Investigation of Small Increase in eFUMI Objective Function Value during Optimization

Changzhe Jiao and Alina Zare

We find that during the alternative updating of eFUMI algorithm, sometimes it happens increase in objective function value, which is counter-intuitive to the optimization of the proposed objective function although it doesn't play a big role in the estimate of target concept. So we did a lot of experiments and found the increase in objective function value came from two sources which were explained in the following parts of this report.

The synthetic data was generated following Alg. 3 and 4 discussed in [1] with parameters $n_{tar} = 100$, $N_b = 0$, $num_{pbags} = 1$, $num_{nbags} = 2$, $num_{points} = 500$, $p_{t\_mean} = 0.5$ and $\sigma = 1$. The constituent target and background endmembers are shown in Fig. 1, where Red Slate was picked as the target endmember. Gaussian white noise was added to the synthetic data that had signal-to-noise ratio equals to 30 dB.

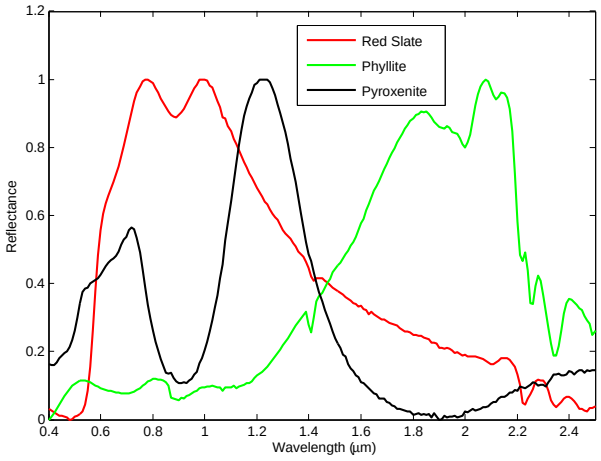


Fig. 1: Signatures from ASTER library used to generate synthetic data

I. Increase in objective function value during updating expectation of z_i in the E-step

Druing the E-step of eFUMI algorithm, the latent variable z_i that indicates the true label of point x_i in the positive bags is estimated using a RBF like function $Prob_Z(l(\mathbf{x}_i) = 1) = 1 - e^{-\beta\left\|\mathbf{x}_i - \sum_{k=1}^{M} p_{ik}\mathbf{e}_k\right\|_2^2}$, where β is a scaling parameter and $\left\|\mathbf{x}_i - \sum_{k=1}^{M} p_{ik}\mathbf{e}_k\right\|_2^2 = r_b$ is the approximation residual between $\mathbf{x}_i$ and its representation using only background endmembers. For the weak target points (points that have small amount presence of target endmember e_t), e.g. $\mathbf{x}_j$ with $p_{it} = 0.05$, its residual rb_j will decrease as the estimate of background endmembers becomes more accurate during the alternative optimization process. And this decrease in rb_j will correspondingly leads to a decrease in $Prob_Z(l(\mathbf{x}_j))$ compared with its value in the previous iteration, but ideally the probability of these weak target points to be true positive should still be one. This degeneration in the estimate of these weak target points' probability to be true positive points during the E-step will lead to an increase in the approximation error of these weak target points. An illustration is shown in Fig. 2 - 6.

Fig. 2 shows the scatter of truth points in a 2-simplex where the target points are scattered in red and non-target in blue. It can be seen that there are many weak target points. Fig. 3 scatters the probability of points from positive bags to be true positive in iteration 4, where we can find that there are many non-target points from positive bags still have non-zero probability to be target points. Fig. 4 scatters the probability of points from positive bags to be true positive in iteration 5, where we can see that all false-positive points from positive bags are successfully differentiated as negative, but at the same time some weak target points have decrease in their probability to be true-positive points, which thus leads to an increase in unmixing error for these points. The objective function value during this period increases from 55.6156 to 59.854. As shown in Fig 5 which is the scatter of change in individual approximation error between iteration 4 and 5, which clearly shows that the weak

C. Jiao and A. Zare are with the Department of Electrical and Computer Engineering, University of Missouri, Columbia, MO, 65211, USA. e-mail: zarea@missouri.edu, cjr25@mail.missouri.edu.

