Minimal Basis Facility Location for Subspace Segmentation Supplementary Material

16th April 2013

1 Message update for ϕ

Following definition, the ϕ messages are written as:

$$\begin{aligned}\phi_j(1) &= \mu_{-C\to e_j}(1)\\ &= \max_{e_k, k\neq j}\left[-C(e_1,\ldots,e_j=1,\ldots,e_M)+\sum_{k\neq j}\xi_k(e_k)\right]\end{aligned} \tag{1}$$

$$\begin{aligned}\phi_j(0) &= \mu_{-C\to e_j}(0)\\ &= \max_{e_k, k\neq j}\left[-C(e_1,\ldots,e_j=0,\ldots,e_M)+\sum_{k\neq j}\xi_k(e_k)\right]\end{aligned} \tag{2}$$

$C(e_1,\ldots,e_j,\ldots,e_M)$ is effectively a feasibility function that restricts only one, two, three or four facilities to be turned on. For (1), since e_j is set as 1, we are looking for zero, one, two or three other e_j's being turned on. For (2), e_j is kept fixed as 0, we are then looking for one, two, three or four e_j's being turned on.

Even though (1) and (2) look combinatorial, the messages can be simplified and updated efficiently. We first observe that finding the max can be achieved by searching for the indices corresponding to the largest one, two, three or four $\xi_k = \xi_k(1) - \xi_k(0)$, for $k = 1,\ldots,M, k \neq j$. We sort $\{\xi_k = \xi_k(1) - \xi_k(0), k = 1,\ldots,M, k \neq j\}$ in descending order. Let $\hat{\xi}$ be the sorted $\{\xi_k = \xi_k(1) - \xi_k(0), k = 1,\ldots,M, k \neq j\}$ and the top four sorted entries be $\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3, \hat{\xi}_4$. Recall that the sorted set $\hat{\xi}$ and resultant top four indices exclude j and hence $\xi_j \notin \{\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3, \hat{\xi}_4\}$. In addition, $\hat{\xi}$ only has $M-1$ number of entries, since index j was omitted.

For the following derivation, we will use the notation $\xi_k(\{0,1\})$ to mean $\xi_k(\{e_k = 0, e_k = 1\})$, for ξ

and other variables. This notation simply means that ξ_k can take on two values, corresponding to $e_k = 0$ and $e_k = 1$.

For ease of notation, we define the cumulative sum operator S_{ij}:

$$S_{ij}(\{0,1\}) = \sum_{k=i}^{j} \hat{\xi}_k(\{0,1\}) \tag{3}$$

For example, $S_{11}(\{0,1\}) = \hat{\xi}_1(\{0,1\})$, $S_{12}(\{0,1\}) = \sum_{k=1}^{2} \hat{\xi}_k(\{0,1\})$, $S_{13}(\{0,1\}) = \sum_{k=1}^{3} \hat{\xi}_k(\{0,1\})$, $S_{23}(\{0,1\}) = \sum_{k=2}^{3} \hat{\xi}_k(\{0,1\})$.

The omit cumulative sum operator $\tilde{S}_i$, where the lower index i indicates the indices from 1 to i that are omitted in the summation:

$$\tilde{S}_i(\{0,1\}) = \sum_{k=i+1}^{M-1} \hat{\xi}_k(\{0,1\}) \tag{4}$$

For example, $\tilde{S}_0(\{0,1\}) = \sum_{k=1}^{M-1} \hat{\xi}_k(\{0,1\})$, $\tilde{S}_1(\{0,1\}) = \sum_{k=2}^{M-1} \hat{\xi}_k(\{0,1\})$, $\tilde{S}_2(\{0,1\}) = \sum_{k=3}^{M-1} \hat{\xi}_k(\{0,1\})$, $\tilde{S}_3(\{0,1\}) = \sum_{k=4}^{M-1} \hat{\xi}_k(\{0,1\})$

