STAT 745 – Fall 2014
Assignment 10

Group 1: Francene Cicia, Doug Raffle, Melissa Smith

December 1, 2014

1 Multiple Dimensional Scaling

1.a Outer vs. Inner Product Eigenvalues

First, we can use a singular values decomposition to decompose X as $X = U\Sigma V^T$, where U and V are $n \times n$ and $p \times p$ orthogonal matrices, respectively, and Σ is an $n \times p$ diagonal matrix such that:

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & & \cdots & & & 0 \\ 0 & \sigma_2 & & & & & \\ & & \ddots & & & & \\ \vdots & & & \sigma_r & & & \vdots \\ & & & & 0 & & \\ & & & & & \ddots & \\ 0 & & & \cdots & & & 0 \end{bmatrix}$$

where r is the rank of X. Using this, we have that:

$$\begin{aligned} XX^T &= U\Sigma V^T \left(U\Sigma V^T\right)^T \\ &= U\Sigma V^T V \Sigma^T U^T \\ &= U\Sigma\Sigma^T U^T \end{aligned}$$

Which is the spectral decomposition of XX^T, so $\Sigma\Sigma^T$ is an $n \times n$ diagonal matrix of the eigenvalues of XX^T. Similarly,

$$\begin{aligned} X^T X &= \left(U\Sigma V^T\right)^T U\Sigma V^T \\ &= V\Sigma^T U^T U \Sigma V^T \\ &= V\Sigma^T \Sigma V^T \end{aligned}$$

where $\Sigma^T\Sigma$ is the $p \times p$ diagonal matrix of the eigenvalues of $X^T X$. If we assume $n > p$ and the columns of X are linearly independent, then $\text{rank}\,(X) = r = p$. We then know that Σ is an $n \times p$ matrix of the form:

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_p \\ 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{bmatrix}$$

From here, the fact that Σ is diagonal lets us easily find the eigenvalues of the inner and outer products of X.

For the outer product, the eigenvalues are:

$$\Sigma\Sigma^T = \begin{bmatrix} \sigma_1^2 & 0 & & \cdots & & & 0 \\ 0 & \sigma_2^2 & & & & & \\ & & \ddots & & & & \\ \vdots & & & \sigma_p^2 & & & \vdots \\ & & & & 0 & & \\ & & & & & \ddots & \\ 0 & & & \cdots & & & 0 \end{bmatrix}$$

Where there are $n-p$ zeros on the diagonal. For X^TX, we have eigenvalues of:

$$\Sigma^T\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_p^2 \end{bmatrix}$$

So we see that the first p eigenvalues of XX^T are the same as the p eigenvalues of X^TX, and the remaining $n-p$ eigenvalues of XX^T are zeros.

1.b Distance Matrix

If we have an $n \times p$ matrix X, the outer product is:

$$XX^T = \begin{bmatrix} \sum_{j=1}^p x_{1j}^2 & \sum_{j=1}^p x_{1j}x_{2j} & \cdots & \sum_{j=1}^p x_{1j}x_{nj} \\ \sum_{j=1}^p x_{2j}x_{1j} & \sum_{j=1}^p x_{2j}^2 & \cdots & \sum_{j=1}^p x_{2j}x_{nj} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^p x_{nj}x_{1j} & \sum_{j=1}^p x_{nj}x_{2j} & \cdots & \sum_{j=1}^p x_{nj}^2 \end{bmatrix} = \begin{bmatrix} x_1^Tx_1 & x_1^Tx_2 & \cdots & x_1^Tx_n \\ x_2^Tx_{1j} & x_2^Tx_2 & \cdots & x_2^Tx_n \\ \vdots & \vdots & \ddots & \vdots \\ x_n^Tx_1 & x_n^Tx_2 & \cdots & x_n^Tx_n \end{bmatrix}$$

And the negative squared Euclidean distances can be expressed as:

$$\begin{aligned} -D_{ij}^2 &= -\sum_{k=1}^p (x_{ik} - x_{jk})^2 \\ &= -\sum_{k=1}^p \left(x_{ik}^2 + x_{jk}^2 - 2x_{ik}x_{jk}\right) \\ &= -\sum_{k=1}^p x_{ik}^2 - \sum_{k=1}^p x_{jk}^2 + 2\sum_{k=1}^p x_{ik}x_{jk} \\ &= -x_i^Tx_i - x_j^Tx_j + 2x_i^Tx_j \end{aligned}$$

