حل مساله‌ی خوشه‌بندی طیفی با استفاده از روش لاگرانژ افزوده‌شده

دانشجو: محمدرضا صادقی

استاد راهنما: دکتر محمدعلی اخایی

دانشکده مهندسی برق و کامپیوتر، دانشگاه تهران

مقدمه

گراف‌ها ابزار ریاضی مناسبی برای بیان روابط متقابل بین داده‌ها هسستند. در تئوری گراف، داده‌ها با گره‌ها و روابط بین آنها با یال‌های گراف مدل می‌شود. نمونه‌ی نمایش داده‌ها به صورت گراف را می‌توان در بسیاری از کاربردها از جمله شبکه‌های اجتماعی و الگوهای ترافیکی در شبکه‌های کامپیوتری مشاهده کرد.

هدف این پروژه، خوشه‌بندی گره‌های یک گراف یال‌دار است به گونه ای که گره‌هایی از گراف که یال های بین آنها وزن بیشتری دارند با احتمال بیشتری در یک خوشه قرار گیرند. به این نوع خوشه‌بندی، خوشه‌بندی طیفی می‌گویند.

مدل پیشنهادی

برای حل مساله‌ی خوشه‌بندی طیفی باید به هر گره‌ی گراف وزن‌دار یک بردار تخصیص داده شود (بردار های $u \in R^{1\times K}$) به گونه ای که در فضای جدید آن دسته از گره هایی که وزن بیشتری روی یال متصل کننده‌ی آن‌ها است، به یکدیگر نزدیکتر باشند. برای این منظور تابع خطایی تعریف می شود که در نقطه‌ی کمینه‌ی آن، بردارهایی بدست آیند که ویژگی‌های یاد شده را دارا باشند. تابع خطای مذکور به شکل زیر است:

$$B = w_{ij}\left|\left|u_i - u_j\right|\right|^2 = Tr(U^T L U)$$
$$s.t \qquad U^T U = I_{K\times K}$$

در رابطه ی فوق $U = [u_1, \dots, u_N]^T$ و L ماتریس لاپلاسین گراف است . این تابع خطا در حالت کمینه به نقاطی همگرا می‌شود که وقتی وزن بین دو گره‌ی گراف زیاد باشد نقاط متناظر آن‌ها در فضای جدید نزدیک به هم قرار می گیرند. محدودیت اضافه شده در مساله‌ی بهینه سازی به این منظور است که ویژگی ها در بردار u از هم ناهمبسته باشند و الگوریتم به جواب جزیی ماتریس صفر همگرا نشود.

برای حل مساله‌ی فوق ابتدا با استفاده از الگوریتم لاگرانژ افزوده‌شده تابع خطایی تعریف می‌کنیم که محدودیت را از بین ببرد. سپس برای حل آن از الگوریتم آدام استفاده می‌کنیم. تابع خطای جدید به فرم زیر است.

$$B = \frac{1}{2}Tr(U^T L U) + Tr(\lambda^T (U^T U - I_{K\times K})) + c Tr((U^T U - I_{K\times K})^T (U^T U - I_{K\times K})))$$

برای حل مساله به روش لاگرانژ افزوده‌شده یک جمله‌ی جریمه به تابع خطا اضافه می‌شود که با آن بتوان مقادیر ماتریس $\lambda \in R^{K\times K}$ را تخمین زد. برای این منظور در هر تکرار مقدار کمینه‌ی تابع خطا (B) با استفاده از الگوریتم آدام پیدا می‌شود و مقدار λ با توجه به کمینه‌ی تابع بروز رسانی می‌شود. در هر تکرار با استفاده از الگوریتم QR نقاط بهینه‌ای که محدودیت را برآورده می‌سازند یافته می‌شوند. در آخر با استفاده از الگوریتم K-means بر روی سطرهای ماتریس U خوشه ها جداسازی می‌شوند.

