CMSE 820 Homework 6. Due November 3, 2020.

I. (OMITTED) Let $\sigma \neq 0$ be a real number and $x^{(1)}, \ldots, x^{(n)} \in \mathbb{R}^p$ be n distinct sample points. Prove that the matrix $K \in \mathbb{R}^{n \times n}$ defined by

$$K_{ij} = \exp\left(-\frac{\|x^{(i)} - x^{(j)}\|^2}{2\sigma^2}\right)$$

has full rank.

Solution. Since the n sample points are distinct from one another, it follows that $\|x^{(i)} - x^{(j)}\|^2 = 0$ only for $j = i$. Therefore, $K_{ii} = 1$ and $K_{ij} \neq 1$ for $j \neq i$. Additionally, $\forall i, j$, we know that $\exp\left(-\frac{\|x^{(i)}-x^{(j)}\|^2}{2\sigma^2}\right) > 0$ from the range of $\exp(\bullet)$. The matrix K must be symmetric because the distance from $x^{(i)}$ to $x^{(j)}$ is the equal to the distance from $x^{(j)}$ to $x^{(i)}$. Thus, K is a symmetric matrix with all positive entries and 1's along the entire diagonal. Furthermore, since the sample point are all distinct, it follows that $\|x^{(i)} - x^{(j)}\|^2 > \|x^{(i)} - x^{(i)}\|^2 = 0$ for any $j \neq i$. Thus,

$$\exp\left(-\frac{\|x^{(i)} - x^{(j)}\|^2}{2\sigma^2}\right) < \exp\left(-\frac{\|x^{(i)} - x^{(i)}\|^2}{2\sigma^2}\right) = 1,$$

for any $j \neq i \Rightarrow 0 < K_{ij} < 1$ for $j \neq i$.

We now use mathematical induction. Consider the base case for $n = 2$ where we have 2 distinct samples in $\mathbb{R}^p$. Let $K_{12} = K_{21} = k_{12} \in (0, 1)$ as per the properties of K outlined above. We can write the matrix K as

$$K = \begin{pmatrix} 1 & k_{12} \\ k_{12} & 1 \end{pmatrix}.$$

If we calculate the determinant of K for $n = 2$, we see $\det(K) = 1 - k_{12}^2 > 0$. Since the determinant is nonzero, K must be invertible and therefore has full rank. Alternatively, if we examine the eigenvalues of $K \in \mathbb{R}^{2\times 2}$, we see $\lambda_1 = 1 - k_{12}, \lambda_2 = 1 + k_{12}$. Since both eigenvalues are positive and nonzero, we can conclude that K is positive definite and thus has full rank.

Suppose we consider one more distinct sample $x^{(3)} \in \mathbb{R}^p$ $(n = 3)$. We write K as

$$K = \begin{pmatrix} 1 & k_{12} & k_{13} \\ k_{12} & 1 & k_{23} \\ k_{13} & k_{23} & 1 \end{pmatrix},$$

where $k_{13}, k_{23} \in (0, 1)$ as per the properties of K stated above. Consider the following,

$$c_1 \begin{pmatrix} 1 \\ k_{12} \\ k_{13} \end{pmatrix} + c_2 \begin{pmatrix} k_{12} \\ 1 \\ k_{23} \end{pmatrix} + c_3 \begin{pmatrix} k_{13} \\ k_{23} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\begin{aligned} \Rightarrow c_1 + c_2 k_{12} + c_3 k_{13} &= 0 \\ c_1 k_{12} + c_2 + c_3 k_{23} &= 0 \\ c_1 k_{13} + c_2 k_{23} + c_3 &= 0 \end{aligned}$$

where $c_1, c_2, c_3 \in \mathbb{R}$. Solving for c_1, c_2, c_3, (omitting some algebraic steps, which I am happy to submit if need be)

$$\begin{aligned}
\Rightarrow & c_1 = -c_2k_{12} - c_3k_{13} \\
& (-c_2k_{12} - c_3k_{13})k_{12} + c_2 + c_3k_{23} = 0 \\
\Rightarrow & c_2 = \frac{c_3k_{13}k_{12} - c_3k_{23}}{1 - k_{12}^2} \\
& -c_2k_{12}k_{13} - c_3k_{13}^2 + c_)2k_{23} + c_3 = 0 \\
& (1 - k_{13}^2)c_3 + (k_{23} - k_{12}k_{13})\left(\frac{c_3k_{13}k_{12} - c_3k_{23}}{1 - k_{12}^2}\right) = 0 \\
& (1 - k_{13}^2)(1 - k_{12}^2)c_3 + (c_3k_{13}k_{12}k_{23} - c_3k_{13}^2k_{12}^2 - c_3k_{23}^2 + c_3k_{23}k_{12}k_{13}) = 0 \\
& c_3(2k_{12}k_{13}k_{23} - k_{12}^2 - k_{13}^2 - k_{23}^2) = 0 \\
\Rightarrow & c_3 = 0 \\
\Rightarrow & c_2 = 0 \\
\Rightarrow & c_1 = 0.
\end{aligned}$$

