Kernel-based Active Search on Graphs

1 Introducing AS on graphs

Here is the energy function used for AS:

$$E(f) = \sum_{i \in \mathcal{L}} (y_i - f_i)^2 D_{ii} + \lambda w_0 \sum_{i \in \mathcal{U}} (f_i - \pi)^2 D_{ii} + \sum_{i,j} (f_i - f_j)^2 A_{ii}$$

Here is the energy function rewritten using matrices, where f_L and f_U are the f-vector portions belonging to the labeled and unlabeled portions respectively (they have been rearranged WLOG):

$$E(f) = \begin{bmatrix} f_L \\ f_U \\ y \\ \pi \end{bmatrix}^T \left[\begin{array}{c|c} \begin{bmatrix} D_L & 0 \\ 0 & \lambda w_0 D_U \end{bmatrix} + \lambda(D - A) & \begin{bmatrix} -D_L & 0 \\ 0 & -\lambda w_0 D_U \end{bmatrix} \\ \hline \begin{bmatrix} -D_L & 0 \\ 0 & -\lambda w_0 D_U \end{bmatrix} & 0 \end{array} \right] \begin{bmatrix} f_L \\ f_U \\ y \\ \pi \end{bmatrix}$$

The minimizer is as follows (not proven here):

$$f^* = (I - A')^{-1} D' y'$$

where

$$A' = \begin{bmatrix} \frac{\lambda}{1+\lambda} I_L & 0 \\ 0 & \frac{1}{1+w_0} I_U \end{bmatrix} D^{-1} A, \quad D' = \begin{bmatrix} \frac{1}{1+\lambda} I_L & 0 \\ 0 & \frac{w_0}{1+w_0} I_U \end{bmatrix}, \quad y' = \begin{bmatrix} y_L \\ \pi \end{bmatrix}$$

If we set $B = \begin{bmatrix} \frac{\lambda}{1+\lambda} I_L & 0 \\ 0 & \frac{1}{1+w_0} I_U \end{bmatrix}$, we have that $A' = BD^{-1}A, \quad D' = I - B$
Thus, we have our optimal solution:

$$f^* = (I - BD^{-1}A)^{-1}(I - B)y'$$

2 Kernel AS – Linear Kernel as similarity

Say $A = X^T X$ where $X = [F(x_1) \dots F(x_n)]$, with n data points and r features.
Then $D = diag(X^T X \mathbb{1})$. (Precomputed in $O(nr)$).

Thus,

$$f^* = (I - \overline{B} X^T X)^{-1} q$$

where $\overline{B} = BD^{-1}, q = (I - B)y'$.
Here, we use the Kailath variant of the matrix inverse lemma:

$$(A + BC)^{-1} = A^{-1} - A^{-1}B(I + CA^{-1}B)^{-1}CA^{-1}$$

We have:

$$\begin{aligned} f^* &= (I - \overline{B}X^T X)^{-1} q \\ &= (I + (\overline{B}X^T)(I - X\overline{B}X^T)^{-1}X)q \\ &= q + \overline{B}X^T(I - X\overline{B}X^T)^{-1}Xq \end{aligned}$$

The inverse can precomputed in $O(r^2 n + r^3)$. So the entire precomputation is in $O(r^2 n)$ assuming $n > r$.
We want to compute the updates in $O(r^2 + nr)$.

2.1 Updates to f

We have precomputed $(I - X\overline{B}X^T)^{-1}$. One element in $\overline{B}$ changes.

$$\overline{B}' = \overline{B} - \gamma e_i e_i^T$$

where e_i is the i^{th} standard basis vector.
Let $K = (I - X\overline{B}X^T)$.
Then,

$$\begin{aligned} K' &:= I - X\overline{B}'X^T \\ &= K + \gamma X e_i e_i^T X^T \\ &= K + \gamma x_i x_i^T \end{aligned}$$

Here, $\gamma = -(\frac{\lambda}{1+\lambda} - \frac{1}{1+w0})D_i i^{-1}$.
Woodbury's Matrix inversion formula:

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}$$

Using this, we have:

$$
\begin{aligned}
K'^{-1} &= K^{-1} - K^{-1}(\gamma x_i)(1+\gamma x_i^T K^{-1} x_i)^{-1} x_i^T K^{-1} \\
&= K^{-1} - \frac{\gamma K^{-1} x_i x_i^T K^{-1}}{1+\gamma x_i^T K^{-1} x_i} \\
&= K^{-1} - \frac{\gamma (K^{-1} x_i)(K^{-1} x_i)^T}{1+\gamma x_i^T K^{-1} x_i}
\end{aligned}
$$

Further, one element in q changes. $q_i' = y_i \dfrac{1}{1+\lambda}$
Thus, with the update to the inverse, $f* = q' + \overline{B}' X^T K'^{-1} X q'$. This takes $O(rn)$.

2.2 Impact factor computation