1 Mean-field for pairwise clustering

Self-consistent equations for the mean-field ⊝

$$\frac{\partial \left\langle E^p \right\rangle_Q}{\partial e_r^{(\alpha)}} = \sum_p \frac{\partial \left\langle m_p^{(\alpha)} \right\rangle_Q}{\partial e_r^{(\alpha)}} e_r^{(\alpha)} \tag{1}$$

useful relations:
①

$$\frac{m_p^{(\alpha)}}{\sum\limits_{\gamma} m_p^{(\gamma)}} = \frac{m_p^{(\alpha)}}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1} \tag{2}$$

② from $\frac{1}{a+b} = \frac{1}{a}\left(1 - \frac{b}{a+b}\right)$ it follows

$$\begin{aligned} \frac{1}{\sum\limits_{\gamma} m_p^{(\gamma)}} &= \frac{1}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + m_p^{(\alpha)}} \\ &= \frac{1}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)}} \left\{ 1 - \frac{m_p^{(\alpha)}}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + m_p^{(\alpha)}} \right\} \\ &= \frac{1}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)}} \left\{ 1 - \frac{m_p^{(\alpha)}}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1} \right\} \end{aligned} \tag{3}$$

derivations of the expected cost:

$$\begin{aligned} \frac{\partial \left\langle E^p \right\rangle_Q}{\partial e_r^{(\alpha)}} &= \frac{1}{M} \sum_{p,\delta,\varepsilon} \underbrace{\frac{\partial}{\partial e_r^{(\alpha)}} \left\langle \frac{m_p^{(\delta)} m_p^{(\varepsilon)}}{\sum\limits_{\gamma} m_p^{(\gamma)}} \right\rangle_Q d_{\delta\varepsilon}}_{\circledast} \\ &= \frac{1}{M} \sum_p \left\{ \sum_{\substack{\delta,\varepsilon \\ \delta \neq \varepsilon; \delta,\varepsilon \neq \alpha}} \circledast + \sum_{\substack{\delta \\ \delta \neq \alpha; \varepsilon = \alpha}} \circledast + \sum_{\substack{\varepsilon \\ \varepsilon \neq \alpha; \delta = \alpha}} \circledast + \sum_{\substack{\delta,\varepsilon \\ \delta = \varepsilon; \delta,\varepsilon \neq \alpha}} \circledast + \sum_{\substack{\delta,\varepsilon \\ \delta = \varepsilon = \alpha}} \circledast \right\} \end{aligned} \tag{4}$$

calculation of the individual expectations

① $\delta, \varepsilon : \delta \neq \varepsilon; \delta, \varepsilon \neq \alpha$

$$\begin{aligned}
\left\langle \frac{m_p^{(\delta)} m_p^{(\epsilon)}}{\sum\limits_{\gamma} m_p^{(\gamma)}} \right\rangle_Q \quad &\underbrace{=}_{\text{using ②}} \left\langle \frac{m_p^{(\delta)} m_p^{(\varepsilon)}}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)}} \left\{ 1 - \frac{m_p^{(\alpha)}}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1} \right\} \right\rangle_Q \\
&= \left\langle \frac{m_p^{(\delta)} m_p^{(\varepsilon)}}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)}} \right\rangle_Q - \left\langle \frac{m_p^{(\delta)} m_p^{(\varepsilon)}}{\left(\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} \right) \left(\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1 \right)} \right\rangle_Q \left\langle m_p^{(\alpha)} \right\rangle_Q
\end{aligned} \tag{5}$$

② $\delta : \delta \neq \alpha; \varepsilon = \alpha$

$$\begin{aligned}
\left\langle \frac{m_p^{(\delta)} m_p^{(\alpha)}}{\sum\limits_{\gamma} m_p^{(\gamma)}} \right\rangle_Q \quad &\underbrace{=}_{\text{using ①}} \left\langle \frac{m_p^{(\delta)} m_p^{(\alpha)}}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1} \right\rangle_Q \\
&= \left\langle \frac{m_p^{(\delta)}}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1} \right\rangle_Q \left\langle m_p^{(\alpha)} \right\rangle_Q
\end{aligned} \tag{6}$$

