Fast Approximate Spectral Clustering for Dynamic Networks: Supplementary Material

Lionel Martin and Andreas Loukas and Pierre Vandergheynst

1 Approximating HR as $h(\mathbf{L})\mathbf{R}$.

We compute the product with each column $\mathbf{r_i}$ of $\mathbf{R}$ independently. To achieve this using Chebyshev polynonials (Shuman et al., 2011), one employs the equation

$$\mathbf{Hr_i} \approx h(\mathbf{L})\mathbf{r_i} = \sum_{m=0}^{c} a_m \mathcal{T}_m(\mathbf{L})\mathbf{r_i},$$

where each $\mathcal{T}_m(\mathbf{L})\mathbf{r_i}$ is computed based on the recursion

$$\mathcal{T}_m(\mathbf{L})\mathbf{r_i} = \left(\frac{4}{\lambda_{\max}}\mathbf{L} - 2\mathbf{I}\right)\mathcal{T}_{m-1}(\mathbf{L})\mathbf{r_i} - \mathcal{T}_{m-2}(\mathbf{L})\mathbf{r_i},$$

having as initial conditions

$$\mathcal{T}_0(\mathbf{L})\mathbf{r_i} = \mathbf{r_i} \quad \text{and} \quad \mathcal{T}_1(\mathbf{L})\mathbf{r_i} = \left(\frac{2}{\lambda_{\max}}\mathbf{L} - \mathbf{I}\right)\mathbf{r_i}.$$

The constant a_m should be selected as $a_m = \frac{2-\delta(m)}{c+1}\sum_{j=0}^{c} s(\frac{\lambda_{\max}}{2}(1+\cos(\pi\frac{2j+1}{2(c+1)})))\cos(c\pi\frac{2j+1}{2(c+1)})$, where $s(x) = \mathbf{1}_{\{x\leq\lambda_k\}}$ is a step-function. The total computational complexity amounts to dc matrix-vector multiplications with a sparse matrix containing m non-zero elements.

2 Proof of Lemma 3.1

Proof. Let $\mathbf{X}_\Phi$ and $\mathbf{X}_\Psi$ be respectively the SC and CSC clustering assignments. Moreover, we denote for compactness the additive error term by $\mathbf{E} = \mathbf{\Psi} - \mathbf{\Phi}\mathbf{I}_{k\times d}\mathbf{Q}$. We have that

$$\begin{aligned}
C_\Psi &= \|\mathbf{\Phi} - \mathbf{X}_\Psi\mathbf{X}_\Psi^\top\mathbf{\Phi}\|_F \\
&= \|(\mathbf{I} - \mathbf{X}_\Psi\mathbf{X}_\Psi^\top)(\mathbf{\Psi} - \mathbf{E})\|_F \\
&\leq \|(\mathbf{I} - \mathbf{X}_\Psi\mathbf{X}_\Psi^\top)\mathbf{\Psi}\|_F + \|(\mathbf{I} - \mathbf{X}_\Psi\mathbf{X}_\Psi^\top)\mathbf{E}\|_F \\
&\leq \|(\mathbf{I} - \mathbf{X}_\Psi\mathbf{X}_\Psi^\top)\mathbf{\Psi}\|_F + \|\mathbf{E}\|_F \\
&\leq \|(\mathbf{I} - \mathbf{X}_\Phi\mathbf{X}_\mathbf{\Phi}^\top)\mathbf{\Psi}\|_F + \|\mathbf{E}\|_F \\
&= \|(\mathbf{I} - \mathbf{X}_\Phi\mathbf{X}_\Phi^\top)(\mathbf{\Phi}\mathbf{I}_{k\times d}\mathbf{Q} + \mathbf{E})\|_F + \|\mathbf{E}\|_F \\
&\leq \|(\mathbf{I} - \mathbf{X}_\Phi\mathbf{X}_\Phi^\top)\mathbf{\Phi}\mathbf{I}_{k\times d}\mathbf{Q}\|_F + 2\,\|\mathbf{E}\|_F \\
&= C_\Phi + 2\,\|\mathbf{\Psi} - \mathbf{\Phi}\mathbf{I}_{k\times d}\mathbf{Q}\|_F
\end{aligned} \tag{1}$$

The lower bound comes from that $\mathbf{X}_\Phi$ in eq. (1) defines the argmin of our cost functions, and thus $C_\Phi \leq C_\Psi$. □

3 Proof of Theorem 3.2

Proof. Let us start by noting that, by the unitary invariance of the Frobenius norm, for any $k \times k$ matrix $\mathbf{M}$

$$\|\mathbf{\Phi M}\|_F = \|\mathbf{U I}_{n\times k}\mathbf{M}\|_F = \|\mathbf{I}_{n\times k}\mathbf{M}\|_F = \|\mathbf{M}\|_F. \tag{2}$$

