P.1 Use the fact that the sum of samples from two Gaussian is again a Gaussian to show why independent component analysis cannot isolate sources perfectly if two or more components are Gaussian.

**Answer:**

Assume $\mathbf{x}$ is a vector of independent Gaussian random variables. A linear mixture of $\mathbf{x}$ gives $\mathbf{y}$:

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

where $\mathbf{y}$ is also a vector of independent Gaussian random variables. The solution of ICA decomposition gives

$$\hat{\mathbf{x}} = \mathbf{B}\mathbf{y}$$

subject to the condition that $E[\hat{\mathbf{x}}\hat{\mathbf{x}}^T] = I$. The condition is due to the fact that $\hat{\mathbf{x}}$ is supposed to be a vector of independent Gaussian random variables. Independent Gaussian random variables have zero cross-correlation (i.e. the off-diagonal components of $E[\hat{\mathbf{x}}\hat{\mathbf{x}}^T]$ is zero). The equal diagonal components of $E[\hat{\mathbf{x}}\hat{\mathbf{x}}^T]$ is due to the fact that the scale of the random variables cannot be recovered and hence their variance are set equal.

In this case, we would see that $\hat{\mathbf{z}} = \mathbf{R}\hat{\mathbf{x}}$ are also an ICA solution as $\hat{\mathbf{z}}$ also satisfies the condition:

$$E[\hat{\mathbf{z}}\hat{\mathbf{z}}^T] = E[\mathbf{R}\hat{\mathbf{x}}\hat{\mathbf{x}}^T\mathbf{R}^T] = \mathbf{R}E[\hat{\mathbf{x}}\hat{\mathbf{x}}^T]\mathbf{R}^T = I$$

where $\mathbf{R}$ is the rotation matrix which has the property that $\mathbf{RR}^T = \mathbf{RR}^{-1} = I$

This means the ICA solution in this case has a rotation ambiguity.

P.2 Let cluster $\mathcal{D}_i$ contain $n_i$ samples, and let $d_{ij}$ be some measure of the distance between two clusters $\mathcal{D}_i$ and $\mathcal{D}_j$. In general, one might expect that if $\mathcal{D}_i$ and $\mathcal{D}_j$ are merged to form a new cluster $\mathcal{D}_k$, then the distance
from $D_k$ to some other cluster $D_h$ is not simply related to $d_{hi}$ and $d_{hj}$.
However, consider the equation

$$d_{hk} = \alpha_i d_{hi} + \alpha_j d_{hj} + \beta d_{ij} + \gamma |d_{hi} - d_{hj}|$$

Show that the following choices for the coefficients $\alpha_i$, $\alpha_j$, $\beta$ and $\gamma$ lead to the distance functions indicated.

(a) $d_{min}$: $\alpha_i = \alpha_j = 0.5$, $\beta = 0$, $\gamma = -0.5$
(b) $d_{max}$: $\alpha_i = \alpha_j = 0.5$, $\beta = 0$, $\gamma = 0.5$

**Answer:**

(a) With the parameter setting, we have

$$d_{hk} = 0.5d_{hi} + 0.5d_{hj} - 0.5|d_{hi} - d_{hj}|$$

If $d_{hi} \leq h_{hj}$,

$$d_{hk} = 0.5d_{hi} + 0.5d_{hj} + 0.5(d_{hi} - d_{hj}) = d_{hi}$$

If $d_{hi} > h_{hj}$

$$d_{hk} = 0.5d_{hi} + 0.5d_{hj} - 0.5(d_{hi} - d_{hj}) = d_{hj}$$

So we can see that this parameter setting leads to $d_{min}$:

$$d_{hk} = \min\{d_{hi}, d_{hj}\}$$

(b) With the parameter setting, we have

$$d_{hk} = 0.5d_{hi} + 0.5d_{hj} + 0.5|d_{hi} - d_{hj}|$$

If $d_{hi} \leq h_{hj}$,

$$d_{hk} = 0.5d_{hi} + 0.5d_{hj} - 0.5(d_{hi} - d_{hj}) = d_{hj}$$

If $d_{hi} > h_{hj}$

$$d_{hk} = 0.5d_{hi} + 0.5d_{hj} + 0.5(d_{hi} - d_{hj}) = d_{hi}$$

So we can see that this parameter setting leads to $d_{max}$:

$$d_{hk} = \max\{d_{hi}, d_{hj}\}$$