Since message passing is based on the value difference, we often work with

$$S_{ij} = S_{ij}(1) - S_{ij}(0) \tag{5}$$

In the derivations below, we will use the following identity frequently:

$$\tilde{S}_i(0) - \tilde{S}_j(0) = S_{(i+1)j}(0) \tag{6}$$

The differential cost between cost C_i and cost C_j is defined as

$$\delta_{ij} = C_i - C_j \tag{7}$$

For (1), since facility j is turned on, either one, two or three other facilities are turned on:

$$\phi_j(1) = \max_{e_k, k \neq j} \left[-C(e_1, \ldots, e_j = 1, \ldots, e_M) + \sum_{k \neq j} \xi_k(e_k) \right] \tag{8}$$

$$= \max \left[\overbrace{-C_1 + \tilde{S}_0(0)}^{\text{1 facility}}, \quad \overbrace{-C_2 + S_{11}(1) + \tilde{S}_1(0)}^{\text{2 facilities}}, \quad \overbrace{-C_3 + S_{12}(1) + \tilde{S}_2(0)}^{\text{3 facilities}}, \quad \overbrace{-C_4 + S_{13}(1) + \tilde{S}_3(0)}^{\text{4 facilities}} \right] \tag{9}$$

For (2), since facility j is turned off, either one, two, three or four other facilities are turned on:

$$\phi_j(0) = \max_{e_k, k\neq j}\left[-C(e_1,\ldots,e_j=0,\ldots,e_M) + \sum_{k\neq j}\xi_k(e_k)\right] \tag{10}$$

$$= \max\left[\overbrace{-C_1 + S_{11}(1) + \tilde{S}_1(0)}^{\text{1 facility}},\; \overbrace{-C_2 + S_{12}(1) + \tilde{S}_2(0)}^{\text{2 facilities}},\; \overbrace{-C_3 + S_{13}(1) + \tilde{S}_3(0)}^{\text{3 facilities}},\; \overbrace{-C_4 + S_{14}(1) + \tilde{S}_4(0)}^{\text{4 facilities}}\right] \tag{11}$$

We change the order of evaluation in computing $\phi_j = \phi_j(1) - \phi_j(0)$ by moving $\phi_j(0)$ into each term of $\phi_j(1)$.

Moving $\phi_j(0)$ into the first term of $\phi_j 1$:

$$\left[-C_1 + \tilde{S}_0(0)\right] - \phi_j(0) \tag{12}$$

$$= \max\begin{cases} -S_{11}(1) + \tilde{S}_0(0) - \tilde{S}_1(0) \\ (C_2 - C_1) - S_{12}(1) + \tilde{S}_0(0) - \tilde{S}_2(0) \\ (C_3 - C_1) - S_{13}(1) + \tilde{S}_0(0) - \tilde{S}_3(0) \\ (C_4 - C_1) - S_{14}(1) + \tilde{S}_0(0) - \tilde{S}_4(0) \end{cases} \tag{13}$$

$$= \max\begin{cases} -[S_{11}(1) - S_{11}(0)] \\ \delta_{21} - [S_{12}(1) - S_{12}(0)] \\ \delta_{31} - [S_{13}(1) - S_{13}(0)] \\ \delta_{41} - [S_{14}(1) - S_{14}(0)] \end{cases} \tag{14}$$

$$= -\max\left[S_{11},\ S_{12} - \delta_{21},\ S_{13} - \delta_{31},\ S_{14} - \delta_{41}\right] \tag{15}$$

Moving $\phi_j(0)$ into the second term of $\phi_j(1)$:

$$\left[-C_2 + S_{11}(1) + \tilde{S}_1(0)\right] - \phi_j(0) \tag{16}$$

$$= \max\begin{cases} -(C_2 - C_1) \\ S_{11}(1) - S_{12}(1) + \tilde{S}_1(0) - \tilde{S}_2(0) \\ (C_3 - C_2) + S_{11}(1) - S_{13}(1) + \tilde{S}_1(0) - \tilde{S}_3(0) \\ (C_4 - C_2) + S_{11}(1) - S_{14}(1) + \tilde{S}_1(0) - \tilde{S}_4(0) \end{cases} \tag{17}$$

$$= \max\begin{cases} -\delta_{21} \\ -[S_{22}(1) - S_{22}(0)] \\ \delta_{32} - [S_{23}(1) - S_{23}(0)] \\ \delta_{42} - [S_{24}(1) - S_{24}(0)] \end{cases} \tag{18}$$

$$= -\max\left[\delta_{21},\ S_{22},\ S_{23} - \delta_{32},\ S_{24} - \delta_{42}\right] \tag{19}$$