③ $\varepsilon : \varepsilon \neq \alpha; \delta = \alpha$

$$\left\langle \frac{m_p^{(\alpha)} m_p^{(\varepsilon)}}{\sum\limits_{\gamma} m_p^{(\gamma)}} \right\rangle_Q \quad \underbrace{=}_{\substack{\text{in analogy} \\ \text{to ②}}} \left\langle \frac{m_p^{(\varepsilon)}}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1} \right\rangle_Q \left\langle m_p^{(\alpha)} \right\rangle_Q \tag{7}$$

④ $\delta, \varepsilon : \delta = \varepsilon; \delta, \varepsilon \neq \alpha$

$$\begin{aligned}
\left\langle \frac{\left(m_p^{(\varepsilon)} \right)^2}{\sum\limits_{\gamma} m_p^{(\gamma)}} \right\rangle_Q \quad &= \left\langle \frac{m_p^{(\varepsilon)}}{\sum\limits_{\gamma} m_p^{(\gamma)}} \right\rangle_Q \\
&\underbrace{=}_{\text{using ②}} \left\{ \left\langle \frac{m_p^{(\varepsilon)}}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)}} \right\rangle_Q - \left\langle \frac{m_p^{(\varepsilon)}}{\left(\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} \right) \left(\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1 \right)} \right\rangle_Q \left\langle m_p^{(\alpha)} \right\rangle_Q \right\}
\end{aligned} \tag{8}$$

⑤ $\delta, \varepsilon : \delta = \varepsilon = \alpha$

$$\begin{aligned}
\left\langle \frac{\left(m_p^{(\alpha)} \right)^2}{\sum\limits_{\gamma} m_p^{(\gamma)}} \right\rangle_Q \quad &= \left\langle \frac{m_p^{(\alpha)}}{\sum\limits_{\gamma} m_p^{(\gamma)}} \right\rangle_Q \\
&\underbrace{=}_{\text{using ①}} \left\langle \frac{1}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1} \right\rangle_Q \left\langle m_p^{(\alpha)} \right\rangle_Q
\end{aligned} \tag{9}$$

inserting this into the derivative we obtain:

$$
\begin{aligned}
\frac{\partial \left\langle E^p \right\rangle_Q}{\partial e_r^{(\alpha)}} &= \frac{1}{M} \sum_p \frac{\partial \left\langle m_p^{(\alpha)} \right\rangle_Q}{\partial e_r^{(\alpha)}} \cdot \\
&\Bigg\{ \underbrace{- \sum_{\substack{\delta,\varepsilon \\ \delta \neq \varepsilon; \delta,\varepsilon \neq \alpha}} \left\langle \frac{m_p^{(\delta)} m_p^{(\varepsilon)}}{\left(\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} \right) \left(\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1 \right)} \right\rangle_Q d_{\delta\varepsilon}}_{\text{from I}} \\
&\underbrace{+ \sum_{\substack{\delta \\ \delta \neq \alpha}} \left\langle \frac{m_p^{(\delta)}}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1} \right\rangle_Q \left(d_{\delta\alpha} + d_{\alpha\delta} \right)}_{\text{from II and III}} \\
&\underbrace{- \sum_{\substack{\delta \\ \delta \neq \alpha}} \left\langle \frac{m_p^{(\delta)}}{\left(\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} \right) \left(\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1 \right)} \right\rangle_Q \left\langle m_p^{(\alpha)} \right\rangle_Q d_{\delta\delta}}_{\text{from IV}} \\
&\underbrace{+ \left\langle \frac{1}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1} \right\rangle_Q d_{\alpha\alpha}}_{\text{from V}} \Bigg\}
\end{aligned}
\tag{10}
$$

