Exercise 1

מכי פרל
031119035

Question 1

(a) Suppose $A = \in \mathbb{R}^{n\times n}$ a symmetric matrix with $(\lambda_1, u_1), (\lambda_2, u_2)$ pairs of eigenvalue-eigenvector of T s.t. $\lambda_1 \neq \lambda_2$. Then $\exists T \in L(V)$ with V the standard inner-product space s.t. T is self-adjoint and $A = M(T)$ relative to the standard base. According to the spectral theorem, T is diagonalizable w.r.t. some orthonormal basis $e_1, .., e_n$. This implies that eigenspaces of different eigenvalues are orthogonal to one another. Therefore $\langle u_1, u_2\rangle = 0 \rightarrow u_1^T u_2 = 0$.

(b) Let $\theta \in \mathbb{R}$ and $B = I + \theta A$. $Bu_1 = Iu_1 + \theta A u_1 = u_1 + \theta\lambda_1 u_1 = (1+\theta\lambda_1)u_1$, meaning that u_1 is an eigenvector of B, with $1+\theta\lambda_1$ its eigvalue.

(c) Define $U = [u_1, ..., u_n]$ where $u_1, .., u_n$ are A's eigenvectors normalized, constituting an orthonormal basis of $\mathbb{R}^n$. Then A can be decomposed as $U\Lambda U^T$ with Λ being an $n\times n$ real diagonal matrix. Since U is orthogonal $U^{-1} = U^T$ and $AA^{-1} = I \rightarrow$ $U\Lambda U^T A^{-1} = I \rightarrow A^{-1} = U\Lambda^{-1}U^T = [u_1, .., u_n]\begin{bmatrix} 1/\lambda_1 & & 0 \\ & \ddots & \\ 0 & & 1/\lambda_n \end{bmatrix}[u_1, .., u_n]^T$, yielding the spectral decomposition $A^{-1} = \frac{1}{\lambda_1}u_1u_1^T + ... + \frac{1}{\lambda_n}u_nu_n^T$, under the assumption that $\lambda_1, .., \lambda_n \neq 0$ (A is invertible).

Question 2

(a) Suppose $X \in \mathbb{R}^{n\times p}$ s.t. $n > p$ and $A = X^T X$, then $(X^T X)^T = X^T X$, meaning that A is a $p\times p$ symmetric square matrix. Hence $\exists T \in (V)$ with $A = M(T)$ where V is the standard inner product space over the reals, for which $\langle T^*Tv, v\rangle = \langle Tv, Tv\rangle \geq 0$ $\forall v \in V$, and T^*T is a positive operator. In addition, if $T^*Tu = \lambda u$ for some $u \in V$ $\lambda \in \mathbb{R}$ s.t. $u \neq 0$ then $\lambda\langle u, u\rangle = \langle \lambda u, u\rangle = \langle T^*Tu, u\rangle = \langle Tu, Tu\rangle \geq 0$, implying that A's eigenvalues are non-negative.

Assume (1) that A is invertible, then $\text{Col}\, A$ span $\mathbb{R}^p$, meaning that A's columns are linearly independent. Assume (2) that A's columns are linearly independent, then $\text{rank}\, A = p \rightarrow \dim \text{null}\, A = 0$ and $\forall v \in V$ s.t. $v \neq 0$ $T^*Tv \neq 0 \rightarrow \langle T^*Tv, v\rangle > 0$, making A a positive definite matrix. Assume now (3) that A is positive definite and $T^*Tv = \lambda v$ for some $v \in V$ $\lambda \in \mathbb{R}$ s.t. $v \neq 0$, then $\langle T^*Tv, v\rangle > 0 \rightarrow \langle \lambda v, v\rangle > 0 \rightarrow \lambda\langle v, v\rangle > 0 \rightarrow \lambda > 0$ and T's eigenvalues are positive. Assume (4) that T has p positive eigenvalues then $\text{null}\, T = \{0\} \rightarrow \text{rank}\, A = p$ and A is invertible.

(b) Let $\theta \in \mathbb{R}$ s.t. $\theta > 0$ then $B = A + \theta I$ is symmetric. Let $\lambda \in \mathbb{R}$ $v \in V$ $v \neq 0$ s.t. $Av = \lambda v$ then $(A + \theta I)v = Av + \theta v = (\lambda + \theta)v > 0$, implying that B is positive definite. Hence, B is invertible based on (a).

Question 3

(a) Suppose $A = \begin{bmatrix} 1 & -2 & 1 \\ 3 & -1 & 0 \\ 1 & 4 & -2 \end{bmatrix}$ and $S = \{u_1, u_2, u_3\} = \left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\2\\3\end{bmatrix}\right\}$ and define $B = [u_1, u_2, u_3]$. Calculating B^{-1} we get:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ 1 & 1 & 3 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 1 & 2 & -1 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & -1 \\ 0 & 1 & 0 & -1 & 2 & -1 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right] \text{ and } B^{-1} = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}.$$

Therefore $[A]_S = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ 3 & -1 & 0 \\ 1 & 4 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} = \begin{bmatrix} -1 & -4 & -2 \\ 1 & -3 & -1 \\ 1 & 3 & 2 \end{bmatrix}$.