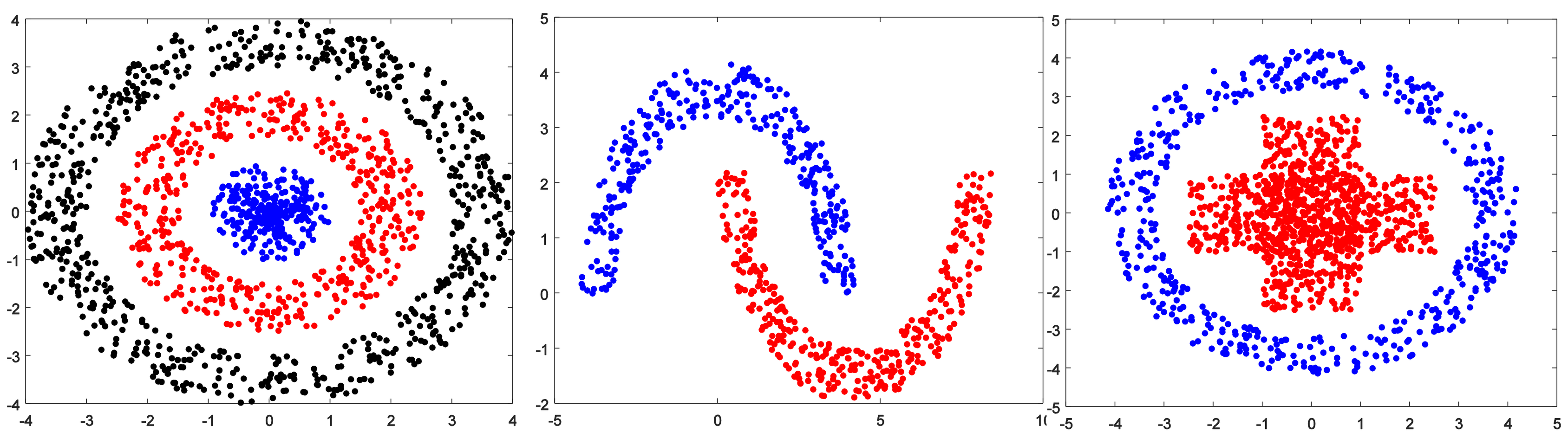
Algorithm 1 Optimization Algorithm

Input: L, c, α, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$
Output: U

$\lambda \leftarrow 0 \in R^{k\times k}$
$U \leftarrow N(0, \sigma^2) \in R^{N\times k}$
while λ *Converge* **do**
  $t \leftarrow 0$
  $m_t \leftarrow 0$
  $v_t \leftarrow 0$
  while U *Converge* **do**
    $t \leftarrow t + 1$
    $g_t \leftarrow \nabla_U B$ from (7)
    $m_t \leftarrow \beta_1 m_t + (1 - \beta_1) g_t$
    $v_t \leftarrow \beta_2 v_t + (1 - \beta_2) g_t \otimes g_t$ ▷ $\otimes$ is element wise product
    $m' = \frac{m_t}{1-\beta_1^t}$
    $v' = \frac{v_t}{1-\beta_2^t}$
    $U \leftarrow U - \alpha\, m' \oslash (\sqrt{v'} + \epsilon)$ ▷ $\oslash$ is element wise division
  $\lambda \leftarrow \lambda + c\,(U^T U - I_{k\times k})$
  $c \leftarrow \mu c$ ▷ $\mu \geq 1$ *is constant*
  $U \leftarrow QR(U)$
Apply K-means algorithm

نتایج

در این قسمت به ارزیابی کارایی الگوریتم ارائه شده پرداخته خواهد شد. الگوریتم ارائه شده برروی ۳ دیتاست ساختگی و ۶ دیتاست واقعی اجرا شده‌است و با ۴ روش دیگر مقایسه شده است. که نتایج آن به شرح زیر است.



شکل ۱: نتایج بر روی دیتاست‌های ساختگی

جدول ۱: مقایسه‌ی روش‌های مختلف روی دیتاست های واقعی

	K-means	SEI [1]	MBSC [2]	RCSF [3]	Proposed
Yeast	45.62	67.99	47.17	64.49	77.96
TOX	57.89	73.68	66.08	72.51	80.11
Glass	51.40	50.00	47.66	50.93	56.54
Ecoli	64.28	79.76	66.07	72.02	84.52
Usps	78.50	82.10	70.40	74.50	86.30
Jaffe	80.28	91.07	85.91	88.26	94.83

جمع بندی

- این پروژه یک روش نوین جهت حل مساله‌ی خوشه‌بندی طیفی ارائه داد که در آن از الگوریتم لاگرانژ افزوده‌شده به همراه روش بهینه‌سازی آدام استفاده شده است. ما با استفاده از آنالیزهای تئوری و همچنین شبیه‌سازی بر روی دیتاست‌های واقعی و مصنوعی کارایی روش خود را به اثبات رساندیم. شبیه‌سازی ها بر روی داده‌های واقعی نشان داد که روش ارائه شده از چهار روش دیگر عملکرد بهتری دارد و می‌تواند جایگزین الگوریتم‌های موجود در این حوزه شود.
- در ادامه‌ی این پروژه پژوهش‌های دیگری قابل انجام است. یکی از این پژوهش‌ها پیاده‌سازی بهینه‌ی این روش بر روی GPU ها است که باعث سریعتر شدن الگوریتم فوق می‌شود. همچنین این روش می‌تواند جایگزین الگوریتم‌های دوستیابی در شبکه‌های اجتماعی شود.

مراجع اصلی

1. El Gheche, Mireille, Giovanni Chierchia, and Pascal Frossard. "Stochastic Gradient Descent for Spectral Embedding with Implicit Orthogonality Constraint." *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019.
2. Han, Yufei, and Maurizio Filippone. "Mini-batch spectral clustering." *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2017.
3. Li, Zhihui, et al. "Rank-constrained spectral clustering with flexible embedding." *IEEE transactions on neural networks and learning systems* , 2018

شماره پوستر: ۲۳