A)}} \right)^2 = f^T L f$$

با جایگذاری این مقدار در رابطه ۲۰-۲ معادله $f^T L f = vol(V) Ncut(A, B)$ بدست می‌آید. می‌توان این معادله را به صورت زیر نوشت:

$$Ncut(A, B) = \frac{f^T L f}{vol(V)} \quad (21-2)$$

با استفاده از قضیه ۲-۳-۳ مقدار متناظر $vol(V)$ در رابطه ۲۱-۲ جایگذاری می‌شود. بنابراین خواهیم داشت:

$$Ncut(A, B) = \frac{f^T L f}{f^T D f}$$

از طرفی قضیه ۲-۳-۴ نیز همواره برقرار است. بنابراین مسئله برش نرمال به صورت مسأله بهینه سازی زیر تعریف می‌شود:

$$\min Ncut(A, B) = \min \frac{f^T Lf}{f^T Df} \quad (22-2)$$

s.t.

$$(Df)^T \vec{j} = 0$$

اگر مسئله ساده شود و در رابطه ۲-۲۲ از مقید بودن f به مقادیر گسسته صرف نظر نموده و f برداری حقیقی در نظر گرفته شود، می توان با حل نمودن دستگاه تجزیه ویژه تعمیم یافته $Lf = \lambda Df$ مینیمم مقدار کسر را در رابطه ۲-۲۲ به دست آورد اما قید $(Df)^T \vec{j} = 0$ نیز می بایست برقرار شود. ابتدا نشان می دهیم چطور حل مسئله برش نرمال منجر به دستگاه ویژه تعمیم یافته می شود سپس درباره نحوه برقراری قید $(Df)^T \vec{j} = 0$ بحث می شود. در مسأله برش نرمال هدف یافتن بردار f به صورتی است که بتواند مقدار تابع هدف $r(f) = \frac{f^T Lf}{f^T Df}$ را مینیمم نماید. بنابراین باید از تابع $Ncut$ نسبت به f مشتق گرفت:

$$\nabla r(f) = \frac{\nabla Lf(f^T Df) - \nabla(f^T Lf)Df}{(f^T Df)^2} = \frac{\nabla Lf - \nabla r(f)Df}{f^T Df} = 0$$

در نتیجه خواهیم داشت:

$$Lf = r(f)Df$$

همانطور که مشاهده می شود جواب مشتق به دستگاه تجزیه ویژه تعمیم یافته برای ماتریس L منتهی می شود.

$r(f)$ در واقع مقدار ویژه است و با λ نمایش داده می شود $r(f) = \lambda$. می توان نوشت:

$$Lf = \lambda Df \quad (23-2)$$

قید $(Df)^T \vec{j} = 0$ با حل دستگاه ویژه تعمیم یافته ۲-۲۳ خود به خود برآورده می شود. برای نشان دادن این موضوع، باید دستگاه ۲-۲۳ به صورت یک مسئله ویژه استاندارد نوشته شود. برای انجام این کار، متغیر $g = D^{-\frac{1}{2}} f$ را تعریف کرده و در دستگاه ویژه تعمیم یافته ۲-۲۳ جایگذاری می کنیم:

$$LD^{-\frac{1}{2}} g = \lambda DD^{-\frac{1}{2}} g \implies LD^{-\frac{1}{2}} g = \lambda D^{\frac{1}{2}} g \implies D^{-\frac{1}{2}} LD^{-\frac{1}{2}} g = \lambda g$$

لذا دستگاه ویژه استاندارد

$$(D^{-\frac{1}{2}}LD^{-\frac{1}{2}})g = \lambda g \quad (24-2)$$

را به دست می آوریم که در آن λ مقدار ویژه و g بردار ویژه برای ماتریس $(D^{-\frac{1}{2}}LD^{-\frac{1}{2}})$ می باشد. نشان می دهیم که $g_0 = D^{-\frac{1}{2}}\vec{j}$ یک بردار ویژه با مقدار ویژه $\lambda = 0$ برای دستگاه ۲-۲۴ می باشد. برای این منظور کافیت ثابت کنیم که

$$(D^{-\frac{1}{2}}LD^{-\frac{1}{2}})g_0 = 0$$

با جایگذاری مقدار g_0 در رابطه فوق داریم:

$$D^{-\frac{1}{2}}LD^{-\frac{1}{2}}D^{-\frac{1}{2}}\vec{j} = D^{-\frac{1}{2}}(D-W)D^{-\frac{1}{2}}D^{-\frac{1}{2}}\vec{j} = D^{-\frac{1}{2}}(D-W)\vec{j} = D^{-\frac{1}{2}}(D\vec{j}-W\vec{j}) = D^{-\frac{1}{2}}(\vec{0}) = 0$$

پس $g_0 = D^{-\frac{1}{2}}\vec{j}$ یک بردار ویژه با مقدار ویژه $\lambda = 0$ برای دستگاه ۲-۲۴ می باشد. با توجه به قضیه ۱-۴-۲ ماتریس L نیمه معین مثبت است و کوچکترین مقدار ویژه آن همیشه صفر است، پس $D^{-\frac{1}{2}}LD^{-\frac{1}{2}}$ نیز نیمه معین مثبت و متقارن است لذا همه مقادیر ویژه آن نامنفی خواهند بود و g_0 کوچکترین بردار ویژه دستگاه ۲-۲۴ می باشد. همچنین $\vec{j} = D^{-\frac{1}{2}}D^{-\frac{1}{2}}\vec{j} = D^{-\frac{1}{2}}g_0 = f_0$ نیز کوچکترین بردار ویژه دستگاه ویژه تعمیم یافته ۲-۲۳ خواهد بود. چون بردارهای ویژه دستگاه ۲-۲۴ بر یکدیگر عمود هستند، داریم:

$$g_0^T g_0 = 0 \Rightarrow (D^{-\frac{1}{2}}f_0)^T D^{-\frac{1}{2}}\vec{j} = f_0^T D\vec{j} = 0$$

که در آن f_0 دومین کوچکترین بردار ویژه دستگاه ویژه تعمیم یافته ۲-۲۳ است پس نتیجه می گیریم که قید $(Df)^T \vec{j} = 0$ به ازای f_0 برقرار است. چون g_0 بر g_0 عمود است با استفاده از قضیه ریلی ریتز (۱-۳-۱۳) داریم:

$$g_1 = \operatorname{argmin}_g \frac{g^T D^{-\frac{1}{2}}LD^{-\frac{1}{2}}g}{g^T g} \quad (25-2)$$

و در نتیجه



شکل ۲-۵: یک تصویر سطح خاکستری از بازی بیس بال

$$f_1 = \operatorname{argmin}_g \frac{f^T L f}{f^T D f} \quad (26-2)$$

بنابراین دومین کوچکترین بردار ویژه دستگاه ویژه تعمیم یافته ۲-۲۳ جواب مسئله تقریب برش نرمال خواهد بود.

شکل ۲-۵ یک تصویر که قصد داریم روی آن قطعه بندی را انجام دهیم نشان می دهد. برای قطعه بندی

تصویر ۲-۵ به صورت زیر عمل می کنیم:

ابتدا گراف $G = (V, E)$ را با در نظر گرفتن هر پیکسل از تصویر به عنوان یک رأس و پیوند بین دو پیکسل به عنوان یال اتصال دهنده رئوس متناظر ایجاد می کنیم درحالیکه وزن هر یال نشان دهنده احتمال متعلق بودن دو پیکسل به یک شیء می باشد. با استفاده از مقدار روشنایی پیکسل ها و موقعیت مکانی آنها، وزن یال اتصال دهنده دو رأس i و j گراف را بصورت زیر تعریف می کنیم:

$$W_{ij} = e^{\frac{-\|F(i)-F(j)\|_1}{\sigma_I}} * \begin{cases} e^{\frac{-\|X(i)-X(j)\|_1}{\sigma_X}} & \text{if } \|X(i) - X(j)\|_1 < r \\ \cdot & \text{if otherwise} \end{cases} \quad (27-2)$$

که در آن $X(i)$ و $F(i)$ به ترتیب میزان شدت روشنایی و موقعیت مکانی پیکسل i ام می باشند. σ_X و σ_I انحراف معیار مرتبط با توابع نرمال اختلاف شدت روشنایی و اختلاف فاصله مکانی دو پیکسل هستند، به عبارت دیگر این پارامترها سرعت کاهش تابع شباهت را با افزایش اختلاف شدت روشنایی و فاصله مکانی دو پیکسل کنترل می کنند. برای بردارهای ویژه متناظر با کوچکترین مقادیر ویژه دستگاه ۲-۲۳ را حل کنید.

شکل ۲-۶ کوچکترین بردارهای ویژه محاسبه شده برای قطعه بندی تصویر ۲-۵ برای دستگاه ویژه تعمیم

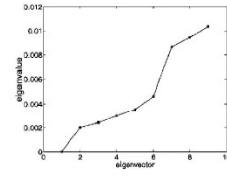
یافته با ماتریس وزن تعریف شده در رابطه ۲-۲۷ را نشان می دهد.



(ج) سومین کوچکترین بردار ویژه



(ب) دومین کوچکترین بردار ویژه



(ا) نمودار کوچکترین بردارهای ویژه دستگاه ویژه تعمیم یافته ۲-۱۲



(و) ششمین کوچکترین بردار ویژه



(ه) پنجمین کوچکترین بردار ویژه



(د) چهارمین کوچکترین بردار ویژه



(ط) نهمین کوچکترین بردار ویژه



(ح) هشتمین کوچکترین بردار ویژه



(ز) هفتمین کوچکترین بردار ویژه

شکل ۲-۶: استفاده از دومین تا نهمین کوچکترین بردار ویژه ماتریس لاپلاسیان گراف شباهت تصویر شکل ۲-۵ برای قطعه‌بندی آن

بعد از محاسبه بردارهای ویژه، گراف را با استفاده از دومین کوچکترین بردار ویژه، به دو قسمت افراز می‌کنیم، در حالت ایده‌آل، بردار ویژه دارای مقادیر گسسته است و علامت مقادیرها چگونگی افراز گراف را نشان می‌دهد، بدین صورت که داده‌هایی که علامت مثبت دارند در یک خوشه و داده‌هایی که علامت منفی دارند درون خوشه دیگر قرار می‌گیرند و بدین ترتیب گراف به دو خوشه مجزا تقسیم می‌شود.

بهر حال، بردارهای ویژه می‌توانند مقادیر پیوسته بگیرند در این صورت برای افراز کردن گراف به دو قسمت، نیاز به یک نقطه جدا کننده داریم. راه‌های مختلفی برای انتخاب نقطه جدا کننده وجود دارد. می‌توان نقطه جدا کننده را صفر یا مقدار میانه نقاط در نظر گرفت یا به جستجوی نقطه جدا کننده‌ای پرداخت که بهترین مقدار $Ncut(A, B)$ را داشته باشد.

پس از اینکه گراف به دو قسمت افراز شد، به طور بازگشتی الگوریتم را روی دو افراز بدست آمده اجرا می‌کنیم یا از بردارهای ویژه بعدی برای افراز زیرگراف‌های ایجاد شده استفاده می‌کنیم.