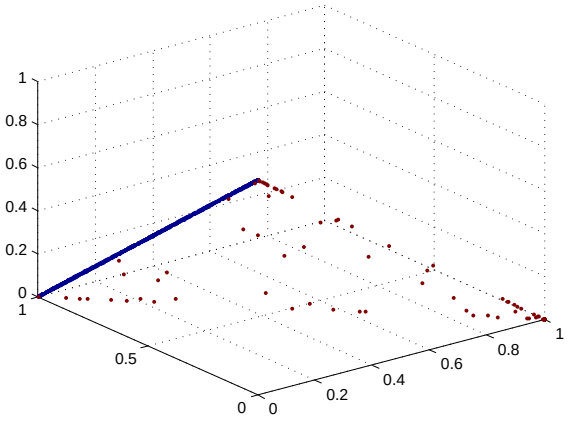


Fig. 2: Scatter of truth points

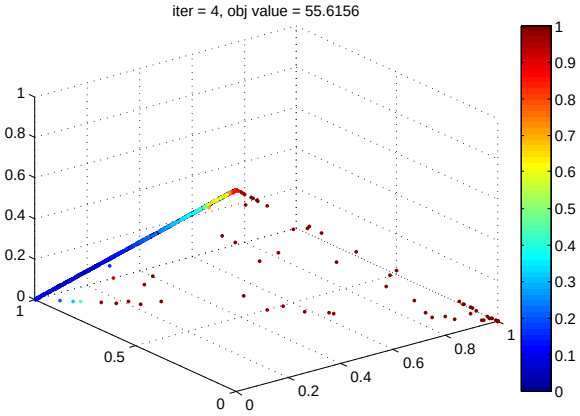


Fig. 3: Scatter of probability in iteration 4

target points (close to the line which has two background endmembers as its vertices) contribute the increase in the objective function value.

To display this more clearly, we scatter the points that have an increase in approximation error larger than 0.01 as red in Fig. 6, where it can be seen that all the red points are weak target points very close to the subspace spanned by the two background endmembers.

Increase the parameter β will be a good way to solve this problem as much as possible, if β is large enough, in $Prob_Z(l(\mathbf{x}_j) = 1) = 1 - e^{-\beta \cdot rb_j}$, although the improvement in estimate of background endmembers will decrease the approximation error rb_j for those weak target points $\mathbf{x}_j$, a large β may still make $Prob_Z(l(\mathbf{x}_j))$ close to 1. And for all non-target points $\mathbf{x}_k$, $rb_k \to 0$, these points still maintain close to 0 probability to be target as the increase of β. In the future work, investigation into an

