10.4

$\sum_{i=1}^m dist_{ij}^2 = tr(\boldsymbol{B}) + mb_{jj}$ [推导]：
$$\begin{aligned}\sum_{i=1}^m dist_{ij}^2 &= \sum_{i=1}^m b_{ii} + \sum_{i=1}^m b_{jj} - 2\sum_{i=1}^m b_{ij} \\ &= tr(B) + mb_{jj}\end{aligned}$$

10.10

$b_{ij} = -\frac{1}{2}(dist_{ij}^2 - dist_{i\cdot}^2 - dist_{\cdot j}^2 + dist_{\cdot\cdot}^2)$ [推导]：由公式（10.3）可得，$b_{ij} = -\frac{1}{2}(dist_{ij}^2 - b_{ii} - b_{jj})$ 由公式（10.6）和（10.9）可得，
$$\begin{aligned}tr(\boldsymbol{B}) &= \frac{1}{2m}\sum_{i=1}^m\sum_{j=1}^m dist_{ij}^2 \\ &= \frac{m}{2}dist_{\cdot\cdot}^2\end{aligned}$$
由公式（10.4）和（10.8）可得，
$$\begin{aligned}b_{jj} &= \frac{1}{m}\sum_{i=1}^m dist_{ij}^2 - \frac{1}{m}tr(\boldsymbol{B}) \\ &= dist_{\cdot j}^2 - \frac{1}{2}dist_{\cdot\cdot}^2\end{aligned}$$
由公式（10.5）和（10.7）可得，
$$\begin{aligned}b_{ii} &= \frac{1}{m}\sum_{j=1}^m dist_{ij}^2 - \frac{1}{m}tr(\boldsymbol{B}) \\ &= dist_{i\cdot}^2 - \frac{1}{2}dist_{\cdot\cdot}^2\end{aligned}$$
综合可得，
$$\begin{aligned}b_{ij} &= -\frac{1}{2}(dist_{ij}^2 - b_{ii} - b_{jj}) \\ &= -\frac{1}{2}(dist_{ij}^2 - dist_{i\cdot}^2 + \frac{1}{2}dist_{\cdot\cdot}^2 - dist_{\cdot j}^2 + \frac{1}{2}dist_{\cdot\cdot}^2) \\ &= -\frac{1}{2}(dist_{ij}^2 - dist_{i\cdot}^2 - dist_{\cdot j}^2 + dist_{\cdot\cdot}^2)\end{aligned}$$

10.14

$$\begin{aligned}\sum_{i=1}^m \|\sum_{j=1}^{d'} z_{ij}\boldsymbol{w}_j - \boldsymbol{x}_i\|_2^2 &= \sum_{i=1}^m \boldsymbol{z}_i^T\boldsymbol{z}_i - 2\sum_{i=1}^m \boldsymbol{z}_i^T\boldsymbol{W}^T\boldsymbol{x}_i + const \\ &\propto -tr(\boldsymbol{W}^T(\sum_{i=1}^m \boldsymbol{x}_i\boldsymbol{x}_i^T)\boldsymbol{W})\end{aligned}$$
[推导]：已知$\boldsymbol{W}^T\boldsymbol{W} = \boldsymbol{I}$和$\boldsymbol{z}_i = \boldsymbol{W}^T\boldsymbol{x}_i$，
$$\begin{aligned}\sum_{i=1}^m \|\sum_{j=1}^{d'} z_{ij}\boldsymbol{w}_j - \boldsymbol{x}_i\|_2^2 &= \sum_{i=1}^m \|\boldsymbol{W}\boldsymbol{z}_i - \boldsymbol{x}_i\|_2^2 \\ &= \sum_{i=1}^m (\boldsymbol{W}\boldsymbol{z}_i)^T(\boldsymbol{W}\boldsymbol{z}_i) - 2\sum_{i=1}^m (\boldsymbol{W}\boldsymbol{z}_i)^T\boldsymbol{x}_i + \sum_{i=1}^m \boldsymbol{x}_i^T\boldsymbol{x}_i \\ &= \sum_{i=1}^m \boldsymbol{z}_i^T\boldsymbol{z}_i - 2\sum_{i=1}^m \boldsymbol{z}_i^T\boldsymbol{W}^T\boldsymbol{x}_i + \sum_{i=1}^m \boldsymbol{x}_i^T\boldsymbol{x}_i \\ &= \sum_{i=1}^m \boldsymbol{z}_i^T\boldsymbol{z}_i - 2\sum_{i=1}^m \boldsymbol{z}_i^T\boldsymbol{z}_i + \sum_{i=1}^m \boldsymbol{x}_i^T\boldsymbol{x}_i \\ &= -\sum_{i=1}^m \boldsymbol{z}_i^T\boldsymbol{z}_i + \sum_{i=1}^m \boldsymbol{x}_i^T\boldsymbol{x}_i \\ &= -tr(\boldsymbol{W}^T(\sum_{i=1}^m \boldsymbol{x}_i\boldsymbol{x}_i^T)\boldsymbol{W}) + \sum_{i=1}^m \boldsymbol{x}_i^T\boldsymbol{x}_i \\ &\propto -tr(\boldsymbol{W}^T(\sum_{i=1}^m \boldsymbol{x}_i\boldsymbol{x}_i^T)\boldsymbol{W})\end{aligned}$$
其中，$\sum_{i=1}^m \boldsymbol{x}_i^T\boldsymbol{x}_i$是常数。