الگوریتم ۲-۲ $Ncut$ دوتایی بازگشتی

ورودی: مجموعه‌ای از ویژگیها به صورت گراف شباهت

خروجی: گراف شباهت افراز شده

گراف وزن دار $G = (V, E)$ را بساز، وزن هر یال را محاسبه کن و ماتریسهای D و W را ایجاد کن.

۲: دستگاه ویژه تعمیم یافته $2-23$ را بساز.

بردار ویژه متناظر با دومین کوچکترین مقدار ویژه را برای افراز گراف به وسیله پیدا کردن نقطه جداکننده‌ای

که کمترین مقدار $Ncut$ را دارد، پیدا کن.

۴: افراز را بصورت بازگشتی تا رسیدن به قطعات مورد نیاز گراف انجام بده.

۴-۲ برش k تایی همزمان با چند بردار ویژه

برای انجام یک خوشه‌بندی k تایی می‌توان این کار را به دو روش بازگشتی و چندگانه انجام داد.

در خوشه‌بندی طیفی چندگانه از k بردار ویژه ماتریس لاپلاسیان به طور همزمان برای افراز گراف استفاده

می‌شود [۱۷] که افراز k تایی نامیده می‌شود. در این روش ابتدا بردارهای ویژه متناظر با k تا از کوچکترین مقادیر

ویژه انتخاب شده و خوشه‌بندی یا براساس علامت داده‌ها در بردارهای ویژه دوم تا k ام انجام می‌گیرد که بر این

اساس، داده‌هایی که تغییر علامت روی این k بردار برای آنها یکسان باشد، در یک گروه قرار خواهند گرفت.

در روش بازگشتی، بعد از تقسیم داده‌ها به دو خوشه بررسی می‌شود که آیا تعداد خوشه‌ها متناسب با نیاز است

یا خیر. اگر لازم باشد، همین فرآیند مجدداً بر روی زیر خوشه‌های بدست آمده اعمال می‌شود و تا رسیدن به تعداد

خوشه مطلوب کار ادامه می‌یابد.

۵-۲ الگوریتم‌های خوشه‌بندی طیفی

الگوریتم‌های خوشه‌بندی طیفی در مقابل الگوریتم‌های خوشه‌بندی مرسوم مانند K-means [۱]، مزایای زیادی

دارند و در اغلب موارد نتایج این الگوریتم‌ها بسیار بهتر می‌باشد. علاوه بر این، اجرا و پیاده‌سازی ساده‌تری داشته

و با نظریه طیفی گراف [۷] رابطه تنگاتنگی دارند. همچنین با استفاده از روش‌های جبرخطی به خوبی قابل تفسیر

و حل می‌باشند.

در این الگوریتم‌ها با استفاده از بردارهای ویژه ماتریس لاپلاسیان یک گراف، گراف را افراز می‌کنیم. چون

کوچکترین مقدار ویژه ماتریس لاپلاسیان صفر است و بردار ویژه واحد متناظر با این مقدار ویژه هیچ اطلاعاتی را

در مورد ساختار گراف نمی‌دهد. بنابراین از بردار ویژه دوم (بردار فیدلر) برای افراز استفاده می‌کنیم. بدین صورت

که با تقسیم کردن عناصر آن بر میانه آنها، دو افراز به دست می‌آید. تمام الگوریتم‌های ذکر شده در ادامه با دومین

کوچکترین بردار ویژه (بردار فیدلر) کار می‌کنند و با دوبخشی کردن گراف سروکار دارند به این معنی که مجموعه V را به دو مجموعه مجزا تقسیم می‌کنند. برای بدست آوردن قطعات بیشتر گراف، این الگوریتم‌ها را می‌توان به صورت بازگشتی به کار برد تا زمانی که قطعات به اندازه کافی کوچک و به تعداد مورد نظر باشد. در حقیقت هدف، پیدا کردن یال مجزای E^T - زیر مجموعه‌ای از E که کمترین وزن را دارد بطوریکه حذف E^T از E گراف G را به دو زیر گراف جدا از هم G_1 و G_2 با رئوس به ترتیب V_1 و V_2 تقسیم می‌کند - است.

طریقه بدست آوردن یال مجزای E^T را در ادامه شرح می‌دهیم.

گراف همبند $G = (V, E)$ را در نظر بگیرید، فرض کنید $\vec{v} = (v_0, v_1, \dots, v_n)$ دومین کوچکترین بردار ویژه ماتریس لاپلاسی (بردار فیدلر) گراف G باشد، چون گراف، همبند است بنا به قضیه ۱-۴-۳ تعداد دفعات تکرار مقدار ویژه صفر ماتریس لاپلاسی، یک بار است لذا دومین مقدار ویژه آن بزرگتر از صفر است.

ایده افزای طیفی پیدا کردن یک مقدار تفکیک کننده s است بطوریکه رئوس G را با استفاده از بردار فیدلر و مقدار بدست آمده برای s افزای کنیم. به این صورت که ابتدا هر یک از رئوس گراف را به هر یک از عناصر بردار فیدلر ماتریس لاپلاسی انتساب می‌دهیم و بعد از به دست آوردن s رئوس متناظر با $v_i \leq s$ را به یک مجموعه و رئوس متناظر با $v_i \geq s$ به یک مجموعه دیگر اختصاص می‌دهیم. در واقع، رئوس G را به دو مجموعه از رئوس V_1 و V_2 تفکیک می‌کنیم که اولین مجموعه از رئوس مربوط به $v_i \leq s$ و دومین مجموعه مربوط به $v_i \geq s$ است. یکی از راههای بدست آوردن s این است که میانه عناصر بردار فیدلر را حساب کنیم. اگر فقط یک جزء مساوی با میانه وجود داشته باشد آنگاه سائز V_1 و V_2 حداکثر در یک عدد متفاوت است و اگر چندین جزء مساوی با میانه وجود داشته باشد به طور دلخواه رأس‌ها را طوری به V_1 و V_2 اختصاص می‌دهیم که سائز V_1 و V_2 حداکثر در یک عدد فرق داشته باشند.

فرض کنید V_1^T نشان دهنده مجموعه رئوسی از V_1 است که با بعضی رئوس در V_2 مجاور هستند و همچنین V_2^T نشان دهنده مجموعه رئوسی از V_2 است که با بعضی رئوس در V_1 مجاور هستند. E_1 را مجموعه یال‌ها با هر دو رأس انتهایی در V_1 در نظر بگیرید و E_2 را مجموعه یال‌ها با هر دو رأس انتهایی در V_2 . فرض کنید $E^T \in E$ مجموعه یال‌های G با یک رأس در V_1^T و یک رأس در V_2^T باشد. بنابراین E^T یک یال جدا کننده در G است و V_1^T و V_2^T رئوس جداکننده متناظر با یال جداکننده E^T هستند. بنابراین برای هر یال جدا کننده دو رأس جدا کننده وجود دارد. در هر صورت ما به این رئوس جداکننده کاری نداریم.

افزای k تایی گراف G تقسیم کردن رئوس آن به k زیر مجموعه مجزا از اندازه تقریباً برابر با $\frac{n}{k}$ هست.

برای این منظور الگوریتم دوبخشی طیفی را روی گراف G تا زمانی که تعداد k زیر گراف مورد نظر بدست بیاید انجام می‌دهیم.

دو حالت وجود دارد: اگر k توانی از دو باشد از دوبخشی طیفی بازگشتی ساده استفاده می‌کنیم و اگر k توانی

الگوریتم ۲-۳ دو بخشی طیفی

ورودی: گراف $G = (V, E)$

خروجی: گرافهای $G_1 = (V_1, E_1)$ و $G_2 = (V_2, E_2)$

ماتریس لاپلاسیان گراف (۱-۴) را بدست بیاور.

بردار فیدلر v را محاسبه کن. (دومین کوچکترین بردار ویژه ماتریس لاپلاسیان گراف (G))

۳: میانه v را پیدا کن.

برای هر رأس i از G

اگر $v_i \leq \text{median}$ رأس i را در افراز V_1 بگذار.

در غیر اینصورت رأس i را در افراز V_2 بگذار.

اگر $|V_1 - V_2| > 1$ تعدادی از رئوس با اجزای مساوی با میانه را از V_1 به V_2 منتقل کن برای اینکه تفاوت اندازه مجموعه‌های V_1 و V_2 به حداکثر یک واحد برسد.

۶: یال جداکننده E^T را بدست آور.

گرافهای $G_1 = (V_1, E_1)$ و $G_2 = (V_2, E_2)$ را بساز.

الگوریتم ۲-۴ دو بخشی طیفی بازگشتی

ورودی: گراف $G = (V, E)$ ؛ یک عدد صحیح k (تعداد افرازهای مورد نظر)

خروجی: گرافهای $G_1 = (V_1, E_1)$ و... و $G_k = (V_k, E_k)$

۱: اگر $\frac{k}{2} > 1$

دو بخشی بازگشتی $(G_1, k/2)$

دو بخشی بازگشتی $(G_2, k/2)$

۲: برگرداندن افرازهای G_1 و... و G_k

۳: پایان

از دو نباشد از دو بخشی طیفی بازگشتی اصلاح شده استفاده می‌کنیم.

الگوریتم‌های خوشه‌بندی طیفی به دلیل اجرای ساده و سرعت قابل قبول مورد توجه قرار گرفته‌اند و مشکل

بهینه‌های محلی را نیز ندارند. این الگوریتم‌ها را می‌توان در سه گروه دسته‌بندی نمود:

• طیفی بازگشتی^۱: در این الگوریتم‌ها با استفاده از یک بردار ویژه مجموعه داده به دو قسمت افراز می‌شود

و برای هر جزء این روند به صورت بازگشتی تا رسیدن به تعداد خوشه مطلوب تکرار می‌گردد.

• طیفی چندگانه^۲: در این روش برای خوشه‌بندی از چند بردار ویژه به صورت همزمان استفاده می‌شود و

داده‌ها مستقیماً به تعداد خوشه مورد نظر افراز می‌شوند.

^۱Recursive spectral

^۲Multway spectral

الگوریتم ۲-۵ دوبخشی طیفی اصلاح شده

- ورودی: گراف $G = (V, E)$ ؛ چارک (از اندازه $|V_1|$)
خروجی: گرافهای $G_1 = (V_1, E_1)$ و $G_2 = (V_2, E_2)$ از اندازه های به ترتیب $|V_1|$ و $|V_2| = |V| - |V_1|$
- ۱: بردار ویژه v ماتریس لاپلاسیان گراف (۴-۱) را محاسبه کن.
 - ۲: چارک v را محاسبه کن.
 - ۳: برای هر رأس i از G
اگر چارک $v_i \leq$ آنگاه رأس i را در افراز V_1 قرار بده
در غیر اینصورت رأس i را در افراز V_2 قرار بده
 - ۴: اگر $|V_1 - V_2| > 1$ تعدادی از رئوس با اجزای مساوی با چارک را از V_1 به V_2 منتقل کن برای اینکه تفاوت اندازه مجموعه های V_1 و V_2 به حداکثر یک واحد برسد.
 - ۵: یال مجزای E^T را بدست آور.
 - ۶: گرافهای $G_1 = (V_1, E_1)$ و $G_2 = (V_2, E_2)$ را بساز.

- غیر طیفی^۱: الگوریتم های ساده که داده ها را سریع خوشه بندی می کنند و در رابطه با الگوریتم های طیفی چندگانه به کار می روند [۱۹].

