Line Search Algorithm for Fused-PGW

1 Line search algorithm

The line search for Fused-GW is defined as

$$\min_{\alpha\in[0,1]} \mathcal{L}((1-\alpha)\gamma^{(k-1)}+\alpha\gamma^{(k)\prime}), \tag{1}$$

where $\mathcal{L}(\gamma)=(1-\text{reg})\langle C,\gamma\rangle+\text{reg}\langle M,\gamma^{\otimes 2}\rangle$, $M=[|C^X_{i,i'}-C^Y_{j,j'}|]_{i,i'\in[1:n],j,j'\in[1:m]}\in\mathbb{R}^{n\times m\times n\times m}$, $(1-\text{reg}),\text{reg}\in[0,1]$ with $(1-\text{reg})+\text{reg}=1$.

Let $\delta\gamma=\gamma^{(k)\prime}-\gamma^{(k-1)}$. The above problem is essentially a quadratic problem with respect to α:

$$\begin{aligned}
&\mathcal{L}((1-\alpha)\gamma^{(k-1)}+\alpha\gamma^{(k)\prime})\\
&=\mathcal{L}(\gamma^{(k-1)}+\alpha\delta\gamma)\\
&=(1-\text{reg})\langle C,(\gamma^{(k-1)}+\alpha\delta\gamma)\rangle+\text{reg}\langle M\circ(\gamma^{(k-1)}+\alpha\delta\gamma),(\gamma^{(k-1)}+\alpha\delta\gamma)\rangle\\
&=\alpha^2\underbrace{\text{reg}\langle M\circ\delta\gamma,\delta\gamma\rangle}_{a}+\alpha\underbrace{\langle 2\text{reg}M\circ\gamma^{(k-1)}+(1-\text{reg})C,\delta\gamma\rangle}_{b}+\underbrace{\langle \text{reg}M\circ\gamma^{(k-1)}+(1-\text{reg})C,\gamma^{(k-1)}\rangle}_{c}
\end{aligned}$$

where $M\circ\gamma=[\sum_{i',j'}M_{i,j,i',j'}\gamma_{i',j'}]_{i\in[1:n,j\in[1:m]]}$, and α^* is given by

$$\alpha^*=\begin{cases}1 & \text{if } a\le 0, a+b\le 0,\\ 0 & \text{if } a\le 0, a+b>0\\ \text{clip}(\frac{-b}{2a},[0,1]), & \text{if } a>0\end{cases}. \tag{2}$$

Next, we further simplify a and b.

We first introduce some fundamental results: Let $f_1(r_1)=r_1^2$, $f_2(r_2)=r_2^2$, $h_1(r_1)=2r_1$, and $h_2(r_2)=r_2$. Then, we have:

$$|r_1-r_2|^2=f_1(r_1)+f_2(r_2)-h_1(r_1)h_2(r_2).$$

Therefore, for all $\gamma\in\mathbb{R}^{n\times m}$, we have:

$$M\circ\gamma=f_1(C^X)\gamma 1_m 1_m^\top+1_n(\gamma^\top 1_n)^\top f_2(C^Y)^\top-h_1(C^X)\gamma h_2(C^Y)^\top.$$

Since $\delta\gamma=\gamma^{(k)\prime}-\gamma^{(k)}$ and $\gamma^{(k)},\gamma^{(k)\prime}\in\Gamma(p,q):=\{\gamma\in\mathbb{R}^{n\times m}:\gamma 1_m=p,\gamma^\top 1_n=q\}$, it follows that:

$$\delta\gamma 1_m=0_n,\quad \delta\gamma^\top 1_n=0_m.$$

Thus, we have:

$$M\circ\delta\gamma=-2C^X\delta\gamma(C^Y)^\top:=-2\text{dot},$$

and therefore:

$$a = -2\mathrm{reg}\langle \mathrm{dot}, \delta\gamma \rangle.$$

Similarly,

$$\begin{aligned} b &= \langle 2\mathrm{reg} M \circ \gamma^{(k-1)} + (1-\mathrm{reg})C, \delta\gamma \rangle \\ &= (1-\mathrm{reg})\langle C, \delta\gamma \rangle + 2\mathrm{reg}\langle M \circ \delta\gamma, \gamma^{(k-1)} \rangle \\ &= (1-\mathrm{reg})\langle C, \delta\gamma \rangle + 2\mathrm{reg}\langle -2\mathrm{dot}, \gamma^{(k-1)} \rangle \\ &= \langle (1-\mathrm{reg})C, \delta\gamma \rangle - 4\mathrm{reg}\langle \mathrm{dot}, \gamma^{(k-1)} \rangle. \end{aligned}$$