Note that each of these terms can be written as a cell from the outer product:

$$-D_{ij}^2 = -\left(XX^T\right)_{ii} - \left(XX^T\right)_{jj} + 2\left(XX^T\right)_{ij}$$

$$-\frac{D_{ij}^2}{2} = -\frac{\left(XX^T\right)_{ii} + \left(XX^T\right)_{jj}}{2} + \left(XX^T\right)_{ij}$$

We can express this for the entire matrix XX^T, where we define V such that $V_{ij} = -\left(x_i^Tx_i + x_j^Tx_j\right)/2$.

$$-\frac{D^2}{2} = V + XX^T$$

1.c Centering

X is defined here as being centered, which implies that there is some uncentered matrix X^*, such that $C_n X^* = X$. From here, the fact that C_n is idempotent tells us:

$$C_n X = C_n \left(C_n X^* \right) = C_n X^* = X$$

Intuitively, we are column-centering an already column-centered matrix, which means we are substracting zero from every element in each column. By the same logic, X^T is row centered, so $X^T C_n = X^T$. Which means:

$$C_n X X^T C_n = X X^T$$

Since the matrices are identical, their eigenvalues are equivalent.

Now, we prove that $C_n V C_n = \vec{0}$. From above, we have that $V_{ij} = -(x_i^T x_i + x_j^T x_j)/2$, giving us the matrix:

$$V = -\frac{1}{2} \begin{bmatrix} x_1^T x_1 + x_1^T x_1 & x_1^T x_1 + x_2^T x_2 & \cdots & x_1^T x_1 + x_p^T x_p \\ x_2^T x_2 + x_2^T x_2 & x_2^T x_2 + x_2^T x_2 & \cdots & x_2^T x_2 + x_p^T x_p \\ \vdots & \vdots & \ddots & \vdots \\ x_n^T x_n + x_1^T x_1 & x_n^T x_n + x_2^T x_2 & \cdots & x_n^T x_n + x_p^T x_p \end{bmatrix}$$

$$= -\frac{1}{2} \left(\begin{bmatrix} x_1^T x_1 & \cdots & x_1^T x_1 \\ \vdots & \ddots & \vdots \\ x_n^T x_n & \cdots & x_n^T x_n \end{bmatrix} + \begin{bmatrix} x_1^T x_1 & \cdots & x_p^T x_p \\ \vdots & \ddots & \vdots \\ x_1^T x_1 & \cdots & x_p^T x_p \end{bmatrix} \right)$$

When we double center this matrix, we get:

$$C_n V C_n = -\frac{1}{2} C_n \left(\begin{bmatrix} x_1^T x_1 & \cdots & x_1^T x_1 \\ \vdots & \ddots & \vdots \\ x_n^T x_n & \cdots & x_n^T x_n \end{bmatrix} + \begin{bmatrix} x_1^T x_1 & \cdots & x_p^T x_p \\ \vdots & \ddots & \vdots \\ x_1^T x_1 & \cdots & x_p^T x_p \end{bmatrix} \right) C_n$$

$$= -\frac{1}{2} \left(C_n \begin{bmatrix} x_1^T x_1 & \cdots & x_1^T x_1 \\ \vdots & \ddots & \vdots \\ x_n^T x_n & \cdots & x_n^T x_n \end{bmatrix} C_n + C_n \begin{bmatrix} x_1^T x_1 & \cdots & x_p^T x_p \\ \vdots & \ddots & \vdots \\ x_1^T x_1 & \cdots & x_p^T x_p \end{bmatrix} C_n \right)$$

For convenience, we define the matrices A and B such that:

$$C_n V C_n = -\frac{1}{2} \left(C_n A C_n + C_n B C_n \right)$$

Note that all of the elements in a row of A are identical, and that all of the elements of a column of B are identical. This means each element of a row in A is the row mean, and each element of a column of B is the column mean. For a matrix X, XC_n row-centers and $C_n X$ column-centers, so $AC_n = C_n B = \vec{0}$, hence:

$$C_n V C_n = -\frac{1}{2} \left(C_n \vec{0} + \vec{0} C_n \right) = -\frac{1}{2} \left(\vec{0} + \vec{0} \right) = -\frac{1}{2} \left(\vec{0} \right) = \vec{0}$$

1.d PCA and MDS Eigenvalues

In PCA, we find the ordered eigenvalues of $X^T X$ such that $\lambda_{(1)} \geq \lambda_{(2)} \geq \cdots \geq \lambda_{(p)}$. Above, we showed that these p eigenvalues are equivalent to the first p eigenvalues of M, and the remaining $n - p$ are zero.

1.e PCA and MDS Preservation

Where PCA preserves the variability of the dimensions, MDS preserves the distances between observations.