(b) Suppose $T \in \mathcal{L}(V, W)$, $F \in \mathcal{L}(W, U)$ and B, C, D are bases of V, W, U respectively. $[F \circ T(v)]_D = [F(T(v))]_D = (D^{-1}M(F)C)(C^{-1}M(T)B)(B^{-1}v) = D^{-1}M(F)M(T)v$ where $M(T), M(F)$, and v are w.r.t. to the standard basis. This shows that matrix multiplication can represent composition of linear maps.

Suppose $T \in (V)$, $A = M(T)$ w.r.t. the standard basis, and $B = [u_1, \dots, u_n]$ where $u_1, \dots, u_n$ is a basis of V. Assume T is invertible, then $B^{-1}AB[T]_B^{-1} = I \rightarrow [T]_B^{-1} = B^{-1}A^{-1}B$ and $[T]_B^{-1}B^{-1}AB = I \rightarrow [T]_B^{-1} = B^{-1}A^{-1}B$, meaning that $[T]_B^{-1} = [T^{-1}]_B$. Assume now that $[T]_B = B^{-1}AB$ is invertible, then $[T]_B^{-1} = (B^{-1}AB)^{-1} = B^{-1}A^{-1}B$, requiring A to be invertible. Therefore, T is invertible iff $[T]_B$ is invertible, in which case $[T]_B^{-1} = [T^{-1}]_B$.

Question 4

(a) Suppose $X \in \mathbb{R}^{n \times p}$ is full column rank, $Y \in \mathbb{R}^n$ and $\beta \in \mathbb{R}^p$. $\exists T \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ s.t. $X = M(T)$. $\operatorname{rank} X = p \rightarrow \operatorname{null} T = \{0\} \rightarrow \operatorname{null} T^*T = \{0\}$ and since T^*T is an injective operator on $\mathbb{R}^n$, it is invertible. $\|Y - T\beta\|^2 = \langle Y - T\beta, Y - T\beta \rangle = \|Y\|^2 - 2\langle Y, T\beta \rangle + \langle T\beta, T\beta \rangle = \|Y\|^2 - 2\langle T^*Y, \beta \rangle + \langle T^*T\beta, \beta \rangle$. Hence $\nabla_\beta \|Y - T\beta\|^2 = -2T^*Y + 2T^*T\beta = 0 \rightarrow T^*Y = T^*T\hat{\beta} \rightarrow \hat{\beta} = (T^*T)^{-1}T^*Y = (X^TX)^{-1}X^TY$. Logically, this is a minimum point since $\|Y - X\beta\|$ measures distance between two vectors. Let $u_1, \dots, u_p$ an orthonormal list in $\mathbb{R}^n$ that span $\operatorname{Im} T$, then $T^*Y = T^*T\hat{\beta} \rightarrow T^*(Y - T\hat{\beta}) = 0$, implying that $Y - T\hat{\beta}$ is orthogonal to $\operatorname{Im} T$, or $\forall i = 1, \dots, p$ $\langle Y - T\hat{\beta}, u_i \rangle = 0 \rightarrow \langle Y, u_i \rangle = \langle T\hat{\beta}, u_i \rangle$. Since $T\hat{\beta} \in \operatorname{Im} T$ then $T\hat{\beta} = \langle T\hat{\beta}, u_1 \rangle u_1 + \dots + \langle T\hat{\beta}, u_p \rangle u_p \rightarrow T\hat{\beta} = \langle Y, u_1 \rangle u_1 + \dots + \langle Y, u_p \rangle u_p = P_{\operatorname{Im} T} Y$.

(b) Let $n \in \mathbb{N}$ and suppose $u_1, u_2, Y \in \mathbb{R}^n$ s.t. $u_1 = [1, 1, \dots, 1]$, $u_2 = [x_1, \dots, x_n]$, $Y = [y_1, \dots, y_n]$ for some $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{R}$ s.t. u_1 and u_2 are linearly independent, and set $X = [u_1, u_2]$. Then, $\hat{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = (X^TX)^{-1}X^TY = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix} = \frac{1}{n\sum x_i^2 - (\sum x_i)^2} \begin{bmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & n \end{bmatrix} \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix} = \frac{1}{n\sum x_i^2 - (\sum x_i)^2} \begin{bmatrix} \sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i \\ n\sum x_i y_i - \sum x_i \sum y_i \end{bmatrix}$. Assuming $X, Y \sim \text{Uniform}(\{1, \dots, n\})$ with $P(X = x_i) = P(Y = y_i) = \frac{1}{n}$, we get $\beta_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2} = \frac{E(XY) - E(X)E(Y)}{E(X^2) - E(X)^2} = \frac{\operatorname{Cov}(X, Y)}{\operatorname{Var}(X)}$ and $\beta_0 = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n\sum x_i^2 - (\sum x_i)^2} = \frac{\sum y_i}{n} - \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2} \frac{\sum x_i}{n} = E(Y) - \beta_1 E(X)$, which are the typical results for the coefficients of a single variable linear regression problem with $\operatorname{var}(x) \neq 0$.