Nyström extension

David Wobrock
david.wobrock@inria.fr

January 30, 2018

In a first phase, we would want to compute the largest eigenvalues of K. However, K is far too big to be held in memory. Instead, let's use the Nyström extension which involves computing the eigenvalues K_{AA} and approximating the largest eigenvectors of K.

[1] defines the approximation $\tilde{K}$ of K as:

$$\tilde{K} = \begin{bmatrix} K_{AA} & K_{AB} \\ K_{AB}^T & K_{AB}^T K_{AA}^{-1} K_{AB} \end{bmatrix}.$$

With K_{AA} of size $p \times p$ and $p \ll N$ with N the number of pixels. So the huge submatrix K_{BB} is approximated by $K_{BB} \approx K_{AB}^T K_{AA}^{-1} K_{AB}$. And let $\tilde{D}$ be a diagonal matrix such as $\tilde{D} = diag(\tilde{K}\mathbb{1})$.

System of linear equations

$$\begin{pmatrix} K_{AA} & K_{AB} \\ K_{AB}^T & K_{AB}^T K_{AA}^{-1} K_{AB} \end{pmatrix} \begin{pmatrix} v_A \\ v_B \end{pmatrix} = \begin{pmatrix} b_A \\ b_B \end{pmatrix}.$$

From the first line of the system

$$\begin{aligned} b_A &= K_{AA} v_A + K_{AB} v_B \\ K_{AA} v_A &= b_A - K_{AB} v_B \\ v_A &= K_{AA}^{-1}(b_A - K_{AB} v_B) \end{aligned} \tag{1}$$

All solution will have this form.

The second line of the linear system

$$\begin{aligned} b_B &= K_{AB}^T v_A + K_{AB}^T K_{AA}^{-1} K_{AB} v_B \\ &= K_{AB}^T K_{AA}^{-1}(b_A - K_{AB} v_B) + K_{AB}^T K_{AA}^{-1} K_{AB} v_B \\ &= K_{AB}^T K_{AA}^{-1} b_A - K_{AB}^T K_{AA}^{-1} K_{AB} v_B + K_{AB}^T K_{AA}^{-1} K_{AB} v_B \\ &= K_{AB}^T K_{AA}^{-1} b_A \end{aligned} \tag{2}$$

Thus if $b_B = K_{AB}^T K_{AA}^{-1} b_A$ then a solution exists, but it is not unique.

So the image of the matrix:

$$Im(\tilde{K}) = \begin{pmatrix} b_A \\ K_{AB}^T K_{AA}^{-1} b_A \end{pmatrix}.$$

Given the Nyström extension [2] and [1] $\tilde{\Phi} = \begin{bmatrix} \Phi_A \\ K_{AB}^T \Phi_A \Pi_A^{-1} \end{bmatrix}$:

$$\begin{aligned} K_{AA} &= \Phi_A \Pi_A \Phi_A^T \\ K_{AA}^{-1} &= \Phi_A \Pi_A^{-1} \Phi_A^T \\ K_{AA}^{-1} \Phi_A &= \Phi_A \Pi_A^{-1} \Phi_A^T \Phi_A \\ K_{AA}^{-1} \Phi_A &= \Phi_A \Pi_A^{-1} \end{aligned} \tag{3}$$

So, as $\tilde{\Phi} \in Im(\tilde{K})$, we can say that the largest eigenvectors

$$\begin{aligned} \tilde{\Phi} &= \begin{bmatrix} \Phi_A \\ K_{AB}^T \Phi_A \Pi_A^{-1} \end{bmatrix} \\ &= \begin{bmatrix} \Phi_A \\ K_{AB}^T K_{AA}^{-1} \Phi_A \end{bmatrix} \end{aligned} \tag{4}$$

Filter approximation Let the filter be defined from the re-normalised Laplacian definition

$$\begin{aligned} W &= I - \frac{1}{\bar{d}}(D - K) \\ &= \begin{bmatrix} W_A & W_{AB} \\ W_{AB}^T & W_B \end{bmatrix} \end{aligned} \tag{5}$$

We want the largest eigenvalues of this filter. They correspond to the smallest eigenvalues of $\mathcal{L} = \frac{1}{\bar{d}}(D - K)$. Still, holding these matrices in memory is not feasable with respect to the resolution of nowaday pictures.