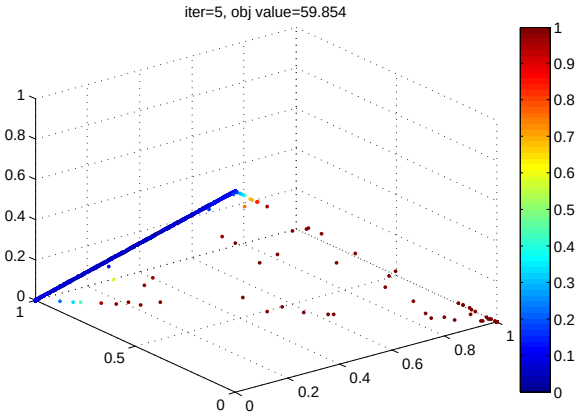


Fig. 4: Scatter of probability in iteration 5

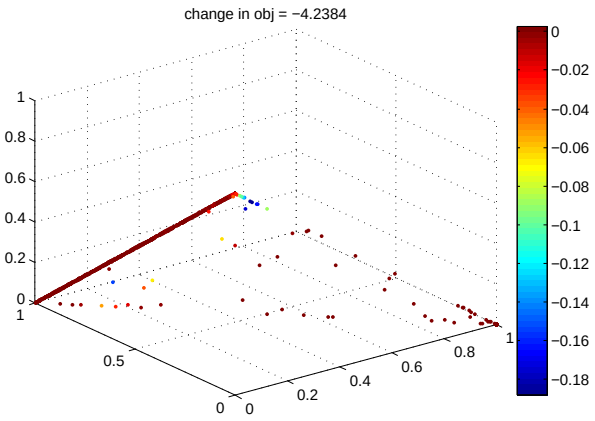


Fig. 5: Scatter of change in estimate error between iteration 4 and 5

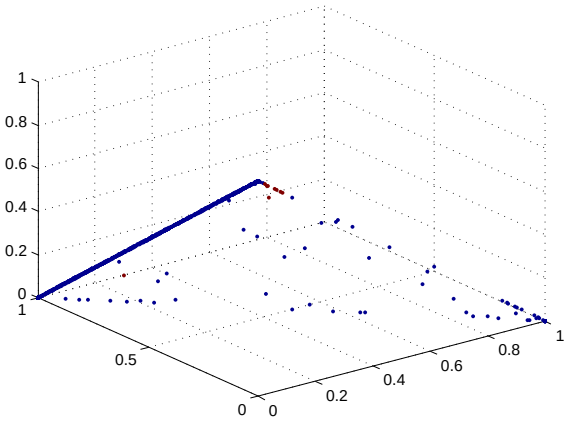


Fig. 6: Scatter of points with increased error large than a set threshold

adaptive β will better solve this problem.

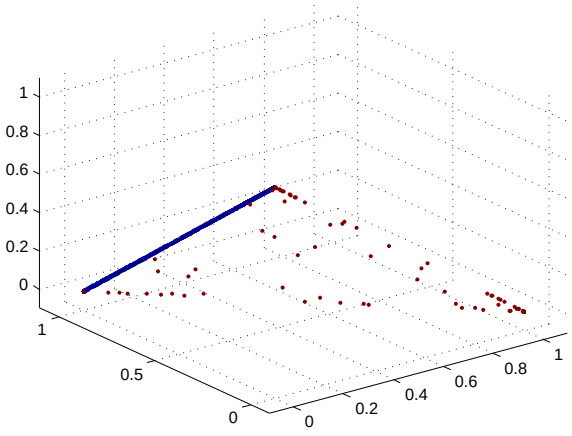


Fig. 7: Scatter of truth points

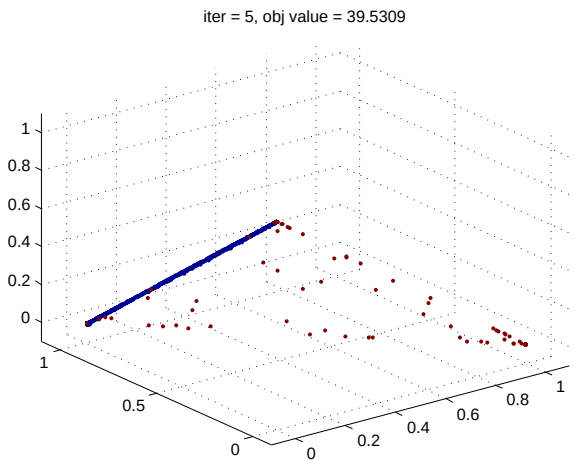


Fig. 8: Scatter of proportion values in iteration 5

II. Increase in updating proportion values $\mathbf{P}$ in the M-step

During the updating of proportion matrix $\mathbf{P}$, a non-negative constraint is added to the element of matrix $\mathbf{P}$. Since solving for the proportion value of one point is not dependent on any other points, the algorithm solving the proportion vector $\mathbf{P}_i$ for $\mathbf{x}_i$ is described in Alg. 1, where $\boldsymbol{\theta}^{t-1}$ is the parameter set acquired from the previous iteration for updating $\mathbf{P}_i$, $\mathbf{P}_i^+$ and $\mathbf{P}_i^-$ are a split of $\mathbf{P}_i$ into positive values and negative values, respectively.