10.17

$$\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{w}_i=\lambda_i\boldsymbol{w}_i$$

[推导]：已知 $\begin{aligned}&\min_{\boldsymbol{W}} -tr(\boldsymbol{W}^T\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{W})\\ &s.t.\ \boldsymbol{W}^T\boldsymbol{W}=\boldsymbol{I}.\end{aligned}$ 运用拉格朗日乘子法可得，$\begin{aligned}J(\boldsymbol{W})&=-tr(\boldsymbol{W}^T\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{W}+\boldsymbol{\lambda}'(\boldsymbol{W}^T\boldsymbol{W}-\boldsymbol{I}))\\ \frac{\partial J(\boldsymbol{W})}{\partial \boldsymbol{W}}&=\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{W}+\boldsymbol{\lambda}'\boldsymbol{W}\end{aligned}$ 令 $\frac{\partial J(\boldsymbol{W})}{\partial \boldsymbol{W}}=\mathbf{0}$，故 $\begin{aligned}\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{W}&=-\boldsymbol{\lambda}'\boldsymbol{W}\\ \boldsymbol{X}\boldsymbol{X}^T\boldsymbol{W}&=\boldsymbol{\lambda}\boldsymbol{W}\end{aligned}$ 其中，$\boldsymbol{W}=\{\boldsymbol{w}_1,\boldsymbol{w}_2,\cdots,\boldsymbol{w}_d\}$和$\boldsymbol{\lambda}=\boldsymbol{diag}(\lambda_1,\lambda_2,\cdots,\lambda_d)$。

10.28

$w_{ij}=\frac{\sum\limits_{k\in Q_i}C_{jk}^{-1}}{\sum\limits_{l,s\in Q_i}C_{ls}^{-1}}$ [推导]：已知 $\begin{aligned}&\min_{\boldsymbol{W}}\sum_{i=1}^m\|\boldsymbol{x}_i-\sum_{j\in Q_i}w_{ij}\boldsymbol{x}_j\|_2^2\\ &s.t.\sum_{j\in Q_i}w_{ij}=1\end{aligned}$ 转换为

$$\begin{aligned}\sum_{i=1}^m\|\boldsymbol{x}_i-\sum_{j\in Q_i}w_{ij}\boldsymbol{x}_j\|_2^2&=\sum_{i=1}^m\|\sum_{j\in Q_i}w_{ij}\boldsymbol{x}_i-\sum_{j\in Q_i}w_{ij}\boldsymbol{x}_j\|_2^2\\ &=\sum_{i=1}^m\|\sum_{j\in Q_i}w_{ij}(\boldsymbol{x}_i-\boldsymbol{x}_j)\|_2^2\\ &=\sum_{i=1}^m\boldsymbol{W}_i^T(\boldsymbol{x}_i-\boldsymbol{x}_j)(\boldsymbol{x}_i-\boldsymbol{x}_j)^T\boldsymbol{W}_i\\ &=\sum_{i=1}^m\boldsymbol{W}_i^T\boldsymbol{C}_i\boldsymbol{W}_i\end{aligned}$$