۲-۵-۱ الگوریتم برش نرمال

این الگوریتم توسط شی و مالیک^۲ برای قطعه بندی تصاویر ارائه شده است و معیار برش نرمال را طبق آنچه که در بخش ۲-۲-۳ ارائه شد، معرفی نمودند. در قسمت ۲-۳ به مسئله افراز و برش گراف و ساده سازی این مسئله پرداخته شد. همانطور که در بخش ۲-۳-۱ بحث گردید، برای حل مسئله برش نرمال می توان از بردارهای ویژه دستگاه ویژه تعمیم یافته ۲-۲۳ و یا بردارهای ویژه دستگاه ویژه استاندارد ۲-۲۴ با در نظر گرفتن $g = D^{\dagger} f$ استفاده نمود.

۲-۶ خوشه بندی کرنلی

نسخه مبتنی بر کرنل بسیاری از الگوریتم های خوشه بندی برای تولید ابرصفحات غیرخطی جداساز ارائه شده است. در این الگوریتم ها با تبدیل فضای ورودی به فضای غیرخطی انتظار می رود تا خوشه ها واضح تر و منسجم تر باشند. بین روش های خوشه بندی طیفی و کرنلی ارتباط تنگاتنگی وجود دارد [۲۰].

^۱Non spectral ^۲Shi & Malik

الگوریتم ۲-۶ الگوریتم برش نرمال با استفاده از بردارهای ویژه دستگاه ویژه تعمیم یافته

ورودی: تصویر

خروجی: گراف خوشه بندی شده

۱: ساخت گراف وزندار $G(V, A)$ با استفاده از مفهوم شباهت داده ها و رابطه زیر

$$a_{ij} = \exp\left(-\frac{\|f_i - f_j\|^2}{2\sigma_1^2}\right) \times \begin{cases} \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma_1^2}\right) & \text{if } \|x_i - x_j\| < R \\ 0 & \text{otherwise} \end{cases} \quad (2-28)$$

که x_i موقعیت i امین پیکسل و f_i بردار ویژگی از صفات پیکسل مانند روشنایی و غیره (با توجه به خواستگاه اولیه این روش در تصاویر) می باشد. با کمک حد آستانه R می توان میزان تنگی ماتریس شباهت را با توجه به تعداد پیکسل های همسایه اثرگذار تعیین کرد.

۲: محاسبه ماتریس درجه D

۳: محاسبه ماتریس لاپلاسیان غیر نرمال

۴: محاسبه دومین بردار ویژه متناظر با دومین کوچکترین مقدار ویژه λ برای دستگاه ویژه تعمیم یافته

$$Lf = \lambda Df$$

۵: قطعه بندی گراف G بدین صورت که رئوس از گراف که متناظر با درایه های هم علامت بردار فیدلر می باشند درون یک خوشه قرار می گیرند.

الگوریتم ۲-۷ الگوریتم برش نرمال با استفاده از بردارهای ویژه دستگاه ویژه استاندارد

ورودی: تصویر به صورت گراف

خروجی: گراف خوشه بندی شده

۱: ساخت گراف وزندار $G(V, A)$ با استفاده از مفهوم شباهت داده ها و رابطه ۲-۱۵

۲: محاسبه ماتریس درجه D

۳: محاسبه ماتریس لاپلاسیان با استفاده از رابطه ۳-۲

۴: محاسبه دومین بردار ویژه متناظر با دومین کوچکترین مقدار ویژه λ برای دستگاه ویژه

$$D^{-\frac{1}{2}} L D^{-\frac{1}{2}} g = \lambda g$$

۵: استفاده از بردار g بدست آمده از گام ۴ و انجام تبدیل $f = D^{-\frac{1}{2}} g$ و قطعه بندی گراف G با استفاده از بردار f حاصل.

فصل ۳

نگاه متحد به K-means کرنلی، خوشه‌بندی طیفی و برش گراف

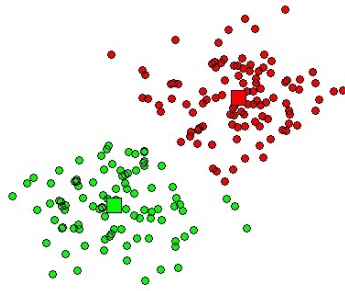
۱-۳ مقدمه

در این فصل ابتدا به صورت خلاصه روش خوشه‌بندی kmeans [۱] را توضیح داده سپس به بیان روش خوشه‌بندی kmeans کرنلی^۱ از دسته الگوریتم‌های خوشه‌بندی کرنلی به علت قرابت آن با الگوریتم‌های خوشه‌بندی طیفی می‌پردازیم و در ادامه نشان می‌دهیم که دوروش K-means کرنلی و افراز گراف که از لحاظ ظاهری با هم متفاوت می‌باشند، برای خوشه‌بندی داده‌هایی که به طور خطی جداپذیر نیستند، هم ارز می‌باشند.

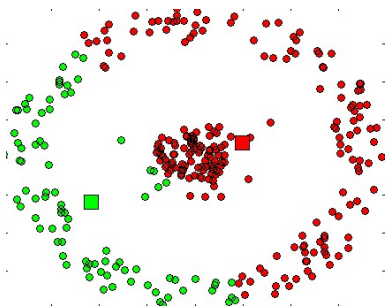
الگوریتم Kmeans کرنلی [۲۱] نسخه تعمیم‌یافته Kmeans در فضای غیرخطی می‌باشد. در [۲۲] از نظر ریاضی اثبات شده است که نسخه وزن‌دار این روش معادل با افراز یک گراف وزن‌دار می‌باشد و ارتباط آن با خوشه‌بندی طیفی بررسی شده است.

به طور کلی در خوشه‌بندی داده‌ها با دو نوع داده مواجه هستیم: داده‌هایی که به طور خطی جداپذیر هستند و داده‌هایی که به طور خطی جداپذیر نیستند. برای خوشه‌بندی داده‌هایی که به طور خطی جداپذیر می‌باشند می‌توان از روش‌های خوشه‌بندی طیفی و روش K-means [۱] استفاده کرد و برای خوشه‌بندی داده‌هایی که به طور خطی جداپذیر نیستند برای راحتی کار می‌توان داده‌ها را به یک فضای با ابعاد بالاتر نگاشت داد سپس با روش‌های موجود به خوشه‌بندی آنها پرداخت.

^۱Kernel Kmeans



شکل ۳-۱: یک نمونه داده جداپذیر خطی



شکل ۳-۲: یک نمونه داده جداناپذیر خطی

الگوریتم K-means کرنلی [۲۱] تعمیم داده شده الگوریتم K-means استاندارد [۲۳] است. الگوریتم K-means کرنلی با نگاشت داده‌ها به فضایی با ابعاد بالاتر، خوشه‌هایی که در فضای ورودی (اولیه) داده‌ها، جدایی پذیر خطی نیستند را پیدا می‌کند. الگوریتم‌های افزایش گراف روی خوشه‌بندی رأس‌های یک گراف تمرکز دارند. روش‌های طیفی برای حل تابع هدف افزایش گراف به عنوان مثال نسبت برش [۱۶] و برش نرمال شده [۱۷] استفاده می‌شوند. از K-means کرنلی وزن‌دار می‌توان برای بهینه‌سازی تابع هدف افزایش گراف استفاده نمود و بالعکس روش‌های طیفی برای K-means کرنلی به کار گرفته شوند. زمانیکه محاسبه بردارهای ویژه، سخت است (مثلاً اگر بخواهیم بردارهای ویژه یک ماتریس بزرگ را محاسبه کنیم)، الگوریتم K-means کرنلی وزن‌دار مطلوب‌تر از روش‌های طیفی است. تجزیه و تحلیل‌های انجام شده نشان می‌دهد که با انتخاب مناسب ماتریس کرنل نقاط، می‌توان از K-means کرنلی وزن‌دار برای بهبود بخشیدن به تابع هدف‌های مسأله افزایش گراف مانند تابع هدف نسبت برش، برش نرمال شده، نسبت وابستگی و... استفاده نمود.

۱-۱-۳ الگوریتم K-means

الگوریتم K-means یکی از روش های خوشه بندی ساده و سریع است [۱] این الگوریتم دارای یک پارامتر به نام k است که تعداد خوشه هایی که باید به دست آید را مشخص می کند. الگوریتم K-means پایه به صورت زیر است:

k داده را به عنوان مرکز خوشه انتخاب می کنیم، سپس فواصل بقیه داده ها با مرکز خوشه ها را تعیین می کنیم و داده هایی که به مرکز هر خوشه نزدیک تر هستند را در آن خوشه قرار می دهیم. میانگین هر خوشه را به عنوان مرکز جدید خوشه انتخاب می کنیم این مراحل را تا زمانی ادامه می دهیم که خوشه ها بدون تغییر باقی بمانند. به طور معمول، مرکز خوشه های اولیه به صورت تصادفی از میان نمونه های اولیه گزینش می شوند. به همین دلیل، مرکز خوشه های اولیه در دو خوشه بندی مستقل K-means می توانند متفاوت باشند. این موضوع موجب می شود که خوشه های به جا مانده از دو اجرای مختلف K-means با هم متفاوت باشند. بنابراین همواره به بهینه سراسری نمی رسد اما ممکن است به بهینه محلی برسد. در الگوریتم K-means می توان از معیارهای فاصله ی گوناگون بهره گرفت و خوبی یا بدی به کارگیری آن معیار بستگی به نوع داده هایی دارد که باید خوشه بندی شوند.

اگر یک ماتریس $n * m$ را به عنوان ورودی به K-means بدهیم K-means این ماتریس را m داده با خصوصیت در نظر می گیرد و اگر K را دو در نظر بگیریم داده ها را به دو دسته تقسیم می کند با این شرط که اعضای در یک دسته قرار می گیرند که خصوصیات آنها به یکدیگر نزدیکتر باشد یعنی سطر متناظر با آنها در ماتریس ورودی شباهت بیشتری به یکدیگر داشته باشد.

مجموعه بردار a_1, a_2, \dots, a_n را در نظر بگیرید، الگوریتم K-means برای پیدا کردن خوشه های $\pi_1, \pi_2, \dots, \pi_k$ با استفاده از min کردن تابع زیر به کار می رود:

$$D(\{\pi_c\}_{c=1}^k) = \sum_{c=1}^k \sum_{a_i \in \pi_c} \|a_i - m_c\|^2$$

که در آن $m_c = \frac{\sum_{a_i \in \pi_c} a_i}{|\pi_c|}$ نشان دهنده میانگین خوشه c ام است و c امین خوشه با π_c نمایش داده می شود. اشکال K-means استاندارد این است که خوشه ها را با یک ابرصفحه از هم جدا می کند، به دلیل اینکه از متر اقلیدسی برای بدست آوردن فاصله بین نقاط استفاده می کند، بنابراین الگوریتم K-means فقط می تواند داده هایی را که به طور خطی جداپذیر هستند، خوشه بندی کند. برای خوشه بندی داده هایی که به طور خطی جداپذیر نیستند

الگوریتم ۱-۳ K-means [۱]

- ورودی: خصوصیات n داده k تعداد دسته ها
خروجی: k دسته که داده های هر دسته از نظر شباهت به هم نزدیک و از دسته های دیگر دورند.
- ۱: داده را به عنوان مرکز خوشه انتخاب می کنیم.
 - ۲: مرحله سوم تا پنجم را تا رسیدن به عدم تغییر در خوشه ها تکرار می کنیم.
 - ۳: فواصل بقیه داده ها با مرکز خوشه ها را تعیین می کنیم.
 - ۴: داده هایی که به مرکز هر خوشه نزدیکترند در آن خوشه قرار می گیرند
 - ۵: میانگین هر خوشه را به عنوان مرکز جدید خوشه در نظر می گیریم.