Algorithm 1 eFUMI proportion vector $\mathbf{P}_i$ updating algorithm

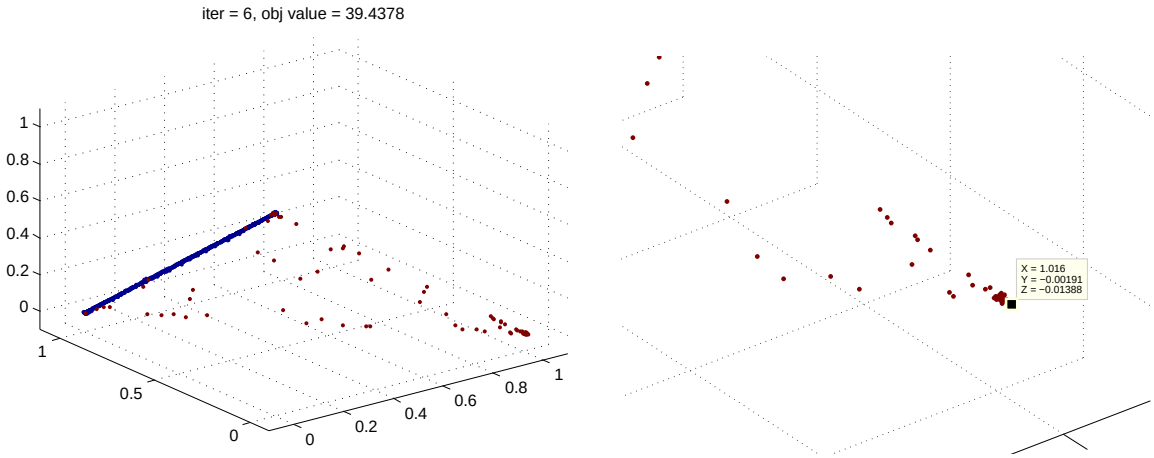
Input: $\boldsymbol{\theta}^{t-1} = \{\mathbf{x}_i, \mathbf{E}, Prob\_Z_i, \cdots\}$
Output: $\mathbf{P}_i$

$\mathbf{P}_i = f(\boldsymbol{\theta}^{t-1})$[1]
if $\mathbf{P}_i$ contains negative value **then**
    Split $\mathbf{P}_i$ into two subsets $\mathbf{P}_i^+$ and $\mathbf{P}_i^-$
    Set $\mathbf{P}_i^- = \mathbf{0}$
    Call this algorithm itself to update $\mathbf{P}_i^+$
end if
return $\mathbf{P}_i$

[1]The update equation used in $\mathbf{P}_i = f(\boldsymbol{\theta}^{t-1})$ are derived in [1].

It can be seen from Alg. 1 that the procefures to enforce non-negative constrain on $\mathbf{P}_i$ are first calculating $\mathbf{P}_i$ without non-negative constraint according the update equation $f(\boldsymbol{\theta}^{t-1})$, then setting the negative values in $\mathbf{P}_i$ to 0 and recursively solving the positive subset of $\mathbf{P}_i$ as $\mathbf{P}_i^+$. The setting of $\mathbf{P}_i^-$ to 0 makes a non-smooth increase in the objective function value. An illustration is shown in Fig. 7 - 11.