Moving $\phi_j(0)$ into the third term of $\phi_j(1)$:

$$\left[-C_3 + S_{12}(1) + \tilde{S}_2(0)\right] - \phi_j(0) \tag{20}$$

$$= \max\begin{cases} -(C_3 - C_1) + S_{12}(1) - S_{11}(1) + \tilde{S}_2(0) - \tilde{S}_1(0) \\ -(C_3 - C_2) \\ S_{12}(1) - S_{13}(1) + \tilde{S}_2(0) - \tilde{S}_3(0) \\ (C_4 - C_3) + S_{12}(1) - S_{14}(1) + \tilde{S}_2(0) - \tilde{S}_4(0) \end{cases} \tag{21}$$

$$= \max\begin{cases} -\delta_{31} + [S_{22}(1) - S_{22}(0)] \\ -\delta_{32} \\ -[S_{33}(1) - S_{33}(0)] \\ \delta_{43} - [S_{34}(1) - S_{34}(0)] \end{cases} \tag{22}$$

$$= -\max\left[\delta_{31} - S_{22},\ \delta_{32},\ S_{33},\ S_{34} - \delta_{43}\right] \tag{23}$$

Moving $\phi_j(0)$ into the fourth term of $\phi_j(1)$:

$$\left[-C_4 + S_{14}(1) + \tilde{S}_4(0)\right] - \phi_j(0) \tag{24}$$

$$= \max\begin{cases} -(C_4 - C_1) + S_{13}(1) - S_{11}(1) + \tilde{S}_3(0) - \tilde{S}_1(0) \\ -(C_4 - C_2) + S_{13}(1) - S_{12}(1) + \tilde{S}_3(0) - \tilde{S}_2(0) \\ -(C_4 - C_3) \\ S_{13}(1) - S_{14}(1) + \tilde{S}_3(0) - \tilde{S}_4(0) \end{cases} \tag{25}$$

$$= \max\begin{cases} -\delta_{41} + [S_{23}(1) - S_{23}(0)] \\ -\delta_{42} + [S_{33}(1) - S_{33}(0)] \\ -\delta_{43} \\ -[S_{44}(1) - S_{44}(0)] \end{cases} \tag{26}$$

$$= -\max\left[\delta_{41} - S_{23},\ \delta_{42} - S_{33},\ \delta_{43},\ S_{44}\right] \tag{27}$$

The ϕ_j message update can now be simplified as

$$\phi_j = \max\begin{cases} -\max\left[S_{11},\ S_{12} - \delta_{21},\ S_{13} - \delta_{31},\ S_{14} - \delta_{41}\right] \\ \\ -\max\left[\delta_{21},\ S_{22},\ S_{23} - \delta_{32},\ S_{24} - \delta_{42}\right] \\ \\ -\max\left[\delta_{31} - S_{22},\ \delta_{32},\ S_{33},\ S_{34} - \delta_{43}\right] \\ \\ -\max\left[\delta_{41} - S_{23},\ \delta_{42} - S_{33},\ \delta_{43},\ S_{44}\right] \end{cases} \tag{28}$$

The underlying pattern can be more easily discerned by discarding the max symbols and presenting the various entries in a matrix:

$$\begin{pmatrix} S_{11} & S_{12}-\delta_{21} & S_{13}-\delta_{31} & S_{14}-\delta_{41} \\ \delta_{21} & S_{22} & S_{23}-\delta_{32} & S_{24}-\delta_{42} \\ \delta_{31}-S_{22} & \delta_{32} & S_{33} & S_{34}-\delta_{43} \\ \delta_{41}-S_{23} & \delta_{42}-S_{33} & \delta_{43} & S_{44} \end{pmatrix} \tag{29}$$

This pattern can be generalized to K facilities

$$\begin{pmatrix} S_{11} & S_{12}-\delta_{21} & \dots & \dots & \dots & \dots & \dots & S_{1K}-\delta_{K1} \\ \delta_{21} & S_{22} & S_{23}-\delta_{32} & \dots & \dots & \dots & \dots & S_{2K}-\delta_{K2} \\ \delta_{31}-S_{22} & \delta_{32} & S_{33} & S_{34}-\delta_{43} & \dots & \dots & \dots & S_{3K}-\delta_{K3} \\ \delta_{41}-S_{23} & \delta_{42}-S_{33} & \delta_{43} & S_{44} & \dots & \dots & \dots & S_{4K}-\delta_{K4} \\ \vdots & \vdots & \vdots & \vdots & & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & & S_{(K-1)K}-\delta_{K(K-1)} \\ \delta_{K1}-S_{2(K-1)} & \delta_{K2}-S_{3(K-1)} & \dots & \dots & \dots & \delta_{K(K-2)}-S_{(K-1)(K-1)} & \delta_{K(K-1)} & S_{KK} \end{pmatrix} \tag{30}$$