می توان از روش K-means کرنلی استفاده کرد.

۲-۳ K-means کرنلی

K-means کرنلی از یک تابع برای نگاشت نقاط به یک فضای ویژگی با بعد بالاتر استفاده می کند. بعد از اینکه نقاط توسط تابع مورد نظر الگوریتم K-means کرنلی به فضای ویژگی (فضای با بعد بالاتر) نگاشت داده شدند، با اعمال شدن K-means در این فضای ویژگی، جداکننده های خطی در این فضای جدید با جداکننده های غیر خطی در فضای ورودی (اولیه) متناظرند.

تابع هدف K-means کرنلی بصورت مینیم سازی رابطه زیر نوشته می شود:

$$D(\{\pi_c\}_{c=1}^k) = \sum_{c=1}^k \sum_{a_i \in \pi_c} \|\phi(a_i) - m_c\|^2 \quad (1-3)$$

که در آن $m_c = \frac{\sum_{a_i \in \pi_c} \phi(a_i)}{|\pi_c|}$ نشان دهنده میانگین خوشه c است و c امین خوشه با π_c نمایش داده می شود.

با گذاشتن مقدار m_c در تابع هدف K-means کرنلی (۱-۳) داریم:

$$\begin{aligned} \|\phi(a_i) - m_c\|^2 &= \left\| \phi(a_i) - \frac{\sum_{a_i \in \pi_c} \phi(a_i)}{|\pi_c|} \right\|^2 & (2-3) \\ &= \phi(a_i) \cdot \phi(a_i) - \frac{2 \sum_{a_j \in \pi_c} \phi(a_i) \cdot \phi(a_j)}{|\pi_c|} + \frac{\sum_{a_j, a_l \in \pi_c} \phi(a_j) \cdot \phi(a_l)}{|\pi_c|^2} \end{aligned}$$

که در آن $|\pi_c|$ تعداد اعضای خوشه c ام را نشان می‌دهد. بنابراین از حاصلضرب داخلی بردارها برای محاسبه فاصله اقلیدسی یک نقطه تا مرکز خوشه استفاده می‌شود. لذا اگر ماتریس کرنل K را در نظر بگیریم طوری که

$$K_{ij} = \phi(a_i) \cdot \phi(a_j)$$

می‌توان فاصله بین نقاط تا مراکز خوشه‌ها را بدون دانستن ضابطه تابع ϕ محاسبه کرد. هر ماتریس نیمه معین مثبت K می‌تواند به عنوان یک ماتریس کرنل در نظر گرفته شود [۲۴]. تابع کرنل \mathcal{K} برای نگاشت نقاط اصلی به فضای ویژگی با استفاده از حاصلضرب داخلی بردارها استفاده می‌شود.

$$\mathcal{K}(a_i, a_j) = K_{i,j}$$

لذا رابطه ۲-۳ می‌تواند بصورت زیر نوشته شود:

$$\|\phi(a_i) - m_c\|^2 = k_{ij} - \frac{2 \sum_{a_j \in \pi_c} k_{ij}}{|\pi_c|} + \frac{\sum_{a_j, a_l \in \pi_c} k_{jl}}{|\pi_c|^2} \quad (3-3)$$

۳-۳ K-means کرنلی وزندار

در این روش، به هریک از نقاط مجموعه داده یک وزن اختصاص داده می‌شود، همانطور که در ادامه خواهید دید این وزن‌ها یک نقش حیاتی را در نشان دادن هم‌ارزی تابع هدف K-means کرنلی و تابع هدف‌های افرازگراف بازی می‌کنند.

تابع هدف K-means کرنلی وزندار به صورت زیر بیان می‌شود:

$$\min \mathcal{D}(\{\pi_c\}_{c=1}^k) = \min \sum_{c=1}^k \sum_{a_i \in \pi_c} w_i \|\phi(a_i) - m_c\|^2$$

که در آن $m_c = \frac{\sum_{a_i \in \pi_c} w_i \phi(a_i)}{\sum_{a_i \in \pi_c} w_i}$ و $w_i \geq 0$ وزن نقطه i ام است و بهترین مرکز خوشه به صورت زیر به دست می‌آید:

$$m_c = \operatorname{argmin}_z \sum_{a_i \in \pi_c} w_i \|\phi(a_i) - z\|^2$$

که در آن

$$z = \{m_{c_j}\}_{j=1}^t$$

بردار حاوی مراکز به دست آمده برای خوشه c ام در طی تکرارهای الگوریتم و t تعداد تکرارهای الگوریتم است. همانطور که قبلاً گفته شد، برای محاسبه فاصله نقاط تا مرکز خوشه فقط از حاصلضرب داخلی استفاده می‌کنیم. با جایگذاری m_c در $\|\phi(a_i) - m_c\|^2$ داریم:

$$\begin{aligned} \|\phi(a_i) - m_c\|^2 &= \left\| \phi(a_i) - \frac{\sum_{a_i \in \pi_c} w_i \phi(a_i)}{\sum_{a_i \in \pi_c} w_i} \right\|^2 & (4-3) \\ &= \phi(a_i) \cdot \phi(a_i) - \frac{2 \sum_{a_j \in \pi_c} w_j \phi(a_i) \cdot \phi(a_j)}{\sum_{a_j \in \pi_c} w_j} + \frac{\sum_{a_j, a_l \in \pi_c} w_j w_l \phi(a_j) \cdot \phi(a_l)}{(\sum_{a_j \in \pi_c} w_j)^2} \end{aligned}$$

با استفاده از ماتریس کرنل K ، تساوی فوق به صورت زیر نوشته می‌شود:

$$\|\phi(a_i) - m_c\|^2 = K_{ii} - \frac{2 \sum_{a_j \in \pi_c} w_j K_{ij}}{\sum_{a_j \in \pi_c} w_j} + \frac{\sum_{a_j, a_l \in \pi_c} w_j w_l K_{jl}}{(\sum_{a_j \in \pi_c} w_j)^2} \quad (5-3)$$

۴-۳ هم‌ارزی تابع هدفهای طیفی با K-means کرنلی وزن‌دار

در این بخش، تابع هدف K-means کرنلی و توابع هدف افزاز گراف را بصورت مسأله ماکزیم‌سازی اثر بیان می‌کنیم سپس نشان می‌دهیم که تابع هدف‌های این دو روش از لحاظ ریاضی کاملاً هم‌ارز هستند.

۱-۴-۳ K-means کرنلی وزن‌دار بصورت مسأله ماکزیم‌سازی اثر

تابع هدف K-means کرنلی وزن‌دار را در نظر گرفته و آن را بصورت یک مسأله ماکزیم‌سازی اثر بیان می‌کنیم. فرض کنید s_c مجموع وزن رئوس خوشه c باشد که به صورت $s_c = \sum_{a_i \in \pi_c} w_i$ نشان داده می‌شود. ماتریس Z از بعد $n \times k$ را به صورت زیر تعریف می‌کنیم:

الگوریتم ۳-۲ الگوریتم کرنل kmeans وزندار

ورودی: K : ماتریس کرنل k : تعداد خوشه‌ها w : ماتریس وزن نقاط t_{max} : ماکزیمم تعداد تکرارها
 $\{\pi_c^{(0)}\}_{c=1}^k$: خوشه‌های اولیه انتخابی

خروجی: $\{\pi_c\}_{c=1}^k$: افراز نهایی نقاط

۱: اگر در ابتدا هیچ خوشه اولیه‌ای وجود ندارد، k خوشه اولیه $\pi_1^{(0)}, \dots, \pi_k^{(0)}$ را در نظر بگیر (مثلاً به صورت تصادفی) و $t = 0$ قرار بده.

۲: به ازای هر خوشه c و هر $a_i \in \pi_c$ مرکز خوشه را بصورت زیر محاسبه کن:

$$m_c = \frac{\sum_{a_i \in \pi_c^{(t)}} w_i \phi(a_i)}{\sum_{a_i \in \pi_c^{(t)}} w_i}$$

۳: برای هر نقطه a_i و تمام مراکز خوشه‌ها

$$d(a_i, m_c) = K_{ii} - \frac{\sum_{a_j \in \pi_c^{(t)}} w_j K_{ij}}{\sum_{a_j \in \pi_c^{(t)}} w_j} + \frac{\sum_{a_j, a_l \in \pi_c^{(t)}} w_j w_l K_{jl}}{(\sum_{a_j \in \pi_c^{(t)}} w_j)^2} \quad \forall c = 1, \dots, k; i = 1, \dots, n$$

را محاسبه کن. (فاصله هر نقطه تا مراکز خوشه‌ها را بدست آور).

۴: $c^*(a_i) = \operatorname{argmin}_c d(a_i, m_c)$ را پیدا کن.

۵: خوشه‌ها را با استفاده از رابطه زیر آپدیت کن:

$$\pi_c^{t+1} = \{a_i : c^*(a_i) = c\}$$

۶: اگر $t = t + 1$ قرار بده و به مرحله ۲ برو، در غیر اینصورت پایان.

$$Z_{ic} = \begin{cases} \frac{1}{s_c} & \text{if } a_i \in \pi_c \\ 0 & \text{otherwise} \end{cases}$$

هر ستون ماتریس Z نشان دهنده اعضای یک خوشه مجزا است، به وضوح، ستون‌های Z بر هم عمودند.

فرض کنید Φ ماتریس حاوی بردارهای $\phi(a_i)$ باشد و M ماتریس قطری وزن رئوس است. ستون i ام ماتریس

$\Phi M Z Z^T$ برابر است با بردار میانگین خوشه c ام، یعنی

$$(\Phi M Z Z^T)_i = m_c$$

بنابراین، تابع هدف K-means کرنلی وزندار بصورت در می‌آید:

$$\begin{aligned} \mathcal{D}(\{\pi_c\}_{c=1}^k) &= \sum_{c=1}^k \sum_{a_i \in \pi_c} w_i \|\phi(a_i) - m_c\|^2 \\ &= \sum_{i=1}^n w_i \|\Phi_i - (\Phi M Z Z^T)_i\|^2 \end{aligned}$$

که در آن Φ_i ، i امین ستون ماتریس Φ است. فرض کنید $\tilde{Y} = M^{\frac{1}{2}} Z$ ، به وضوح \tilde{Y} متعامد نرمال است ($\tilde{Y}^T \tilde{Y} = I_k$)، زیرا M قطری است پس نرمال است یعنی $MM^T = M^T M$ و Z نیز همانطور که قبلاً گفته شد متعامد است.