其中，$\boldsymbol{W}_i=(w_{i1},w_{i2},\cdots,w_{ik})^T$，$k$是$Q_i$集合的长度，$\boldsymbol{C}_i=(\boldsymbol{x}_i-\boldsymbol{x}_j)(\boldsymbol{x}_i-\boldsymbol{x}_j)^T$，$j\in Q_i$。

$$\sum_{j\in Q_i}w_{ij}=\boldsymbol{W}_i^T\mathbf{1}_k=1$$

其中，$\mathbf{1}_k$为k维全1向量。 运用拉格朗日乘子法可得，

$$J(\boldsymbol{W})==\sum_{i=1}^m\boldsymbol{W}_i^T\boldsymbol{C}_i\boldsymbol{W}_i+\lambda(\boldsymbol{W}_i^T\mathbf{1}_k-1)$$

$\frac{\partial J(\boldsymbol{W})}{\partial \boldsymbol{W}_i}=2\boldsymbol{C}_i\boldsymbol{W}_i+\lambda'\mathbf{1}_k$ 令 $\frac{\partial J(\boldsymbol{W})}{\partial \boldsymbol{W}_i}=0$，故 $\begin{aligned}\boldsymbol{W}_i&=-\frac{1}{2}\lambda\boldsymbol{C}_i^{-1}\mathbf{1}_k\\ \boldsymbol{W}_i&=\lambda\boldsymbol{C}_i^{-1}\mathbf{1}_k\end{aligned}$ 其中，λ为一个常数。利用$\boldsymbol{W}_i^T\mathbf{1}_k=1$，对$\boldsymbol{W}_i$归一化，可得

$$\boldsymbol{W}_i=\frac{\boldsymbol{C}_i^{-1}\mathbf{1}_k}{\mathbf{1}_k\boldsymbol{C}_i^{-1}\mathbf{1}_k}$$

10.31

$\min\limits_{\boldsymbol{Z}} tr(\boldsymbol{Z}\boldsymbol{M}\boldsymbol{Z}^T)$
$s.t.\ \boldsymbol{Z}^T\boldsymbol{Z} = \boldsymbol{I}.$

[推导]：

$$
\begin{aligned}
\min_{\boldsymbol{Z}} \sum_{i=1}^{m} \|\boldsymbol{z}_i - \sum_{j\in Q_i} w_{ij}\boldsymbol{z}_j\|_2^2 &= \sum_{i=1}^{m} \|\boldsymbol{Z}\boldsymbol{I}_i - \boldsymbol{Z}\boldsymbol{W}_i\|_2^2 \\
&= \sum_{i=1}^{m} \|\boldsymbol{Z}(\boldsymbol{I}_i - \boldsymbol{W}_i)\|_2^2 \\
&= \sum_{i=1}^{m} (\boldsymbol{Z}(\boldsymbol{I}_i - \boldsymbol{W}_i))^T\boldsymbol{Z}(\boldsymbol{I}_i - \boldsymbol{W}_i) \\
&= \sum_{i=1}^{m} (\boldsymbol{I}_i - \boldsymbol{W}_i)^T\boldsymbol{Z}^T\boldsymbol{Z}(\boldsymbol{I}_i - \boldsymbol{W}_i) \\
&= tr((\boldsymbol{I} - \boldsymbol{W})^T\boldsymbol{Z}^T\boldsymbol{Z}(\boldsymbol{I} - \boldsymbol{W})) \\
&= tr(\boldsymbol{Z}(\boldsymbol{I} - \boldsymbol{W})(\boldsymbol{I} - \boldsymbol{W})^T\boldsymbol{Z}^T) \\
&= tr(\boldsymbol{Z}\boldsymbol{M}\boldsymbol{Z}^T)
\end{aligned}
$$

其中，$\boldsymbol{M} = (\boldsymbol{I} - \boldsymbol{W})(\boldsymbol{I} - \boldsymbol{W})^T$。 [解析]：约束条件$\boldsymbol{Z}^T\boldsymbol{Z} = \boldsymbol{I}$是为了得到标准化（标准正交空间）的低维数据。