The columns of $K \in \mathbb{R}^{3\times 3}$ are linearly independent and thus K has full rank.

Suppose now the induction hypothesis holds true and $K \in \mathbb{R}^{n\times n}$ for n distinct samples in $\mathbb{R}^p$ has full rank. Let $\tilde{K} \in \mathbb{R}^{n+1\times n+1}$ be the matrix K after we include one additional sample $x^{(n+1)}$ that is distinct from all $x^{(1)}, \ldots, x^{(n)}$. We can write $\tilde{K}$ as

$$\tilde{K} = \begin{pmatrix} & & & k_{1,n+1} \\ & K & & \vdots \\ & & & k_{n,n+1} \\ k_{1,n+1} & \ldots & k_{n,n+1} & 1 \end{pmatrix},$$

where $k_{1,n+1}, \ldots, k_{n,n+1} \in (0,1)$ satisfy the properties stated above. We can exchange two rows or two columns in $\tilde{K}$ which corresponds to changing the sign of the determinant but not the value of the determinant. Since there is a 1 at the $n+1$th diagonal, we can use Gaussian elimination to turn the elements in the $n+1$th column to 0 in order to isolate the pivot. Therefore, we can rewrite this matrix equivalently by letting

$$\tilde{K}' = \begin{pmatrix} 1 & k_{1,n+1} & \ldots & k_{n,n+1} \\ 0 & & & \\ \vdots & & K' & \\ 0 & & & \end{pmatrix},$$

where K' represents the matrix K after performing Gaussian elimination (*Note:* this is problematic because Gaussian Elimination *can* effect K such that the determinant is not guaranteed to be nonzero because we are using a row not in K to perform the elimination). Thus, $\det(\tilde{K}') = 1\det(K') = \pm\det(\tilde{K})$. Since $\det(K) \neq 0$ because K has full rank, it follows that $\det(K') \neq 0$ because Gaussian elimination entails only elementary row operations which preserve the determinant, and thus $\det(\tilde{K}) \neq 0$. We conclude that $\tilde{K}$ must have full rank. ■

II. Let $D \in \mathbb{R}^{n\times n}$ be a real symmetric matrix. Prove that D is a square distance matrix (that is, there exist sample points $x^{(1)}, \ldots, x^{(n)} \in \mathbb{R}^p$ such that $D_{ij} = \|x^{(i)} - x^{(j)}\|^2$) if and only if

$$D_{ij} = K_{ii} + K_{jj} - 2K_{ij}$$

for some symmetric positive semi-definite matrix $K \in \mathbb{R}^{n\times n}$.

Solution. Suppose that D is a square distance matrix, that is, there exist sample points $x^{(1)},\dots,x^{(n)} \in \mathbb{R}^p$ such that $D_{ij} = \|x^{(i)} - x^{(j)}\|^2$. Thus, $D_{ij} = (x^{(i)} - x^{(j)})^T(x^{(i)} - x^{(j)}) = (x^{(i)})^2 + (x^{(j)})^2 - 2x^{(i)}x^{(j)}$. Let $K_{ij} = x^{(i)} \cdot x^{(j)}$. We see that K must be symmetric because the dot product is symmetric. We can write $K = X^TX$ where $X \in \mathbb{R}^{p\times n}$ is the sample matrix formed by letting the sample points $x^{(1)},\dots,x^{(n)}$ be the columns of X. For arbitrary nonzero vector $v \in \mathbb{R}^n$,

$$\begin{aligned} v^TKv &= v^TX^TXv \\ &= (Xv)^T(Xv) \geq 0, \end{aligned}$$

so K is positive semi-definite. Thus, we can write $D_{ij} = K_{ii} + K_{jj} - 2K_{ij}$ where K is a symmetric positive semi-definite matrix.