بنابراین تابع هدف را می‌توان به صورت زیر بازنویسی کرد:

$$\begin{aligned} \mathcal{D}(\{\pi_c\}_{c=1}^k) &= \sum_{i=1}^n w_i \|\Phi_i - (\Phi M Z Z^T)_i\|^2 \\ &= \sum_{i=1}^n w_i \|\Phi_i - (\Phi M M^{-\frac{1}{2}} \tilde{Y} \tilde{Y}^T M^{-\frac{1}{2}})_i\|^2 \\ &= \sum_{i=1}^n w_i \|\Phi_i - (\Phi M^{\frac{1}{2}} \tilde{Y} \tilde{Y}^T M^{-\frac{1}{2}})_i\|^2 \\ &= \sum_{i=1}^n \|\Phi_i w_i^{\frac{1}{2}} - (\Phi M^{\frac{1}{2}} \tilde{Y} \tilde{Y}^T)_i\|^2 \\ &= \|\Phi M^{\frac{1}{2}} - \Phi M^{\frac{1}{2}} \tilde{Y} \tilde{Y}^T\|_F^2 \end{aligned}$$

از سه خاصیت زیر نرم فروبینیوس (۱-۱) برای ادامه روند اثبات استفاده می‌کنیم.

$$\|A\|_F^2 = \text{trace}(A^T A) = \text{trace}(A A^T) \quad .۱$$

$$\text{trace}(A + B) = \text{trace}(A) + \text{trace}(B) \quad .۲$$

$$\text{trace}(AB) = \text{trace}(BA) \quad .۳$$

$$\begin{aligned}
\|\Phi M^{\dagger} - \Phi M^{\dagger} \tilde{Y} \tilde{Y}^T\|_F^2 &= \text{trace}((\Phi M^{\dagger} - \Phi M^{\dagger} \tilde{Y} \tilde{Y}^T)^T (\Phi M^{\dagger} - \Phi M^{\dagger} \tilde{Y} \tilde{Y}^T)) \\
&= \text{trace}((M^{\dagger} \Phi^T - \tilde{Y} \tilde{Y}^T M^{\dagger} \Phi^T) (\Phi M^{\dagger} - \Phi M^{\dagger} \tilde{Y} \tilde{Y}^T)) \\
&= \text{trace}(M^{\dagger} \Phi^T \Phi M^{\dagger} - M^{\dagger} \Phi^T \Phi M^{\dagger} \tilde{Y} \tilde{Y}^T - \tilde{Y} \tilde{Y}^T M^{\dagger} \Phi^T \Phi M^{\dagger} \\
&\quad + \tilde{Y} \tilde{Y}^T M^{\dagger} \Phi^T \Phi M^{\dagger} \tilde{Y} \tilde{Y}^T) \\
&= \text{trace}(M^{\dagger} \Phi^T \Phi M^{\dagger}) - \text{trace}(\tilde{Y}^T M^{\dagger} \Phi^T \Phi M^{\dagger} \tilde{Y})
\end{aligned}$$

با توجه به اینکه $K = \Phi^T \Phi$ و اینکه $\text{trace}(M^{\dagger} K M^{\dagger})$ یک مقدار ثابت است، مینیمم تابع هدف K-means کرنلی وزندار، هم‌ارز است با:

$$\max \text{trace}(\tilde{Y}^T M^{\dagger} K M^{\dagger} \tilde{Y}) \quad (3-6)$$

که در آن \tilde{Y} یک ماتریس $n \times k$ متعامد نرمال است که متناسب با ریشه دوم ماتریس وزن M است همانطور که در قبل گفته شد.

۳-۵ هم‌ارزی تابع هدف های افراز‌گراف با K-means کرنلی وزندار

در این بخش هر تابع هدف مسأله افراز‌گراف را به صورت مسأله ماکزیم‌سازی اثر می‌نویسیم و سپس هم‌ارزی آنها را با K-means کرنلی وزندار نشان می‌دهیم.

۳-۵-۱ نسبت وابستگی

برای تبدیل کردن تابع هدف نسبت وابستگی به مسأله ماکزیم‌سازی اثر ابتدا بردار شاخص X_c را برای افراز π_c معرفی می‌کنیم:

$$X_c = \begin{cases} 1 & a_i \in \pi_c \quad \forall i = 1, \dots, n \\ 0 & otherwise \end{cases}$$

تابع هدف نسبت وابستگی (۷-۲) با استفاده از این بردار شاخص بصورت رابطه ۷-۳ در می آید:

$$\begin{aligned} & \max \sum_{c=1}^k \frac{W(\pi_c, \pi_c)}{|\pi_c|} \\ &= \max \sum_{c=1}^k \frac{X_c^T W X_c}{X_c^T X_c} \\ &= \max \sum_{c=1}^k \tilde{X}_c^T W \tilde{X}_c \end{aligned} \quad (۷-۳)$$

که در آن W ماتریس مجاورتی گراف است و $\tilde{X}_c = \frac{X_c}{(X_c^T X_c)^{\frac{1}{2}}}$ تساوی های فوق برقرارند زیرا $X_c^T X_c$ برابر با سایز خوشه c و $X_c^T W X_c$ برابر با مجموع وزن یال های درون خوشه c است. با توجه به اینکه ماتریس $\tilde{X}^T W \tilde{X}$ یک ماتریس قطری از بعد $k \times k$ است که درایه c ام قطر اصلی آن برابر است با $\tilde{X}_c^T W \tilde{X}_c$ و تابع هدف مسأله نسبت وابستگی که به صورت رابطه ۷-۳ بیان شد، برابر است با مجموع عناصر روی قطر اصلی ماتریس $\tilde{X}^T W \tilde{X}$ ، رابطه ۷-۳ را می توان به صورت زیر نوشت:

$$\max \text{trace}(\tilde{X}^T W \tilde{X}) \quad (۸-۳)$$

بدیهی است که $\tilde{X}^T \tilde{X} = I_k$. اکنون تابع هدف مسأله K-means کرنلی وزندار (۶-۳) را در نظر می گیریم. می دانیم که $\tilde{Y} = M^{\frac{1}{2}} Z$ پس مقدار $M^{\frac{1}{2}} Z$ را در تابع هدف، جایگزین \tilde{Y} می کنیم:

$$\begin{aligned} \tilde{Y}^T M^{\frac{1}{2}} K M^{\frac{1}{2}} \tilde{Y} &= Z^T M^{\frac{1}{2}} M^{\frac{1}{2}} K M^{\frac{1}{2}} M^{\frac{1}{2}} Z \\ &= Z^T M K M Z \end{aligned}$$

بنابراین تابع هدف K-means کرنلی وزندار (۶-۳) بصورت زیر در می آید:

$$\max \text{trace}(Z^T M K M Z)$$

طبق آنچه در قبل گفتیم Z یک بردار شاخص برای هر اعضای هر خوشه است لذا MZ را نیز می توان یک بردار شاخص در نظر گرفت و ماتریس W نیز به جای ماتریس کرنل استفاده شده است و وزن رئوس در نسبت

وابستگی، یک در نظر گرفته شده است.

بنابراین می توان نتیجه گرفت که تابع هدف مسأله نسبت وابستگی و مسأله K-means کرنلی وزن دار (۳-۶) باهم هم ارز هستند پس می توان K-means کرنلی وزن دار را روی ماتریس مجاورت برای بهینه سازی یکنواخت نسبت وابستگی در گراف اجرا کرد البته توجه به این نکته ضروری است که ماتریس مجاورت باید معین مثبت باشد. معین مثبت بودن شرط استفاده از ماتریس مجاورت به عنوان یک ماتریس کرنل است و می توان این ماتریس را به صورت $W = \Phi^T \Phi$ نوشت.

اگر ماتریس مجاورت معین مثبت نباشد هیچ تضمینی برای همگرایی تابع هدف K-means کرنلی نخواهیم داشت. البته معین مثبت بودن شرط کافی برای همگرایی است اما شرط لازم نیست.

۲-۵-۳ نسبت برش

تابع هدف مسأله نسبت برش (۲-۸) را در نظر می گیریم.

$$\min \sum_{c=1}^k \frac{W(\pi_c, v \setminus \pi_c)}{|\pi_c|}$$

با استفاده از بردار شاخص X_c تعریف شده در قبل تابع هدف فوق را می توان به صورت زیر نوشت:

$$\min \sum_{c=1}^k \frac{X_c^T (D - W) X_c}{X_c^T X_c}$$

یا استفاده از تغییر متغیر $\tilde{X}_c = \frac{X_c}{(X_c^T X_c)^{\frac{1}{2}}}$ تابع هدف نسبت برش به صورت زیر در می آید:

$$\min \sum_{c=1}^k \tilde{X}_c^T L \tilde{X}_c$$

چون $\tilde{X}_c^T L \tilde{X}_c$ درایه c ام ماتریس $\tilde{X}^T L \tilde{X}$ است و ماتریس $\tilde{X}^T L \tilde{X}$ قطری است، مسأله مینیمم سازی برش را به صورت زیر نیز می توان نوشت:

$$\min \text{trace}(\tilde{X}^T L \tilde{X}) \quad (۳-۹)$$

این مسأله مینیمم سازی را می توان با در نظر گرفتن $I - L$ به صورت زیر به یک مسأله ماکزیمم سازی تبدیل کرد:

$$\text{trace}(\tilde{X}^T(I - L)\tilde{X}) = \text{trace}(\tilde{X}^T\tilde{X}) - \text{trace}(\tilde{X}^TL\tilde{X})$$

چون \tilde{X} نرمال متعامد است، $\text{trace}(\tilde{X}^T\tilde{X}) = k$ ، بنابراین ماکزیمم سازی $\text{trace}(\tilde{X}^T(I - L)\tilde{X})$ با مینیمم سازی $\text{trace}(\tilde{X}^TL\tilde{X})$ هم ارز است. با قرار دادن همه اینها در کنار هم، به یک مسأله ماکزیمم سازی اثر هم ارز برای نسبت برش می رسیم.

نتیجه: مینیمم سازی نسبت برش برای ماتریس مجاورت با ماکزیمم سازی

$$\text{trace}(\tilde{X}^T(I - L)\tilde{X})$$

هم ارز است. بنابراین ماکزیمم سازی مسأله K-means کرنلی بدون وزن با ماکزیمم سازی

$$\text{trace}(\tilde{X}^T(I - L)\tilde{X})$$

هم ارز است که $(I - L)$ به جای ماتریس کرنل در نظر گرفته می شود.

۳-۵-۳ برش نرمال

تابع هدف مسأله برش نرمال (۵-۲) را که به صورت معادل زیر نیز نوشته می شود، در نظر می گیریم:

$$\min \sum_{c=1}^k \frac{W(\pi_c, v \setminus \pi_c)}{\text{vol}(\pi_c)}$$

که در آن $W(\pi_c, v \setminus \pi_c) = \sum_{(i \in \pi_c, j \in v \setminus \pi_c)} W_{ij}$ و $\text{vol}(\pi_c) = \sum_{j \in \pi_c} D_j$ و با توجه به اینکه

$$W(\pi_c, v \setminus \pi_c) = \text{vol}(\pi_c) - W(\pi_c, \pi_c)$$

پس مینیمم سازی $W(\pi_c, v \setminus \pi_c)$ برابر است با ماکزیمم سازی $W(\pi_c, \pi_c)$ ، لذا تابع برش نرمال را می توان

به صورت زیر نوشت:

$$\max \left\{ \sum_{c=1}^k \frac{W(\pi_c, \pi_c)}{\text{vol}(\pi_c)} = \sum_{c=1}^k \frac{X_c^T W X_c}{X_c^T D X_c} = \sum_{c=1}^k \tilde{X}_c W \tilde{X}_c \right\} = \max \text{trace}(\tilde{X} W \tilde{X})$$

که در آن $\tilde{X}_c = \frac{X_c}{(X_c^T D X_c)^{\frac{1}{2}}}$ فرض می‌کنیم $\tilde{Y} = D^{-\frac{1}{2}} \tilde{X}$ که یک ماتریس نرمال متعامد است. لذا مسأله برش نرمال را می‌توان بصورت

$$\max_Y \text{trace}(\tilde{Y}^T D^{-\frac{1}{2}} W D^{-\frac{1}{2}} \tilde{Y})$$

نوشت.

