SOLD: Sub-Optimal Low-rank Decomposition for Efficient Video Segmentation Supplymental material

Chenglong Li[1], Liang Lin[2*], Wangmeng Zuo[3], Shuicheng Yan[4], Jin Tang[1]

[1]School of Computer Science and Technology, Anhui University, Hefei,China

[2]School of Advanced Computing, Sun Yat-sen University, Guangzhou, China

[3]School of Computer Science and Technology, Harbin Institute of Technology, China

[4]Department of ECE, National University of Singapore, Singapore

lcl1314@foxmail.com, linliang@ieee.org, {cswmzuo, ahhftang}@gmail.com, eleyans@nus.edu.sg

1. Alternating optimization in SOLD

The objective function of SOLD is

$$\min_{\mathbf{A},\mathbf{B},\mathbf{E}} \frac{1}{2}||\mathbf{X} - \mathbf{XAB} - \mathbf{E}||_F^2 + \lambda||\mathbf{E}||_1 + \frac{\beta}{2}||\mathbf{AB}||_F^2 + \gamma\, tr((\mathbf{AB})^T\mathbf{Q}), \tag{1}$$

We adopt the alternating optimization method to optimize Eq. 1, and denote

$$J(\mathbf{A},\mathbf{B},\mathbf{E}) = \frac{1}{2}||\mathbf{X} - \mathbf{XAB} - \mathbf{E}||_F^2 + \lambda||\mathbf{E}||_1 + \frac{\beta}{2}||\mathbf{AB}||_F^2 + \gamma\, tr((\mathbf{AB})^T\mathbf{Q}). \tag{2}$$

Given $\mathbf{E}$, taking the derivative of $J(\mathbf{A},\mathbf{B},\mathbf{E})$ w.r.t. $\mathbf{B}$, and setting it to zero, we obtain

$$-\mathbf{A}^T\mathbf{X}^T(\mathbf{X} - \mathbf{XAB} - \mathbf{E}) + \beta\mathbf{A}^T\mathbf{AB} + \gamma\mathbf{A}^T\mathbf{Q} = 0. \tag{3}$$

Eq. 3 can be rewritten as follows:

$$\mathbf{B} = (\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{S}_2, \tag{4}$$

where

$$\begin{aligned} \mathbf{S}_1 &= \mathbf{X}^T\mathbf{X} + \beta\mathbf{I}, \\ \mathbf{S}_2 &= (\mathbf{X}^T(\mathbf{X} - \mathbf{E}) - \gamma\mathbf{Q}). \end{aligned} \tag{5}$$

By substituting Eq. 4 back into Eq. 1, the subproblem on $\mathbf{A}$ becomes

$$\min_{\mathbf{A}} \frac{1}{2}||(\mathbf{X} - \mathbf{E}) - \mathbf{XA}(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{S}_2||_F^2 + \frac{\beta}{2}||\mathbf{A}(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{S}_2||_F^2 + \gamma\, tr(\mathbf{S}_2^T\mathbf{A}(\mathbf{AS}_1^T\mathbf{A}^T)^{-1}\mathbf{A}^T\mathbf{Q}). \tag{6}$$

Note that $||\mathbf{x}||_F^2 = tr(\mathbf{x}^T\mathbf{x})$, we have

$$\begin{aligned} \min_{\mathbf{A}} tr(&(\mathbf{X} - \mathbf{E})^T(\mathbf{X} - \mathbf{E}) - 2(\mathbf{X} - \mathbf{E})^T\mathbf{X}^T\mathbf{A}(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{S}_2 + \mathbf{S}_2^T\mathbf{A}(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{XX}^T\mathbf{A}(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1} \\ &\mathbf{A}^T\mathbf{S}_2) + \beta tr(\mathbf{S}_2^T\mathbf{A}(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{A}(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{S}_2) + 2\gamma\, tr(\mathbf{S}_2^T\mathbf{A}(\mathbf{AS}_1^T\mathbf{A}^T)^{-1}\mathbf{A}^T\mathbf{Q}). \end{aligned} \tag{7}$$