We can thus rewrite the feature error as

$$\begin{aligned}\|\mathbf{\Psi} - \mathbf{\Phi I}_{k\times d}\mathbf{Q}\|_F &= \|\mathbf{\Phi\Phi}^\top\mathbf{R} - \mathbf{\Phi I}_{k\times d}\mathbf{Q}\|_F \\ &= \|\mathbf{\Phi}^\top\mathbf{R} - \mathbf{I}_{k\times d}\mathbf{Q}\|_F \\ &= \|\mathbf{I}_{k\times n}\mathbf{U}^\top\mathbf{R} - \mathbf{I}_{k\times d}\mathbf{Q}\|_F \\ &= \|\mathbf{R}' - \mathbf{I}_{k\times d}\mathbf{Q}\|_F. \end{aligned} \tag{3}$$

We claim that there is a unitary matrix $\mathbf{Q}$ that satisfies eq. (9). We describe this matrix as follows. Let $\mathbf{R}' = \mathbf{Q}_L\mathbf{\Sigma}\mathbf{Q}_R^\top$ be the singular value decomposition of $\mathbf{R}'$ and set

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_L & 0 \\ 0 & \mathbf{I}_{d-k} \end{pmatrix} \mathbf{Q}_R^\top. \tag{4}$$

Substituting this to the feature error, we have that

$$\begin{aligned}\|\mathbf{R}' - \mathbf{I}_{k\times d}\mathbf{Q}\|_F &= \|\mathbf{Q}_L\mathbf{\Sigma}\mathbf{Q}_R^\top - \mathbf{I}_{k\times d}\mathbf{Q}\|_F \\ &= \|\mathbf{\Sigma} - \mathbf{Q}_L^\top\mathbf{I}_{k\times d}\mathbf{Q}\mathbf{Q}_R\|_F \\ &= \|\mathbf{\Sigma} - \mathbf{Q}_L^\top\mathbf{I}_{k\times d}\begin{pmatrix} \mathbf{Q}_L & 0 \\ 0 & \mathbf{I}_{d-k} \end{pmatrix}\mathbf{Q}_R^\top\mathbf{Q}_R\|_F \\ &= \|\mathbf{\Sigma} - \mathbf{Q}_L^\top\begin{pmatrix} \mathbf{Q}_L & 0 \end{pmatrix}\|_F \\ &= \|\mathbf{\Sigma} - \mathbf{I}_{k\times d}\|_F, \end{aligned} \tag{5}$$

which is the claimed result. □

4 Proof of Corollary 3.2

Proof. To obtain the following extremal inequality for the singular values of $\mathbf{R}'$, we note that $\mathbf{R}'$ is composed of i.i.d. Gaussian random variables with zero mean and variance $1/d$, and thus use Cor. 3.1 setting $\mathbf{R}' = \mathbf{N}/d$ providing for every i,

$$\sigma_i(\mathbf{R}') = \sigma_i(\mathbf{N})/\sqrt{d} \le 1 + \frac{\sqrt{k}+\varepsilon}{\sqrt{d}}. \tag{6}$$

By simple algebraic manipulation, we then find that

$$\begin{aligned}\|\mathbf{\Sigma} - \mathbf{I}_{k\times d}\|_F^2 &= \sum_{i=1}^{k}\left(\sigma_i(\mathbf{R}') - 1\right)^2 \\ &\le k\left(\frac{\sqrt{k}+\varepsilon}{\sqrt{d}}\right)^2 = \frac{k}{d}(\sqrt{k}+\varepsilon)^2, \end{aligned} \tag{7}$$

which, after taking a square root, matches the claim. □

5 Relation Between Edge Similarity and Spectral Similarity

Corollary 5.1 (adapted from Cor. 4 (Hunter and Strohmer, 2010))**.** *Let* $\mathbf{H}_{t-1}$ *and* $\mathbf{H}_t$ *be the orthogonal projection on to the span of* $[\mathbf{U}_k]_{t-1}(=\mathbf{\Phi}_{t-1})$ *and* $[\mathbf{U}_k]_t(=\mathbf{\Phi}_t)$*. If there exists an* $\alpha > 0$ *such that* $\alpha \leq \lambda_{k+1}^{(t-1)} - \lambda_k^t$ *and* $\alpha \leq \lambda_k^t$*, then,*

$$\|\mathbf{H}_t - \mathbf{H}_{t-1}\|_F \leq \frac{\sqrt{2}}{\alpha}\|\mathbf{L}_t - \mathbf{L}_{t-1}\|_F. \tag{8}$$

Note that the bounds on α are those described in their Thm. 3.

References

Hunter, B. and Strohmer, T. (2010). Performance analysis of spectral clustering on compressed, incomplete and inaccurate measurements. *arXiv preprint arXiv:1011.0997*.

Shuman, D. I., Vandergheynst, P., and Frossard, P. (2011). Chebyshev polynomial approximation for distributed signal processing. In *2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS)*, pages 1–8. IEEE.