بنابراین اگر در تابع هدف K-means کرنلی وزن دار $(3-6)$ $M = D$ و $K = D^{-1} W D^{-1}$ قرار دهیم،

داریم:

$$\begin{aligned} \tilde{Y}^T M^{\frac{1}{2}} K M^{\frac{1}{2}} \tilde{Y} &= \tilde{Y}^T D^{\frac{1}{2}} D^{-1} W D^{-1} D^{\frac{1}{2}} \tilde{Y} \\ &= \tilde{Y}^T D^{-\frac{1}{2}} W D^{-\frac{1}{2}} \tilde{Y} \end{aligned}$$

به این نتیجه می‌رسیم که ماکزیمم‌سازی اثر تابع هدف K-means کرنلی وزن دار $(3-6)$ با مینیمم‌سازی تابع هدف برش نرمال، هم‌ارز است.

۳-۵-۴ نسبت وابستگی وزن دار

مسأله نسبت وابستگی وزن دار را نیز می‌توان به صورت مسأله ماکزیمم‌سازی اثر بیان کرد. برای این منظور تابع هدف مسأله را در نظر می‌گیریم:

$$\max \left\{ \sum_{c=1}^k \frac{W(\pi_c, \pi_c)}{w(\pi_c)} = \sum_{c=1}^k \frac{X_c^T W X_c}{X_c^T M X_c} = \sum_{c=1}^k \tilde{X}_c^T W \tilde{X}_c \right\} = \max \text{trace}(\tilde{X}^T W \tilde{X})$$

که در آن $\tilde{X}_c = X_c(X_c^T M X_c)^{-\frac{1}{2}}$ ، فرض می‌کنیم $Y = M^{-\frac{1}{2}} \tilde{X}$ بنابراین داریم:

$$W Assoc(G) = \max_Y \text{trace}(Y^T M^{-\frac{1}{2}} W M^{-\frac{1}{2}} Y) \quad (۱۰-۳)$$

لذا اگر ماتریس W را ماتریس کرنل در نظر بگیریم به این نتیجه می‌رسیم که مسأله نسبت وابستگی وزن دار با Krenel Kmeans (۶-۳) وزن دار هم‌ارز است.

۵-۵-۳ نسبت برش وزن دار

با استفاده از مفهوم مشابه برای M و X_c تابع هدف مسأله نسبت برش را به صورت زیر می‌نویسیم:

$$\begin{aligned} Wcut(G) &= \min \sum_{c=1}^k \frac{W(\pi_c, v \setminus \pi_c)}{w(\pi_c)} \\ &= \min \sum_{c=1}^k \frac{X_c^T (D - W) X_c}{X_c^T M X_c} \\ &= \min \sum_{c=1}^k \tilde{X}_c^T L \tilde{X}_c \\ &= \min \text{trace}(\tilde{X}^T L \tilde{X}) \\ &= \min_Y \text{trace}(Y^T M^{-\frac{1}{2}} L M^{-\frac{1}{2}} Y) \end{aligned}$$

با این فرض که $Y = M^{-\frac{1}{2}} \tilde{X}$.

مسأله $Wcut$ می‌تواند به صورت $Assoc$ بیان شود زیرا:

$$\begin{aligned} \text{trace}(Y^T M^{-\frac{1}{2}} (M - L) M^{-\frac{1}{2}} Y) &= \text{trace}(Y^T M^{-\frac{1}{2}} M M^{-\frac{1}{2}} Y - Y^T M^{-\frac{1}{2}} L M^{-\frac{1}{2}} Y) \\ &= \text{trace}(Y^T Y - Y^T M^{-\frac{1}{2}} L M^{-\frac{1}{2}} Y) \\ &= k - \text{trace}(Y^T M^{-\frac{1}{2}} L M^{-\frac{1}{2}} Y) \end{aligned}$$

بنابراین مینیمم‌سازی $Wcut$ با ماکزیمم‌سازی $\text{trace}(Y^T M^{-\frac{1}{2}} (M - L) M^{-\frac{1}{2}} Y)$ هم‌ارز است. از اینرو مینیمم‌سازی $Wcut$ روی W هم‌ارز با ماکزیمم‌سازی $Assoc$ روی ماتریس مجاورتی $M - L$ است.

می‌دانیم که

$$Y = M^{\frac{1}{2}} \tilde{X} = \frac{M^{\frac{1}{2}} X}{(X^T M X)^{\frac{1}{2}}}$$

باتوجه به قبل می‌دانیم که $\tilde{Y} = M^{\frac{1}{2}} Z$ یک ماتریس نرمال متعامد است و

$$\tilde{Y}_{ic} = \begin{cases} \frac{w_i^{\frac{1}{2}}}{s_c^{\frac{1}{2}}} & a_i \in \pi_c \\ 0 & otherwise \end{cases}$$

که s_c را می‌توان بصورت زیر نوشت:

$$s_c = X_c^T M X_c$$

لذا داریم:

$$\tilde{Y} = \frac{M^{\frac{1}{2}} X}{(X^T M X)^{\frac{1}{2}}}$$

لذا ماکزیمم‌سازی اثر برای K-means کرنلی وزن‌دار ۳-۶ برابر است با

$$\max_Y \text{trace}(\tilde{Y}^T M^{-\frac{1}{2}} W M^{-\frac{1}{2}} \tilde{Y})$$

است با این شرط که $W = M K M$ که در آن K ماتریس کرنل و M ماتریس وزن نقاط است. بنابراین تابع هدف K-means کرنلی وزن‌دار (۳-۶) با تابع هدف نسبت وابستگی وزن‌دار (۳-۹)، هم‌ارز است و همانطور که در قبل گفتیم ماکزیمم‌سازی نسبت وابستگی وزن‌دار با مینیمم‌سازی نسبت برش وزن‌دار، هم‌ارز است، پس تابع هدف مسأله نسبت برش وزن‌دار نیز با تابع هدف K-means کرنلی وزن‌دار، هم‌ارز است.

فصل ۴

دست‌آوردهای پژوهش

یکی از مشکلات خوشه‌بندی طیفی، تنظیم پارامترهای مورد نیاز برای ایجاد ماتریس مجاورتی گراف است. در این فصل یک شیوه مؤثر برای بدست آوردن این پارامترها در قطعه‌بندی تصاویر ام آر آی مبتنی بر الگوریتم ژنتیک ارائه شده است.^۱ همچنین با توجه به اندازه بزرگ ماتریس‌های مجاورتی تصاویر ام آر آی از روش نایستروم برای تخمین بردارها و مقادیر ویژه مورد نیاز در خوشه‌بندی طیفی استفاده شده است. نتایج آزمایشات انجام شده بر روی چند تصویر کارایی شیوه پیشنهادی در تخمین پارامترهای خوشه‌بندی با استفاده از الگوریتم ژنتیک بر روی یک تصویر و تعمیم آن بر روی تصاویر مشابه را نشان داده است. قطعه‌بندی تصویر یکی از معروف‌ترین مسائل در حوزه پردازش تصویر و بینایی ماشین است که کاربردهای بسیار متعددی دارد و تحقیقات مختلفی را به خود معطوف نموده است [۲۶]. کاربرد موردنظر در این روش پیشنهادی، شناسایی ناحیه چربی در تصاویر ام آر آی ران ورزشکاران قبل و بعد از انجام یک دوره تمرینات ورزشی و مشخص شدن تاثیر نوع تمرین بر کاهش یا افزایش بافت چربی داوطلبان بوده است. در این فصل شیوه خودکاری برای تخمین پارامترهای موردنیاز در ایجاد گراف خوشه‌بندی تصویر بیان شده و کارایی آن در جداسازی ناحیه چربی تصاویر ام آر آی نشان داده شده است. حتی برای اندازه‌های کوچک تصاویر، ماتریس مجاورتی بسیار بزرگ شده و با کمبود حافظه در عمل مواجه می‌شویم. در اینجا از روش نایستروم که در [۲۷] برای مواجهه با این مشکل پیشنهاد شده است، استفاده شده است.

در این فصل یک تکنیک تقریبی قابل اجرا برای روش‌های گروه‌بندی طیفی ارائه می‌دهیم که بار محاسباتی را تا حد زیادی کاهش می‌دهد. رویکرد مورد نظر ما روش کلاسیک حل مساله مقدار ویژه شناخته شده تحت عنوان روش نایستروم است. به طور خلاصه، تقریب با اولین جواب مساله برای یک زیرمجموعه تصادفی کوچک از پیکسل‌ها کار میکند و سپس این جواب با مجموعه تمام پیکسل‌های تصویر مقایسه می‌شود. روش طیفی ای که در این جا روی آن تمرکز می‌کنیم روش برش نرمال است.

^۱ مطالب این فصل در قالب مقاله [۲۵] در کنفرانس تحقیق در عملیات ارائه شده است

یکی از مشکلات عملی خوشه‌بندی طیفی در قطعه‌بندی تصویر، اندازه بسیار بزرگ ماتریس مجاورتی W است. به عنوان مثال یک تصویر با ابعاد 1000×1000 دارای 10^6 پیکسل بوده و ماتریس مجاورتی این پیکسل‌ها به عنوان رئوس گراف یک ماتریس $10^6 \times 10^6$ خواهد بود که به صورت معمول در حافظه نرم‌افزارهایی همچون MATLAB جا نمی‌گیرد. برای محاسبه بردارها و مقادیر ویژه ماتریس لاپلاسیان در چنین حالاتی از روش نایستروم [۲۸] استفاده میشود.

۱-۴ روش نایستروم

روش نایستروم^۱ بردارها و مقادیر ویژه یک ماتریس بزرگ را با استفاده از یک زیر ماتریس کوچک A تخمین می‌زند [۲۸]. اگر رتبه ماتریس A ، r باشد، تخمین بدست آمده دقیق است.