using eq. ⊝ and comparing coefficients, we find

$$
\begin{aligned}
e_p^{(\alpha)} &= \frac{1}{M} \left\langle \frac{1}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1} \right\rangle_Q d_{\alpha\alpha} + \frac{1}{M} \sum_{\delta \neq \alpha} \left\langle \frac{m_p^{(\delta)}}{\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1} \right\rangle_Q \left\{ d_{\delta\alpha} + d_{\alpha\delta} \right\} \\
&\quad - \frac{1}{M} \sum_{\delta \neq \alpha} \left\langle \frac{m_p^{(\delta)}}{\left(\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} \right) \left(\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1 \right)} \right\rangle_Q d_{\delta\delta} \\
&\quad - \frac{1}{M} \sum_{\substack{\delta,\varepsilon \\ \delta \neq \varepsilon; \delta,\varepsilon \neq \alpha}} \left\langle \frac{m_p^{(\delta)} m_p^{(\varepsilon)}}{\left(\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} \right) \left(\sum\limits_{\gamma \neq \alpha} m_p^{(\gamma)} + 1 \right)} \right\rangle_Q d_{\delta\varepsilon}
\end{aligned}
\tag{11}
$$

approximations for large numbers of data points (neglecting terms Q_{y_p})

$$
\begin{aligned}
\left\langle \frac{1}{\sum\limits_{\gamma\neq\alpha} m_p^{(\gamma)}+1} \right\rangle_Q &\approx \frac{1}{\sum\limits_{\gamma} \left\langle m_p^{(\gamma)} \right\rangle_Q} && \text{proof via Taylor expansion} \\
\left\langle \frac{m_p^{(\delta)}}{\sum\limits_{\gamma\neq\alpha} m_p^{(\gamma)}+1} \right\rangle_Q &\approx \frac{\left\langle m_p^{(\delta)} \right\rangle_Q}{\sum\limits_{\gamma} \left\langle m_p^{(\gamma)} \right\rangle_Q} && \text{using above and relation ①} \\
\left\langle \frac{m_p^{(\delta)} m_p^{(\varepsilon)}}{\left(\sum\limits_{\gamma\neq\alpha} m_p^{(\gamma)}\right)\left(\sum\limits_{\gamma\neq\alpha} m_p^{(\gamma)}+1\right)} \right\rangle_Q &\approx \frac{\left\langle m_p^{(\delta)} \right\rangle_Q \left\langle m_p^{(\varepsilon)} \right\rangle_Q}{\left(\sum\limits_{\gamma} \left\langle m_p^{(\gamma)} \right\rangle_Q\right)^2} && \text{using above and relation ①}
\end{aligned}
\tag{12}
$$

leads to the self-consistent equations for the mean-fields:

$$
\begin{aligned}
e_p^{(\alpha)} \;=\; & \frac{1}{M}\left[\frac{1}{\sum\limits_{\gamma} \left\langle m_p^{(\gamma)} \right\rangle_Q} \left\{ d_{\alpha\alpha} - \frac{1}{\sum\limits_{\gamma} \left\langle m_p^{(\gamma)} \right\rangle_Q} \sum_{\delta\neq\alpha} \left\langle m_p^{(\delta)} \right\rangle_Q d_{\delta\delta} \right\} \right. \\
& + \frac{1}{\sum\limits_{\gamma} \left\langle m_p^{(\gamma)} \right\rangle_Q} \sum_{\delta\neq\alpha} \left\langle m_p^{(\delta)} \right\rangle_Q \left\{ \left(d_{\delta\alpha} + d_{\alpha\delta} \right) \right. \\
& \left. \left. - \frac{1}{2} \frac{1}{\sum\limits_{\gamma} \left\langle m_p^{(\gamma)} \right\rangle_Q} \sum_{\substack{\delta,\varepsilon \\ \delta\neq\varepsilon;\delta,\varepsilon\neq\alpha}} \left\langle m_p^{(\varepsilon)} \right\rangle_Q \left(d_{\delta\varepsilon} + d_{\varepsilon\delta} \right) \right\} \right]
\end{aligned}
\tag{13}
$$