Fig. 7 shows the scatter of truth points in a 2-simplex the same as shown in Sec. 1. Fig. 8 scatters the proportion values updated in iteration 5 which has an objective function value of 39.5309. Fig. 9(a) scatters the proportion values updated in iteration 6 before the non-negative constraint is enforced and has an objective function value of 39.4378 which is lower than the value in the previous iteration. This shows that the alternative optimization indeed decreases the objective function value. But we can see from Fig. 9(a) that before the non-negative constraint is conducted, there exists points fall outside of the 2-simplex spanned by the constituent endmembers. Fig. 9(b) zooms the lower right corner of Fig. 9(a), where we can see that the noted point has proportion value [1.016, -0.00191, -0.01388]. Fig. 10 shows the proportion values updated in iteration 6 after the non-negative constraint is enforced which has an objective function value of 39.641. It is clear that enforcing the non-negative constraint leads to an increase of 0.20324 in the objective function value. Fig. 11(a) shows the change in individual approximation error between before and after the enforcement of non-negative constraint in iteration 6, which clearly indicates that the points fall outside the 2-simplex (have negative proportion values) account for the increase in the objective function value. Fig. 11(b) zooms the lower left corner of Fig. 11(a), where it can be seen that the extreme point [1.016, -0.00191, -0.01388] with both negative values with respect to two background endmembers has the largest increase in approximation error.



(a) a (b) b

Fig. 9: Scatter of proportion values in iteration 6 before enforcing non-negative constraint

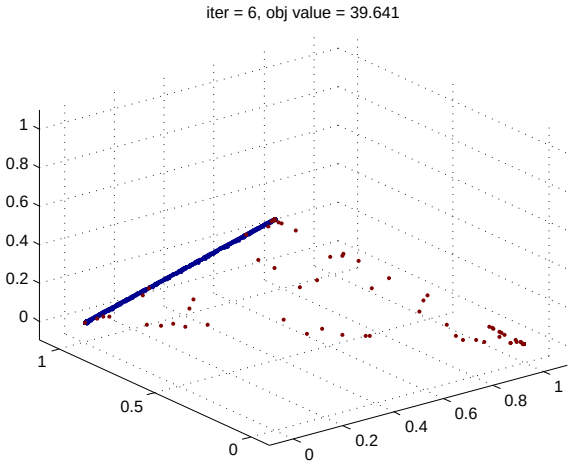
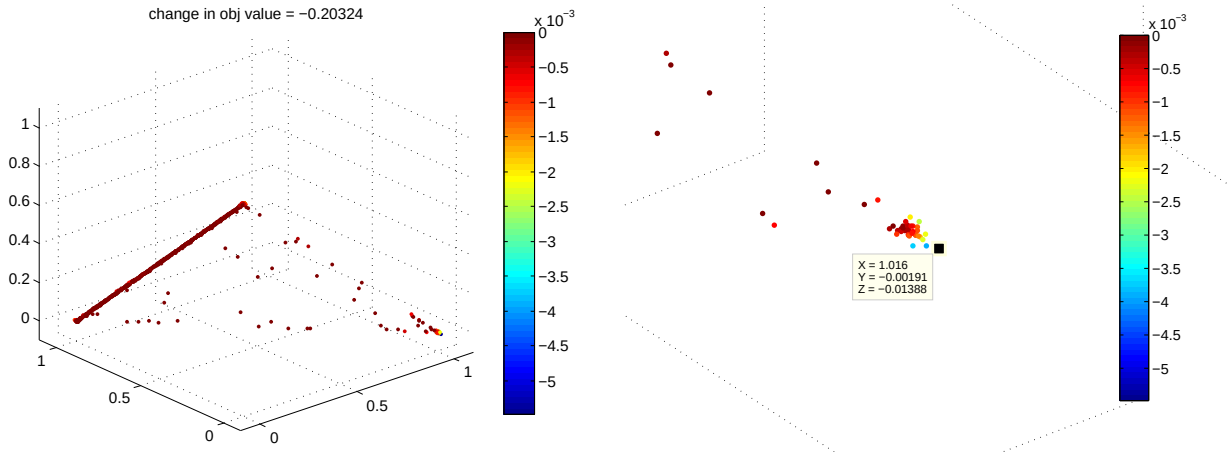


Fig. 10: Scatter of proportion values in iteration 6 after enforcing non-negative constraint



(a) a (b) b

Fig. 11: Scatter of change in estimate error between before and after enforcing non-negative constraint in iteration 6

REFERENCES

[1] C. Jiao and A. Zare, "Functions of multiple instances for learning target signatures," *IEEE Trans. Geosci. Remote Sens.*, 2015, (in press).