فرض کنید \tilde{K}_{mm} یک ماتریس کرنل با رتبه $r \ll m$ باشد. چون \tilde{K} نیمه معین مثبت است پس گرام است و می‌تواند بصورت $\tilde{K} = ZZ^T$ نوشته شود که $Z \in M_{mr}$ و از رتبه r است. سطرها را در Z طوری مرتب می‌کنیم که r تای اول مستقل خطی باشند، این کار باعث مرتب شدن سطرها و ستونها در \tilde{K} و ایجاد یک ماتریس جدید K که متقارن است می‌شود در نتیجه زیر ماتریس اصلی $A \in S_r$ از K (که خودش ماتریس گرام r سطر اول Z است) دارای رتبه کامل است. اکنون فرض می‌کنیم $n \equiv m - r$ ، ماتریس K را بصورت زیر می‌نویسیم:

$$\begin{bmatrix} A_{rr} & B_{rn} \\ B_{nr}^T & C_{nn} \end{bmatrix}$$

چون ماتریس A دارای رتبه کامل است r سطر $[A_{rr} \ B_{rn}]$ مستقل خطی اند و از آنجایی که رتبه r ، K است، n سطر $[B_{nr}^T \ C_{nn}]$ بصورت

$$[B_{nr}^T \ C_{nn}] = H_{nr} [A_{rr} \ B_{rn}]$$

تجزیه می‌شود. با در نظر گرفتن r ستون اول بصورت $H = B^T A^{-1}$ و n ستون بعدی $C = B^T A^{-1} B$ داریم:

$$K_{mm} = \begin{bmatrix} A & B \\ B^T & B^T A^{-1} B \end{bmatrix} = \begin{bmatrix} A \\ B^T \end{bmatrix} A_{rr}^{-1} [A \ B] \quad (۱-۴)$$

^۱Nyström

با استفاده از تجزیه ویژه A تجزیه ویژه تقریبی K_{mm} را (برای مقادیر ویژه ناصفر) می‌سازیم. $A = U\Lambda U^T$ تجزیه ویژه A است که در آن U ماتریس ستونی بردارهای ویژه A و Λ ماتریس قطری مقادیر ویژه A است. بنابراین فرمول ۴-۱ را بصورت زیر بازنویسی می‌کنیم:

$$K_{mm} = \begin{bmatrix} U \\ B^T U \Lambda^{-1} \end{bmatrix} \Lambda [U^T \quad \Lambda^{-1} U^T B] \equiv D \Lambda D^T \quad (۲-۴)$$

رابطه ۴-۲ یک تجزیه تقریبی K را نشان می‌دهد، D ماتریس ستونی بردارهای ویژه K و Λ ماتریس قطری مقادیر ویژه K است. اگر ستونهای ماتریس D متعامد باشند، تجزیه ویژه دقیق K بدست می‌آید. اگر به جای A

$$Q_{rr} = A + A^{-1/2} B B^T A^{-1/2} \equiv U_Q \Lambda_Q U_Q^T$$

را تجزیه کنیم، ماتریس بردارهای ویژه ستونی متعامد K بصورت زیر بدست می‌آید:

$$V_{mr} = \begin{bmatrix} A \\ B^T \end{bmatrix} A^{-1/2} U_Q \Lambda_Q^{-1/2} \quad (۳-۴)$$

بنابراین $K_{mm} = V \Lambda_Q V^T$ تجزیه دقیق K است و $V V^T = I_{rr}$. در پیاده‌سازی انجام شده از این شیوه برای تخمین بردارها و مقادیر ویژه ماتریس لاپلاسیان استفاده شده است.

۲-۴ روش پیشنهادی

همانطور که قبلاً گفته شد ماتریس مجاورتی گراف بر طبق رابطه ۲-۲۷ و بر اساس اختلاف شدت روشنایی و اختلاف موقعیت مکانی پیکسل‌ها بدست می‌آید، اما عنوان نشد که پارامترهای σ آن چگونه برآورد می‌شوند. در بخش ۴ مرجع [۱۷] ذکر شده است که عموماً مقدار σ بین ۱۰ تا ۲۰ درصد کل برد تابع فاصله در نظر گرفته می‌شود. بازه مشخص شده بازه مناسبی است و با سعی و خطا می‌توان مقداری که خروجی مناسبی تولید نماید را بدست آورد. در روش پیشنهادی به جای استفاده از شیوه سعی و خطا از یک الگوریتم ژنتیک برای تخمین این پارامترها استفاده شده است. اما به دلیل اینکه قطعه‌بندی تصویر با روش خوشه‌بندی طیفی زمان نسبتاً زیادی لازم

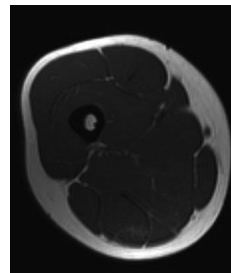
دارد در اینجا به جای تخمین پارامترها روی تصاویر اصلی پارامترها روی نمونه‌های کوچک‌تری از تصاویر تخمین زده می‌شوند و سپس برای تصویر اولیه اصلاح می‌شوند. با توجه به اینکه طیف شدت روشنایی پیکسلها در ابعاد مختلف تصویر تغییری ندارد، تخمین σ_I نیز اصلاح نمی‌شود. اما از آنجا که با مثلا دو برابر شدن اندازه تصویر، فواصل نقاط مشابه نیز دو برابر می‌شود، کفایت پارامتر σ_X در مقیاس تغییر اندازه تصویر ضرب شود تا مقدار مناسب تصویر اولیه حاصل شود.

عمل تخمین پارامترها با استفاده از یک الگوریتم ژنتیک با دو متغیر ورودی انجام شده است. یک نکته مهم در الگوریتم ژنتیک تعریف تابع برازندگی (تابع هدف) مناسب مسئله می‌باشد. در اینجا برای یک تصویر از مجموعه تصاویر مورد نظر عمل قطعه‌بندی به صورت دستی انجام پذیرفته است و در تابع تطابق الگوریتم ژنتیک، خروجی خوشه‌بندی هر جواب ممکن با این جواب از پیش مشخص مقایسه می‌شود. اختلاف کمتر به منزله برازندگی بیشتر کروموزوم خواهد بود. در چنین کاربردهایی که قطعه‌بندی تصاویر متعددی از یک حوزه خاص مدنظر باشد چنین روشی معقولانه است.

پارامترهای بدست آمده از الگوریتم ژنتیک در کاربرد خاص این مقاله به راحتی قابل تعمیم به سایر تصاویر می‌باشند. شکل ۴-۱ تصویر ورودی به الگوریتم ژنتیک و نتیجه قطعه‌بندی آن بر اساس پارامترهای بدست آمده از الگوریتم ژنتیک و با یک قطعه‌بندی از قبل مشخص را نشان می‌دهد. اعمال پارامترهای بدست آمده بر روی سایر تصاویر ام آر آی خروجی رضایت‌بخشی را تولید نموده است. ۴-۲ دو نمونه تصویر دیگر و خروجی قطعه‌بندی آنها با شیوه خوشه‌بندی طیفی و بر اساس پارامترهای بدست آمده از اجرای الگوریتم ژنتیک روی تصویر شکل ۴-۱ را نشان می‌دهد. همان‌گونه که مشاهده می‌شود، ناحیه چربی به صورت مناسبی استخراج شده است.

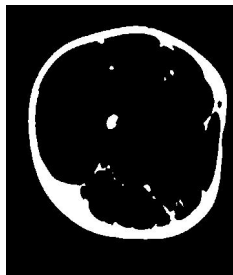


(ب) خروجی قطعه‌بندی شده

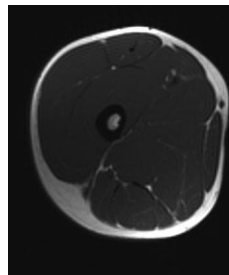


(آ) تصویر اصلی

شکل ۴-۱: یک نمونه تصویر ام آر آی و خروجی قطعه‌بندی شده آن با استفاده از خوشه‌بندی طیفی بر اساس پارامترهای تخمین زده شده توسط الگوریتم ژنتیک



(ب) خروجی قطعه بندی شده



(آ) تصویر ام آر آی

شکل ۴-۲: یک تصویر دیگر از مجموعه تصاویر ام آر آی و خروجی قطعه بندی شده آن که ماتریس مجاورتی خوشه بندی طیفی آنها بر اساس پارامترهای بدست آمده از اجرای الگوریتم ژنتیک بر روی شکل ۴-۱ به دست آمده است.

فهرست منابع

- [1] Alpaydin, Ethem. *Introduction to Machine Learning*. The MIT Press, 2nd ed. , 2010.
- [۲] گریمالدی، رالف پی. ریاضیات گسسته و ترکیباتی. ترجمه‌ی زاده، علی اکبر عالم. موسسه نشر علوم نوین، ویرایش چهارم، ۱۳۸۰.
- [3] Golub, G.H and Van Loan, C.F. *Matrix Computations*. Johns Hopkins University Press, Baltimore, 2nd ed. , 1989.
- [۴] کاردل، مرتضی. رفاهی شیخانی، امیرحسین. روش مستقل از معکوس وزن دار برای حل معادله مقدار ویژه تعمیم یافته. مجله ریاضی کاربردی واحد لاهیجان، ۲۰۱۰.
- [5] M.Apostol, Tom. *Multi Variable Calculus and Linear Algebra with Applications to Differential Equations and Probability*, vol. 2. Indiana University, 2nd ed. , 1968.
- [6] Mohar, Bojan. The laplacian spectrum of graphs. in *Graph Theory, Combinatorics, and Applications*, pp. 871–898. Wiley, 1991.
- [7] Chung, F.R.K. *Spectral Graph Theory*. American Mathematical Society, 1997.
- [8] Luxburg, Ulrike. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, December 2007.
- [9] Kabelíková, Pavla. Graph partitioning using spectral methods. tech. rep., VŠB - Technical University of Ostrava Faculty of Electrical Engineering and Computer Science Department of Computer Science, 2006.
- [10] J.Zaki, Mohammed and Meira Jr, Wagner. *Data Mining and Analysis: Fundamental Concepts and Algorithms*. Cambridge University Press, New York, NY, USA, 2014.
- [11] Zelnik-manor, Lihi and Perona, Pietro. Self-tuning spectral clustering. in *Advances in Neural Information Processing Systems 17*, pp. 1601–1608. MIT Press, 2004.
- [12] Nascimento, M. C. V. and de Carvalho, A. C. P. L. F. Spectral methods for graph clustering - a survey. *European Journal of Operational Research* 211(2), pp. 221–231, 2011.

- [13] Stoer, Mechthild and Wagner, Frank. A simple min-cut algorithm. *J. ACM*, 44(4):585–591, July 1997.
- [14] Wu, Z. and Leahy, R. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(11):1101–1113, November 1993.
- [15] Leighton, T. and Rao, S. An approximate max-flow min-cut theorem for uniform multicommodity flow problems with applications to approximation algorithms. *Proceedings of the 29th Annual Symposium on Foundations of Computer Science, IEEE Computer Society*, pp. 422–431, 1988.
- [16] Chan, K. P., and Schlag, M. Spectral k-way ratio-cut partitioning – part i: Preliminary results. tech. rep., University of California at Santa Cruz, 1992.
- [17] Shi, Jianbo and Malik, Jitendra. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(8):888–905, August 2000.
- [18] Ding, Q., C. H., He, X., and et al. A min-max cut algorithm for graph partitioning and data clustering. *Proceedings of the 2001 IEEE International Conference on Data Mining, IEEE Computer Society*, pp. 107–114, 2001.
- [19] M.Meila. and D.Verma. A comparison of spectral clustering algorithms. *University of Washington Department of Computer Science*, 2001.
- [20] Camastra, F., M.Filippone, and et al. A survey of kernel and spectral methods for clustering. *Pattern Recognition*41(1), pp. 176–190, 2008.
- [21] B.Schölkopf, A.Smola, and K-R.Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation*10(5), 1996.
- [22] Dhillon, Inderjit, Guan, Yuqiang, and Kulis, Brian. A unified view of kernel k-means, spectral clustering and graph cuts. tech. rep., University of Texas at Austin Department of Computer Sciences Austin, 2005.
- [23] J.MacQueen. Some methods for classification and analysis of multivariate observations. in *Proceedings of the Fifth Berkeley Symposium on Math., Stat. and Prob*, p. 281–296, 1967.
- [24] Cristianini.Nello and j.Shawe Taylor. *An Introduction to Support Vector Machines: And Other Kernel-based Learning Methods*. Cambridge University Press, New York, NY, USA, 2000.