2 Message update for ξ

$$\begin{aligned} \xi_j(1) &= \mu_{E_j \to e_j}(1) \\ &= \max_{h_{:j}} \left[E_j(h_{:j}, e_j = 1) + \sum_i \rho_{ij}(h_{ij}) \right] \end{aligned} \tag{31}$$

$$\begin{aligned} \xi_j(0) &= \mu_{E_j \to e_j}(0) \\ &= \max_{h_{:j}} \left[E_j(h_{:j}, e_j = 0) + \sum_i \rho_{ij}(h_{ij}) \right] && (32) \\ &= \sum_i \rho_{ij}(0), \text{ since facility } j \text{ is not turned on} && (33) \end{aligned}$$

$$
\begin{aligned}
\xi_j &= \xi_j(1) - \xi_j(0) && (34)\\
&= \max_{h_{:j}} \left[E_j(h_{:j}, e_j = 1) + \sum_i \rho_{ij}(h_{ij}) \right] - \sum_i \rho_{ij}(0) && (35)\\
&= \max_{h_{:j}} \left[E_j(h_{:j}, e_j = 1) + \sum_i \rho_{ij}(h_{ij}) - \sum_i \rho_{ij}(0) \right] && (36)\\
&= \rho_{kj} + \sum_{i \neq k} \max(0, \rho_{ij}), \text{ since at least one facility must be turned on and} && (37)\\
&\quad k \text{ is the index of the largest } \rho \text{ value}
\end{aligned}
$$

Since e_j is turned on, one of the h_{ij} must be turned on as well, otherwise the consistency constraint is violated. The max operation is therefore taken over all combinations of h_{ij}, with at least one of the h_{ij}'s set to 1. The one h_{ij} turned on can be readily identified as the largest value in ρ, say ρ_k. The rest of the ρ_{ij}'s only contribute to the sum if they are positive

For efficient Matlab implementation, ξ_j can be written as

$$
\begin{aligned}
\rho_{kj} + \sum_{i \neq k} \max(0, \rho_{ij}) &= \rho_{kj} - \max(0, \rho_{kj}) + \sum_i \max(0, \rho_{ij}) && (38)\\
&= \min(0, \rho_{kj}) + \sum_i \max(0, \rho_{ij}), && (39)\\
&\quad \text{using the relationship } x - \max(x, y) = \min(0, x - y) && (40)
\end{aligned}
$$

3 Message update for α

The other message update that is affected by the global facility function is α.

$$
\begin{aligned}
\alpha_{ij}(1) &= \max_{h_{kj}, k \neq i} \left[E(h_{1j}, \ldots, h_{ij} = 1, \ldots, h_{Nj}, e_j) + \sum_{k \neq i} \rho_{kj}(h_{kj}) + \phi_j(e_j) \right] && (41)\\
&= \sum_{k \neq i} \max_{h_{kj}} \rho_{kj}(h_{kj}) + \phi_j(1) && (42)
\end{aligned}
$$

$$
\begin{aligned}
\alpha_{ij}(0) &= \max_{h_{kj}, k \neq i} \left[E(h_{1j}, \ldots, h_{ij} = 0, \ldots, h_{Nj}, e_j) + \sum_{k \neq i} \rho_{kj}(h_{kj}) + \phi_j(e_j) \right] && (43)\\
&= \max \left[\sum_{k \neq i} \max_{h_{kj}} \rho_{kj}(h_{kj}) + \phi_j(1), \sum_{k \neq i} \rho_{kj}(0) + \phi_j(0) \right] && (44)
\end{aligned}
$$

$$
\begin{aligned}
\alpha_{ij} &= \alpha_{ij}(1) - \alpha_{ij}(0) && (45)\\
&= \min\left[0, \sum_{k\neq i} \max_{h_{kj}} \rho_{kj}(h_{kj}) + \phi_j(1) - \sum_{k\neq i} \rho_{kj}(0) - \phi_j(0)\right] && (46)\\
&= \min\left[0, \sum_{k\neq i} \max(0, \rho_{kj}) + \phi_j\right] && (47)
\end{aligned}
$$

where we have used the notation $\max_{h_{kj}} \rho_{kj}(h_{kj})$ to mean $\max\left[\rho_{kj}(h_{kj}=1), \rho_{kj}(h_{kj}=0)\right]$