① a)

~~Solution w_i for $\vec{x}_i$~~

Reconstruction weights w_i for $\vec{x}_i$ are found like:

$$w_i = \frac{C^{-1}\mathbb{1}}{\mathbb{1}^T C^{-1}\mathbb{1}} \quad \text{with } C_{jk} = (\vec{x}_i - \eta_j)^T (x_i - \eta_k)$$

$j,k \in N_i$ $\quad N_i \hat{=}$ number of neighbors for each $\vec{x}_i$

Scaling all ~~for~~ data points by α affects the matrix C in the following way:

$$C_{jk} = (\alpha\vec{x}_i - \alpha\eta_j)^T(\alpha x_i - \alpha\eta_k) = \left(\alpha(\vec{x}_i - \vec{\eta}_j)\right)^T\left(\alpha(\vec{x}_i - \vec{\eta}_k)\right)$$

$$= \alpha^2 (\vec{x}_i - \vec{\eta}_j)^T(\vec{x}_i - \vec{\eta}_k) = \alpha^2 C_{jk}$$

So w_i for $\alpha\vec{x}_i$ amounts to

$$w_i = \frac{\left(\frac{1}{\alpha^2}C\right)^{-1}\mathbb{1}}{\mathbb{1}^T\left(\frac{1}{\alpha^2}C\right)^{-1}\mathbb{1}} = \frac{\cancel{\frac{1}{\alpha^2}}\, C^{-1}\mathbb{1}}{\cancel{\frac{1}{\alpha^2}}\, \mathbb{1}^T C^{-1}\mathbb{1}}$$

which is the same as before scaling of the data

b) For translation of all data points we can show that the matrix C is unaffected.

$$C_{jk} = \left((\vec{x}_i + \vec{v}) - (\eta_j + \vec{v})\right)^T\left((\vec{x}_i + \vec{v}) - (\eta_k + \vec{v})\right)$$

$$= (\vec{x}_i - \vec{\eta}_j)^T(\vec{x}_i - \vec{\eta}_k)$$

as w_i depends only on C and C is unaffected by translation w_i isn't either.

c) To show that C is unaffected by rotation of data points we use the Facts: ~~that~~

1. Rotations are described by orthogonal matricies.
2. For an orthogonal matrix U $U^T U = I$ since $U^{-1} = U^T$

$$C_{jk} = (U\vec{x}_i - U\vec{\eta}_j)^T (U\vec{x}_i - U\vec{\eta}_j) = (U(\vec{x}_i - \vec{\eta}_j))^T (U(\vec{x}_i - \vec{\eta}_j))$$

$$= (\vec{x}_i - \vec{\eta}_j)^T U^T U (\vec{x}_i - \vec{\eta}_j) = (\vec{x}_i - \vec{\eta}_j)^T (\vec{x}_i - \vec{\eta}_j)$$

② a) $\mathcal{E} = \left| \vec{x} - \sum_j^N w_j y_j \right|^2$ $N \hat{=}$ number of neighbors

$\vec{x} = \sum_j w_j \vec{x}$ iff $\sum_j w_j = 1$

$\mathcal{E} = \left| \sum_j^N w_j (\vec{x} - \vec{y}_i) \right|^2 = \left(\sum_j^N w_j (\vec{x} - \vec{y}_i) \right) \left(\sum_k^N w_k (\vec{x} - \vec{y}_k) \right)$

$= \sum_{jk} w_j w_k \underbrace{(\vec{x} - \vec{y}_j)(\vec{x} - \vec{y}_k)}_{C_{jk}} = w^T C w = \mathcal{E}$ given $\mathbb{1}^T w = 1$

C_{jk} for $C = (\mathbb{1}\vec{x}^T - y)(\mathbb{1}\vec{x}^T - y)^T$

b) $\min_w w^T C w$ subject to $w^T \mathbb{1} = 1$

$\min_w w^T C w - \lambda (w^T \mathbb{1} - 1)$

$\frac{\partial}{\partial w} \left(w^T C w - \lambda (w^T \mathbb{1} - 1) \right) = 2 C w - \lambda \mathbb{1} \stackrel{!}{=} 0$

$C w = \frac{\lambda}{2} \mathbb{1} \Leftrightarrow w = \underbrace{\frac{\lambda}{2}}_{\substack{\text{constant} \\ \text{scaling factor}}} C^{-1} \mathbb{1}$ $\Rightarrow$ direction of $w = C^{-1} \mathbb{1}$

Restraint $\mathbb{1}^T w = 1$ $w = \frac{C^{-1} \mathbb{1}}{a}$ $1 = \mathbb{1}^T w = \mathbb{1}^T \frac{C^{-1} \mathbb{1}}{a}$

$\Rightarrow a = \mathbb{1}^T C^{-1} \mathbb{1}$ $w = \frac{C^{-1} \mathbb{1}}{\mathbb{1}^T C^{-1} \mathbb{1}}$

c) $2Cw + \lambda \mathbb{1} \stackrel{!}{=} 0$

$$Cw = \mathbb{1} \cdot \left(-\frac{\lambda}{2}\right)$$

$$\left(-\frac{2}{\lambda}\right) Cw = \mathbb{1} \iff C\left(-\frac{2}{\lambda}\right)w = \mathbb{1}$$

The constant factor infront of w does not change its direction but merely its magnitude. Since we rescale w after solving $Cw = \mathbb{1}$ anyways, the solution stays the same

3a) $\frac{\partial}{\partial q_{ij}} \sum_{ij}^{N} p_{ij} \log\left(\frac{p_{ij}}{q_{ij}}\right) = \frac{\partial}{\partial q_{ij}} p_{ij} \log\left(\frac{p_{ij}}{q_{ij}}\right)$

$= \cancel{p_{ij}} \frac{\cancel{q_{ij}}}{\cancel{p_{ij}}} \cdot p_{ij} \cdot \left(-\frac{1}{q_{ij}^2}\right) = \underline{\underline{-\frac{p_{ij}}{q_{ij}}}}$

b) $C = \sum_{ij}^{N} p_{ij} \log\left(\frac{p_{ij} \sum_{kl} \exp(z_{kl})}{\exp(z_{ij})}\right)$

$= \sum_{ij}^{N} p_{ij} \left(\log\left(p_{ij} \sum_{kl} \exp(z_{kl})\right) - z_{ij}\right)$

$\frac{\partial C}{\partial z_{ab}} = \underbrace{\frac{\partial}{\partial z_{ab}}\left(\sum_{ij}^{N} p_{ij} \log\left(p_{ij} \sum_{kl} \exp(z_{kl})\right)\right)}_{} - \underbrace{\frac{\partial}{\partial z_{ab}}\left(\sum_{ij}^{N} p_{ij} z_{ij}\right)}_{}$

$= \frac{\partial}{\partial z_{ab}} (p_{ab} z_{ab}) = \underline{\underline{p_{ab}}}$

$\sum_{ij} p_{ij} \frac{1}{\cancel{p_{ij}} \sum_{kl} \exp(z_{kl})} \cancel{p_{ij}} \exp(z_{ab})$

$= \sum_{ij} p_{ij} q_{ab} = \underline{q_{ab}} \cdot \underbrace{\sum_{ij} p_{ij}}_{=1}$

$\frac{\partial C}{\partial z_{ab}} = q_{ab} - p_{ab}$

c)