Merging the third and the fourth term, we have

$$\min_{\mathbf{A}} tr((\mathbf{X} - \mathbf{E})^T(\mathbf{X} - \mathbf{E}) - 2(\mathbf{X} - \mathbf{E})^T\mathbf{X}^T\mathbf{A}(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{S}_2 + \mathbf{S}_2^T\mathbf{A}(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{S}_2) + 2\gamma\, tr(\mathbf{S}_2^T\mathbf{A}(\mathbf{AS}_1^T\mathbf{A}^T)^{-1}\mathbf{A}^T\mathbf{Q}). \tag{8}$$

*Corresponding author is Liang Lin.

978-1-4673-6964-0/15/$31.00 ©2015 IEEE

Substituting first $\mathbf{S}_2$ to $\mathbf{X}^T(\mathbf{X}-\mathbf{E})-\gamma\mathbf{Q}$ in the third term of Eq. 8 and employing $tr(\mathbf{x}^T)=tr(\mathbf{x})$, we obtain

$$\min_{\mathbf{A}} tr((\mathbf{X}-\mathbf{E})^T(\mathbf{X}-\mathbf{E})-(\mathbf{X}-\mathbf{E})^T\mathbf{X}^T\mathbf{A}(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{S}_2-\gamma\mathbf{Q}^T\mathbf{A}(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{S}_2)+2\gamma\, tr(\mathbf{S}_2^T\mathbf{A}(\mathbf{A}\mathbf{S}_1^T\mathbf{A}^T)^{-1}\mathbf{A}^T\mathbf{Q}), \tag{9}$$

and it equals to

$$\min_{\mathbf{A}} tr((\mathbf{X}-\mathbf{E})^T(\mathbf{X}-\mathbf{E})-(\mathbf{X}-\mathbf{E})^T\mathbf{X}^T\mathbf{A}(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{S}_2+\gamma\mathbf{Q}^T\mathbf{A}(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{S}_2). \tag{10}$$

Thus, we have

$$\min_{\mathbf{A}} tr((\mathbf{X}-\mathbf{E})^T(\mathbf{X}-\mathbf{E})-\mathbf{S}_2^T\mathbf{A}(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{S}_2). \tag{11}$$

According to Eq. 11, we utilize the fact that $tr(\mathbf{xy})=tr(\mathbf{yx})$, and solve $\mathbf{A}$ by the following program:

$$\mathbf{A}^* = \arg\max_{\mathbf{A}} tr\{(\mathbf{A}^T\mathbf{S}_1\mathbf{A})^{-1}\mathbf{A}^T\mathbf{S}_2\mathbf{S}_2^T\mathbf{A}\}. \tag{12}$$

Eq. 12 can be transformed to a generalized eigen-problem. Its global optimal solution is the top r eigenvectors of $\mathbf{S}_1^\dagger\mathbf{S}_2\mathbf{S}_2^T$ corresponding to the nonzero eigenvalues, where $\mathbf{S}_1^\dagger$ denotes the pseudo-inverse of $\mathbf{S}_1$.

Given $\mathbf{A}$ and $\mathbf{B}$, the matrix $\mathbf{E}$ can be solved by the soft-threshold (or shrinkage) method in [1]:

$$\mathbf{E}^* = \arg\min_{\mathbf{E}} \lambda\|\mathbf{E}\|_1 + \frac{1}{2}\|\mathbf{E}-(\mathbf{X}-\mathbf{XAB})\|_F^2. \tag{13}$$

A sub-optimal solution can be obtained by alternating between the updating of $\{\mathbf{A},\mathbf{B}\}$ and the updating of $\mathbf{E}$

References

[1] Z. Lin, A. Ganesh, J. Wright, M. Chen, L. Wu, and Y. Ma. Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix. *UIUC Technical Report UILU-ENG-09-2214*, July 2009.

2