Let's instead, compute the eigenvalues of $W_A = I - \frac{1}{\bar{d}}(D_A - K_A$ and extend them using Nyström: $\begin{bmatrix} V_A \\ W_{AB}^T K_{AA}^{-1} V_A \end{bmatrix}$.

We want the smallest eigenvalue $\tilde{\mu}$ of $\tilde{\mathcal{L}}$:

$$\begin{aligned} \tilde{\mu} &= min \frac{\left\langle \tilde{\mathcal{L}} \begin{pmatrix} b_A \\ K_{AB}^T K_{AA}^{-1} b_A \end{pmatrix}, \begin{pmatrix} b_A \\ K_{AB}^T K_{AA}^{-1} b_A \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} b_A \\ K_{AB}^T K_{AA}^{-1} b_A \end{pmatrix} \right\|} \\ &= min \frac{\left\langle \tilde{\mathcal{L}} \begin{pmatrix} b_A \\ K_{AB}^T K_{AA}^{-1} b_A \end{pmatrix}, \begin{pmatrix} b_A \\ K_{AB}^T K_{AA}^{-1} b_A \end{pmatrix} \right\rangle}{\langle b_A, b_A \rangle + \langle K_{AB}^T K_{AA}^{-1} b_A, K_{AB}^T K_{AA}^{-1} b_A \rangle} \end{aligned} \tag{6}$$

$\tilde{\mathcal{L}} = \frac{1}{\tilde{d}} \tilde{D} - \tilde{K}$. The system of linear equations:

$$\begin{pmatrix} \tilde{D_A} - K_{AA} & -K_{AB} \\ -K_{AB}^T & \tilde{D_C} - K_{AB}^T K_{AA}^{-1} K_{AB} \end{pmatrix} \begin{pmatrix} w_A \\ w_B \end{pmatrix} = \begin{pmatrix} c_A \\ c_B \end{pmatrix}.$$

Again, from the first line:

$$\begin{aligned} c_A &= (\tilde{D_A} - K_{AA}) w_A - K_{AB} w_B \\ (\tilde{D_A} - K_{AA}) w_A &= c_A + K_{AB} w_B \\ w_A &= (\tilde{D_A} - K_{AA})^{-1} (c_A + K_{AB} w_B) \end{aligned} \tag{7}$$

And with the second line:

$$\begin{aligned} c_B &= -K_{AB}^T w_A + (\tilde{D_C} - K_{AB}^T K_{AA}^{-1} K_{AB}) w_B \\ &= -K_{AB}^T (\tilde{D_A} - K_{AA})^{-1} (c_A + K_{AB} w_B) + (\tilde{D_C} - K_{AB}^T K_{AA}^{-1} K_{AB}) w_B \end{aligned} \tag{8}$$

With $\tilde{\mathcal{L}_A} = \tilde{D_A} - K_{AA}$ and $\tilde{\mathcal{L}_C} = \tilde{D_C} - K_{AB}^T K_{AA}^{-1} K_{AB}$ we write in a more readable manner

$$c_B = \tilde{\mathcal{L}_C} w_B - K_{AB}^T \tilde{\mathcal{L}_A}^{-1} (c_A + K_{AB} w_B) \tag{9}$$

References

[1] Hossein Talebi and Peyman Milanfar. "Global Image Denoising". In: *IEEE Transactions on Image Processing* 23.2 (Feb. 2014), pp. 755–768. ISSN: 1057-7149, 1941-0042. DOI: `10.1109/TIP.2013.2293425`. URL: `http://ieeexplore.ieee.org/document/6678291/` (visited on 10/03/2017).

[2] Charless Fowlkes et al. "Spectral grouping using the Nystrom method". In: *IEEE transactions on pattern analysis and machine intelligence* 26.2 (2004), pp. 214–225. URL: `http://ieeexplore.ieee.org/abstract/document/1262185/` (visited on 10/03/2017).