Suppose instead that $D_{ij} = K_{ii} + K_{jj} - 2K_{ij}$ for some symmetric positive semi-definite matrix $K \in \mathbb{R}^{n\times n}$. Since K is symmetric positive semi-definite, it follows that there exists a matrix $X \in \mathbb{R}^{n\times n}$ such that $K = X^TX$. From this, we can write

$$K_{ij} = \sum_{k=1}^{n} X_{ik}X_{kj} = \sum_{k=1}^{n} X_{ik}X_{jk},$$

and

$$D_{ij} = \sum_{k=1}^{n} X_{ik}X_{ik} + X_{jk}X_{jk} - 2X_{ik}X_{jk} = \sum_{k=1}^{n}(X_{ik} - X_{jk})^2.$$

From this expression, we see that the matrix D must be a square distance matrix. ■

III. Let $H = I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T$ be the centering matrix where I_n is the $n \times n$ identity matrix and $\mathbf{1} = (1, 1, \dots, 1)^T \in \mathbb{R}^n$. A real symmetric matrix $D \in \mathbb{R}^{n\times n}$ is called conditionally negative semi-definite if (and only if)

$$u^TDu \leq 0$$

for any $u \in \text{Ran}(H)$, where $\text{Ran}(H)$ is the range of H (see HW5, 1.b.). Prove that a real symmetric matrix D is conditionally negative semi-definite if and only if $-\frac{1}{2}HDH$ is positive semi-definite.

Solution. We know that $\text{Ran}(H) = \{u \in \mathbb{R}^n : u^T\mathbf{1} = 0\}$. Suppose a real symmetric matrix D is conditionally negative semi-definite. Let $x \in \mathbb{R}^n$ be an arbitrary nonzero vector and $B = -\frac{1}{2}HDH$.

$$\begin{aligned} x^TBx &= x^T(-\frac{1}{2}HDH)x \\ &= -\frac{1}{2}x^THDHx \\ &= -\frac{1}{2}x^THDH^Tx \\ &= -\frac{1}{2}(H^Tx)^TD(H^Tx). \end{aligned}$$

Let $y = H^Tx$. It follows that $\mathbf{1}^T \cdot y = \mathbf{1}^T \cdot (H^Tx) = \mathbf{1}^T \cdot (x - \frac{1}{n}\mathbf{1}\mathbf{1}^Tx) = \mathbf{1}^T \cdot x - \mathbf{1}^T \cdot x = 0$. Therefore, $y \in \text{Ran}(H)$. Since D is conditionally negative semi-definite, then $y^TDy \leq 0 \Rightarrow -\frac{1}{2}y^TDy \geq 0$. Thus, $x^TBx \geq 0$ for arbitrary choice of x and B is positive semi-definite.

Suppose now instead that $B = -\frac{1}{2}HDH$ is positive semi-definite. For any $u \in \text{Ran}(H)$, $Hu = H^T u = (I - \frac{1}{n}\mathbf{1}\mathbf{1}^T)u = u$. Therefore,

$$\begin{aligned} u^T D u &= (H^T u)^T D (H^T u) \\ &= u^T H D H u. \end{aligned}$$

Since B is positive semi-definite, $x^T(-\frac{1}{2}HDH)x \geq 0, \forall x$, and thus, $x^T(HDH)x \leq 0, \forall x$. It follows that $u^T D u \leq 0$, which implies D is conditionally negative semi-definite. ■

IV. Let $D \in \mathbb{R}^{n \times n}$ be a real symmetric matrix and $d \in \mathbb{R}^n$ be the diagonal of D as a vector. Define

$$C = D - \frac{1}{2}d\mathbf{1}^T - \frac{1}{2}\mathbf{1}d^T.$$

(a) Prove that $-\frac{1}{2}HDH = -\frac{1}{2}HCH$.

Solution.