[۲۵] امین طوسی، محمود و فیاض، طیبه. محاسبه پارامترهای خوشه بندی طیفی در تصاویر ام آر آی با الگوریتم ژنتیک. در هشتمین کنفرانس بین المللی انجمن ایرانی تحقیق در عملیات، صفحات ۷۱-۷۳، مشهد، ۱۳۹۴.

[26] Govindan, V. K. and Camilus, K. Santle. A review on graph based segmentation. *International Journal of Image, Graphics and Signal Processing (IJIGSP)*, 5:1–13, 2012.

[27] Fowlkes, Charless, Belongie, Serge, Chung, Fan, and Malik, Jitendra. Spectral grouping using the Nyström method. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(2):214–225, January 2004.

[28] Burges, Christopher J. C. Dimension reduction: A guided tour. *Foundations and Trends in Machine Learning*, 2(4), 2010.

پیوست آ

برنامه های نوشته شده

در این قسمت بخش اصلی برنامه های نوشته شده برای پایان نامه ذکر شده اند.

برنامه آ-۱: کد خوشه بندی داده ها با استفاده از الگوریتم *kmeans* کرنلی

```
۱ % Modified by M.Amintoosi
۲ % http://www.dcs.gla.ac.uk/~srogers/firstcourseml/matlab/
   chapter6/kernelkmeans.html#5
۳ clear all;
۴ close all;
۵
۶ produceNLData
۷ figure(1);hold off
۸ plot(X(:,1),X(:,2),'ko');
۹
۱۰ %% Compute the kernel
۱۱
۱۲ N = size(X,1);
۱۳ Ke = zeros(N);
۱۴ gam = 1;
۱۵ for n = 1:N
۱۶     for n2 = 1:N
۱۷         Ke(n,n2) = exp(-gam*sum((X(n,:)-X(n2,:)).^2)); %
           Gaussian
۱۸ %         Ke(n,n2) = X(n,:)*X(n2,:)' ; % Linear
۱۹ %         Ke(n,n2) = sum(x(n,:).^2)-sum(x(n2,:).^2); %
           Point Distances
۲۰     end
۲۱ end
۲۲
۲۳ %% Run Kernel K-means
```

```

24
25 converged = 0;
26 % Assign all objects into one cluster except one
27 % Kernel K-means is *very* sensitive to initial conditions.
    Try altering
28 % this initialisation to see the effect.
29 K = 2;
30 Z = repmat([1 0],N,1);
31 s = sum(X.^2,2);
32 pos = find(s==min(s));
33 Z(pos,:) = [0 1];
34 di = zeros(N,K);
35 cols = {'r','b'};
36
37 %% Plot the assignments
38
39 figure(1);hold off
40 for k = 1:K
41     pos = find(Z(:,k));
42     plot(X(pos,1),X(pos,2),'ko','markerfacecolor',cols{k});
43     hold on
44 end
45
46 %%
47 while ~converged
48
49     Nk = sum(Z,1);
50     for k = 1:K
51         % Compute kernelised distance
52         di(:,k) = diag(Ke) - (2/(Nk(k)))*sum(repmat(Z(:,k)
53             ',N,1).*Ke,2) + ...
54             Nk(k)^(-2)*sum(sum((Z(:,k)*Z(:,k)')).*Ke));
55     end
56     oldZ = Z;
57     Z = (di == repmat(min(di,[],2),1,K));
58     Z = 1.0*Z;
59     if sum(sum(oldZ~=Z))==0
60         converged = 1;
61     end
62
63 % Plot the assignments
64
65 figure(1);hold off
66 for k = 1:K
67     pos = find(Z(:,k));
68     plot(X(pos,1),X(pos,2),'ko','markerfacecolor',cols{k}

```

```

        k});
۶۸     hold on
۶۹     end
۷۰     pause (4);
۷۱ end

```

برنامه آ-۲: کد محاسبه بردارهای ویژه یک ماتریس بزرگ به روش نایستروم

```

۱ function labels = nystrom_driver(data, sigma, nsamp, k, n, t)
۲ % This function calculates the nystromized approximation to
۳ % the
۴ % Normalized cut partitioning algorithm.
۵ % data is the matrix with the data, where each row of the
۶ % matrix
۷ % contains a data point
۸ %
۹ % sigma is the standard deviation for the gaussian
۱۰ %
۱۱ % nsamp is the number samples that should whie performing
۱۲ % the
۱۳ % approximation
۱۴ % k is the number of desired groups
۱۵ %
۱۶ % n is the maximum number of iterations you would like the
۱۷ % kmeans
۱۸ % algorithm to run on top of the eigenvectors
۱۹ %
۲۰ % t is the threshold at which you would like the kmeans
۲۱ % algorithm
۲۲ % to terminate
۲۳ %
۲۴ % this function assumes the availability of a function
۲۵ % "your_own_distance"
۲۶ %
۲۷ % which will take two matrices A and B, with row vectors
۲۸ % corresponding to datapoints and calculate the all pairs
۲۹ % distance
۳۰ % between them.
۳۱ %
۳۲ % an example is included in this directory, its called "
۳۳ % pairdist"
۳۴ %
۳۵ % labels is an array of cluster labels for each point.
۳۶ %
۳۷ %
۳۸ %
۳۹ %
۴۰ %
۴۱ [N, Ncol]=size(data); % the dimensions of the dataset

```

```

۳۲             % each row is a new data point
۳۳
۳۴ ind_all=randperm(N);
۳۵ ind_samp=ind_all(1:nsamp);
۳۶ ind_rest=ind_all(nsamp+1:end);
۳۷ %samp_pos=[x(ind_samp)' y(ind_samp)'];
۳۸
۳۹
۴۰
۴۱ %d2A=your_own_distance(FB(ind_samp,:),FB(ind_samp,:)); %
    use your own distance function to get the pairwise
    distances
۴۲ %d2B=your_own_distance(FB(ind_samp,:),FB(ind_rest,:));
۴۳
۴۴ d2A=pairdist(data(ind_samp,:),data(ind_samp,:),'L2'); % use
    your own distance function to get the pairwise
    distances
۴۵ d2B=pairdist(data(ind_samp,:),data(ind_rest,:),'L2');
۴۶
۴۷ A=exp(-d2A/(2*sigma^2)); % choose your sigma
۴۸ B=exp(-d2B/(2*sigma^2));
۴۹ %imagesc(A);
۵۰ % the joys of nystrom
۵۱
۵۲ [V,ss]=nystrom_ncut(A,B);
۵۳
۵۴ % depermute the nystrom result
۵۵ V(ind_all,:)=V;
۵۶ V1=V(:,1:k); % choose the k eigenvectors
۵۷ Vnormalized=V1./repmat(sqrt(sum(V1.^2,2)),1,k);
۵۸ [centers,labels,error]=km(Vnormalized',k,n,t);

```

برنامه آ-۳: کد خوشه‌بندی داده‌ها به روش برش نرمال

```

۱ function [V1,D1]=ncut(W);
۲ % [V1,D1]=ncut(W,nv);
۳ %
۴ % solve generalized eigenproblem Wy={\mu}My
۵ %
۶ % optional third argument specifies minimum allowed weight
    (e.g. 0.01)
۷
۸
۹ m=sum(W,1);
۱۰ N=length(m);
۱۱ M=sparse(1:N,1:N,m);

```

```

12 B=inv(sqrt(M));
13
14 % solve generalized eigensystem
15 OPTIONS.tol=1e-4;
16 OPTIONS.maxit=20;
17 %OPTIONS.disp=0;
18 C=B*W*B;
19 format long
20 [V1,D1]=eig(C);
21
22 D1=1-D1;
23 V1=B*V1;
24
25 % sort the eigenvalues
26 [dvalues,dindex]=sort(diag(D1));
27 %keyboard;
28 V1=V1(:,dindex);
29 D1=diag(dvalues);
30 format

```


واژه‌نامه فارسی به انگلیسی

Genetic Algorithm	الگوریتم ژنتیک
Graph cut	برش گراف
Algebraic multiplicity	تکرر جبری
Kernel Clustering	خوشه‌بندی کرنلی
Spectral Clustering	خوشه‌بندی طیفی
Similarity Graph	گراف شباهت
Image Segmentation	قطعه بندی تصویر
Kernel Matrix	ماتریس کرنل
Laplacian Matrix	ماتریس لاپلاسیان
Minimum Cut Problem	مسأله برش کمینه
ratio cut problem	مسأله نسبت برش
Normaliezed Cut Problem	مسأله برش نرمال شده
Ratio Association Problem	مسأله نسبت وابستگی
connected	همبند

واژه‌نامه انگلیسی به فارسی

Algebraic multiplicity	تکرر جبری
Connected	همبند
Graph cut	برش گراف
Genetic Algorithm	الگوریتم ژنتیک
Image Segmentation	قطعه بندی تصویر
Kernel Matrix	ماتریس کرنل
Kernel Clustering	خوشه بندی کرنلی
Laplacian Matrix	ماتریس لاپلاسیان
Minimum Cut Problem	مسأله برش کمینه
Normaliezed Cut Problem	مسأله برش نرمال شده
Ratio Association Problem	مسأله نسبت وابستگی
Ratio cut problem	مسأله نسبت برش
Spectral Clustering	خوشه بندی طیفی
Similarity Graph	گراف شباهت

Hakim Sabzevari University

An Outline of MSc. Thesis



Surname:Fayaz	Name:Tayebe	Student No.:9213133019
Supervisor: Dr. Mahmood Amintoosi		
Advisor: Dr. Yaser Alizadeh		
Faculty of Mathematics and Computer Science	Applied Mathematics	Operational Research
Title of thesis: Spectral Clustering For Image Segmentation		
Keywords: Image Segmentation, Graph Partitioning, Spectral Methods, Genetic Algorithm		
<p>Abstract: Clustering is one of the most important problems in the field of artificial intelligence. One of the applications of the clustering is image segmentation. The purpose of image segmentation is separating the pixels of image into two or several related groups. Transformation of the image processing issue to graph issue is one of the image segmentation methods that is called spectral clustering. Spectral clustering refers to a class of techniques which rely on the eigenstructure of a similarity matrix to partition points into disjoint clusters with points in the same cluster having high similarity and points in different clusters having low similarity. Recently, a variety of clustering algorithms have been proposed to handle data that is not linearly separable. Spectral clustering and kernel k-means are two such methods that are seemingly quite different. In this thesis we presented spectral partitioning methods as a tool for graph partitioning, and we showed that a general weighted kernel k-means objective is mathematically equivalent to a weighted graph partitioning objective. We used one of the spectral clustering method is called normalized cut for segmentation of MRI image and we presented a solution for determining one parameter of normalized cut using the Genetic Algorithm.</p>		



Hakim Sabzevari University

Faculty of Mathematics and Computer Science

**A Thesis Submitted in Partial Fulfilment of the Requirement for the
Degree of Master of Science in Applied Mathematics**

Spectral Clustering For Image Segmentation

Supervisor:

Dr. Mahmood Amintoosi

Advisor:

Dr. Yaser Alizadeh

By:

Tayebe Fayaz

September 2015