4. Classical scaling:

$$\underset{y_1,\dots y_m}{\text{minimize}} \sum_{i=1}^{m}\sum_{j=1}^{n} (\tilde{x}_i^T \tilde{x}_j - y_i^T y_j)^2 \quad \text{where } \tilde{x}_i = x_i - (1/n)\sum_{j=1}^{n} x_j$$

a) The above formulation can be rewritten in matrix form,

$$\min_{Y^TY} \| \tilde{X}^T\tilde{X} - Y^TY \|$$

The matrix $\tilde{X}^T\tilde{X}$ is symmetric, therefore can be applied Eigen Value decomposition.

From Eckart Young Theorem,

$$\min_{\tilde{A}} \| A - \tilde{A} \|^2 \quad \text{s.t} \quad \text{rank}(\hat{A}) = k < r$$

is $\hat{A} = \sum_{i=1}^{k} \sigma_i u_i v_i^T$ $\qquad \hat{A} = U_1 \Sigma_1 V_1^T$

This is the best approximation of A of rank k.

In our problem we use eigen decomposition instead of SVD.

$$Y^TY = U_1 \Lambda_1 U_1^T = (\Lambda_1^{1/2} U_1^T)^T \Lambda_1^{1/2} U_1^T$$

$$Y = \Lambda_1^{1/2} U_1^T$$

$\Lambda_1 =$ First k eigen values of Λ

$U_1 =$ First k eigen vectors of U

b) $D_{ij} = \|x_i - x_j\|^2 = \|x_i\|^2 + \|x_j\|^2 - 2x_i^T x_j$

$$= b1^T + 1b^T - 2X^TX$$

Let $b = \begin{bmatrix} \|x_1\|^2 \\ \|x_2\|^2 \\ \vdots \\ \|x_m\|^2 \end{bmatrix}$

Now, $X^TX = \frac{1}{2}(b1^T + 1b^T - D)$

~~Now~~ Let H be $(I - \frac{1}{m}11^T)$, so, $\tilde{X} = XH$ (Centered data)

Now, $\tilde{X}^T\tilde{X} = H^TX^TXH$

$$= \frac{1}{2}(Hb1^TH + H1b^TH - \underline{HDH}) \quad \begin{bmatrix} 1^TH = 0 \\ H^T1 = 0 \end{bmatrix}$$

$$= -\frac{1}{2}(I - \frac{1}{m}11^T) D (I - \frac{1}{m}11^T)$$

$$= -\frac{1}{2}B.$$