$$\begin{aligned} HCH &= H\left(D - \frac{1}{2}d1^T - \frac{1}{2}1d^T\right)H \\ &= \left(I - \frac{1}{n}11^T\right)\left(D - \frac{1}{2}d1^T - \frac{1}{2}1d^T\right)\left(I - \frac{1}{n}11^T\right) \\ &= \left(D - \frac{1}{2}d1^T - \frac{1}{2}1d^T - \frac{1}{n}11^TD + \frac{1}{2n}11^Td1^T + \frac{1}{2n}11^T1d^T\right)\left(I - \frac{1}{n}11^T\right) \\ &= \left(D - \frac{1}{2}d1^T - \frac{1}{2}1d^T - \frac{1}{n}11^TD + \frac{1}{2n}11^Td1^T + \frac{1}{2}1d^T\right)\left(I - \frac{1}{n}11^T\right) \\ &= \left(D - \frac{1}{2}d1^T - \frac{1}{n}11^TD + \frac{1}{2n}11^Td1^T\right)\left(I - \frac{1}{n}11^T\right) \\ &= \left(D - \frac{1}{2}d1^T - \frac{1}{n}11^TD + \frac{1}{2n}11^Td1^T - \frac{1}{n}D11^T + \frac{1}{2n}d1^T11^T + \frac{1}{n^2}11^TD11^T - \frac{1}{2n^2}11^Td1^T11^T\right) \\ &= \left(D - \frac{1}{2}d1^T - \frac{1}{n}11^TD + \frac{1}{2n}11^Td1^T - \frac{1}{n}D11^T + \frac{1}{2}d1^T + \frac{1}{n^2}11^TD11^T - \frac{1}{2n}11^Td1^T\right) \\ &= \left(D - \frac{1}{n}11^TD - \frac{1}{n}D11^T + \frac{1}{n^2}11^TD11^T\right) \\ &= D - \frac{1}{n}D11^T - \left(\frac{1}{n}11^TD - \frac{1}{n^2}11^TD11^T\right) \\ &= D\left(I - \frac{1}{n}11^T\right) - \frac{1}{n}11^TD\left(I - \frac{1}{n}11^T\right) \\ &= \left(D - \frac{1}{n}11^TD\right)\left(I - \frac{1}{n}11^T\right) \\ &= \left(I - \frac{1}{n}11^T\right)D\left(I - \frac{1}{n}11^T\right) = HDH. \end{aligned}$$

Since $HCH = HDH$, it follows that $-\frac{1}{2}HCH = -\frac{1}{2}HDH$. ■

(b) Denote $B = -\frac{1}{2}HCH$. Prove that

$$C_{ij} = B_{ii} + B_{jj} - 2B_{ij}$$

where C_{ij} and B_{ij} are the (i,j)-entry of C and B, respectively.

Solution.

$$\begin{aligned}
B &= -\frac{1}{2}HCH \\
&= -\frac{1}{2}\left(I - \frac{1}{n}\mathbf{1}\mathbf{1}^T\right)C\left(I - \frac{1}{n}\mathbf{1}\mathbf{1}^T\right) \\
&= -\frac{1}{2}\left(C - \frac{1}{n}\mathbf{1}\mathbf{1}^TC - \frac{1}{n}C\mathbf{1}\mathbf{1}^T + \frac{1}{n^2}\mathbf{1}\mathbf{1}^TC\mathbf{1}\mathbf{1}^T\right). \\
\Rightarrow -2B &= C - \frac{1}{n}\mathbf{1}\mathbf{1}^TC - \frac{1}{n}C\mathbf{1}\mathbf{1}^T + \frac{1}{n^2}\mathbf{1}\mathbf{1}^TC\mathbf{1}\mathbf{1}^T.
\end{aligned}$$

We can calculate $\mathbf{1}\mathbf{1}^TC$ and $C\mathbf{1}\mathbf{1}^T$ as

$$\begin{aligned}
\mathbf{1}\mathbf{1}^TC &= \begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}\begin{pmatrix}1 & 1 & \dots & 1\end{pmatrix}\begin{pmatrix}c_{11} & c_{12} & \dots & c_{1n}\\ c_{21} & c_{22} & \dots & c_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ c_{n1} & c_{n2} & \dots & c_{nn}\end{pmatrix} \\
&= \begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}\begin{pmatrix}\sum_{j=1}^n c_{j1} & \sum_{j=1}^n c_{j2} & \dots & \sum_{j=1}^n c_{jn}\end{pmatrix} \\
&= \begin{pmatrix}\sum_{j=1}^n c_{j1} & \sum_{j=1}^n c_{j2} & \dots & \sum_{j=1}^n c_{jn}\\ \sum_{j=1}^n c_{j1} & \sum_{j=1}^n c_{j2} & \dots & \sum_{j=1}^n c_{jn}\\ \vdots & \vdots & \ddots & \vdots\\ \sum_{j=1}^n c_{j1} & \sum_{j=1}^n c_{j2} & \dots & \sum_{j=1}^n c_{jn}\end{pmatrix}. \\
C\mathbf{1}\mathbf{1}^T &= (\mathbf{1}\mathbf{1}^TC)^T \\
&= \begin{pmatrix}\sum_{j=1}^n c_{1j} & \sum_{j=1}^n c_{1j} & \dots & \sum_{j=1}^n c_{1j}\\ \sum_{j=1}^n c_{2j} & \sum_{j=1}^n c_{2j} & \dots & \sum_{j=1}^n c_{2j}\\ \vdots & \vdots & \ddots & \vdots\\ \sum_{j=1}^n c_{nj} & \sum_{j=1}^n c_{nj} & \dots & \sum_{j=1}^n c_{nj}\end{pmatrix}.
\end{aligned}$$

Additionally,

$$\begin{aligned}
\mathbf{1}\mathbf{1}^TC\mathbf{1}\mathbf{1}^T &= \mathbf{1}\mathbf{1}^T\begin{pmatrix}\sum_{j=1}^n c_{1j} & \sum_{j=1}^n c_{1j} & \dots & \sum_{j=1}^n c_{1j}\\ \sum_{j=1}^n c_{2j} & \sum_{j=1}^n c_{2j} & \dots & \sum_{j=1}^n c_{2j}\\ \vdots & \vdots & \ddots & \vdots\\ \sum_{j=1}^n c_{nj} & \sum_{j=1}^n c_{nj} & \dots & \sum_{j=1}^n c_{nj}\end{pmatrix} \\
\Rightarrow (\mathbf{1}\mathbf{1}^TC\mathbf{1}\mathbf{1}^T)_{ij} &= \sum_{k=1}^n\sum_{l=1}^n c_{kl},
\end{aligned}$$

which is just the sum of all entries in the matrix C. Therefore,

$$-2B_{ij} = C_{ij} - \frac{1}{n}\sum_{k=1}^{n} c_{kj} - \frac{1}{n}\sum_{k=1}^{n} c_{ik} + \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n} c_{kl}.$$

It follows that

$$-2B_{ii} = C_{ii} - \frac{1}{n}\sum_{k=1}^{n} c_{ki} - \frac{1}{n}\sum_{k=1}^{n} c_{ik} + \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n} c_{kl},$$
$$-2B_{jj} = C_{jj} - \frac{1}{n}\sum_{k=1}^{n} c_{kj} - \frac{1}{n}\sum_{k=1}^{n} c_{jk} + \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n} c_{kl},$$

and thus,

$$\begin{aligned} B_{ii} + B_{jj} - 2B_{ij} &= -\frac{1}{2}C_{ii} + \frac{1}{2n}\sum_{k=1}^{n} c_{ki} + \frac{1}{2n}\sum_{k=1}^{n} c_{ik} - \frac{1}{2}C_{jj} + \frac{1}{2n}\sum_{k=1}^{n} c_{kj} + \frac{1}{2n}\sum_{k=1}^{n} c_{jk} + \\ &\quad C_{ij} - \frac{1}{n}\sum_{k=1}^{n} c_{kj} - \frac{1}{n}\sum_{k=1}^{n} c_{ik} \\ &= -\frac{1}{2}C_{ii} - \frac{1}{2}C_{jj} + C_{ij} - \frac{1}{2n}\sum_{k=1}^{n} c_{ik} - \frac{1}{2n}\sum_{k=1}^{n} c_{kj} + \frac{1}{2n}\sum_{k=1}^{n} c_{ki} + \frac{1}{2n}\sum_{k=1}^{n} c_{jk}. \end{aligned}$$

Since $C = D - \frac{1}{2}d\mathbf{1}^T - \frac{1}{2}\mathbf{1}d^T$ and D is a real symmetric matrix, this further simplifies to

$$B_{ii} + B_{jj} - 2B_{ij} = -\frac{1}{2}C_{ii} - \frac{1}{2}C_{jj} + C_{ij}.$$

We know that

$$C_{ii} = D_{ii} - \frac{1}{2}D_{ii} - \frac{1}{2}D_{ii} = 0,$$

and thus,

$$B_{ii} + B_{jj} - 2B_{ij} = C_{ij}. \blacksquare$$

(c) If D is conditionally negative semi-definite, prove that C is a square distance matrix.

Solution. If D is conditionally negative semi-definite, then $B = -\frac{1}{2}HCH = -\frac{1}{2}HDH$ is positive semi-definite (by III).

Since B is positive semi-definite, it follows that there exists an $X \in \mathbb{R}^{n\times n}$ such that $B = X^T X$. The matrix B must also be symmetric. Since the elements of C can be written in terms of a symmetric positive semi-definite matrix B as in II, it follows that C must be a square distance matrix. $\blacksquare$

V. See attached jupyter notebook for source code and discussion.