ECE 269 - Solution to Homework 3

Fall 2021

Problem 1 It is given that $\mathcal{V}$ is a subspace of $\mathbb{R}^n$, $\mathcal{V}^\perp = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}^T\mathbf{y} = 0, \forall \mathbf{y} \in \mathcal{V}\}$.

(a) Consider $\alpha, \beta \in \mathbb{R}, \mathbf{v_1}, \mathbf{v_2} \in \mathcal{V}^\perp$. So $\forall \mathbf{u} \in \mathcal{V}$:
$(\alpha\mathbf{v_1} + \beta\mathbf{v_2})^T\mathbf{u} = \alpha\mathbf{v_1}^T\mathbf{u} + \beta\mathbf{v_2}^T\mathbf{u} = 0$ since $\mathbf{v_1}^T\mathbf{u} = 0 = \mathbf{v_2}^T\mathbf{u}$.
Also, $\mathbf{0} \in \mathcal{V}^\perp$ since $\mathbf{0}^T\mathbf{u} = 0, \forall \mathbf{u} \in \mathcal{V}$. Thus, $\mathcal{V}^\perp$ is a subspace of $\mathbb{R}^n$.

(b) Using the given matrix $\mathbf{A}$, notice that:
$\mathcal{V} = \{\sum_{i=1}^k x_i\mathbf{v_i} : x_1, ..., x_k \in \mathbb{R}\} =$
$\{\mathbf{A} \cdot (x_1, ..., x_k)^T : x_1, ..., x_k \in \mathbb{R}\} =$
$\{\mathbf{A}\mathbf{x} : \mathbf{x} \in \mathbb{R}^k\} = \mathcal{R}(\mathbf{A})$. Meaning $\mathcal{V}$ is the range space of $\mathbf{A}$.
Similarly:
$\mathcal{V}^\perp = \{\mathbf{y} \in \mathbb{R}^n : \mathbf{y}^T \sum_{i=1}^k x_i\mathbf{v_i} = 0, \forall x_1, ..., x_k \in \mathbb{R}\}$
$= \{\mathbf{y} \in \mathbb{R}^n : (\sum_{i=1}^k x_i\mathbf{v_i}^T)\mathbf{y} = 0, \forall x_1, ..., x_k \in \mathbb{R}\}$
$= \{\mathbf{y} \in \mathbb{R}^n : \mathbf{x}^T\mathbf{A}^T\mathbf{y} = 0, \forall \mathbf{x} \in \mathbb{R}^k\}$
$= \{\mathbf{y} \in \mathbb{R}^n : \mathbf{A}^T\mathbf{y} = 0\}$
$= \mathcal{N}(\mathbf{A}^T)$.
Meaning the orthogonal complement of the range space is the null space of the transpose.

(c) Consider $\mathbf{x} \in \mathcal{V}$. By definition of the orthogonal subspace $\mathbf{x}^T\mathbf{y} = 0, \forall \mathbf{y} \in \mathcal{V}^\perp \rightarrow \mathbf{x} \in (\mathcal{V}^\perp)^\perp$. Thus, $\mathcal{V} \subseteq (\mathcal{V}^\perp)^\perp$. Using the result of section (d), $dim(\mathcal{V}) + dim(\mathcal{V}^\perp) = n \rightarrow dim(\mathcal{V}^\perp) + dim((\mathcal{V}^\perp)^\perp) = n \rightarrow dim((\mathcal{V}^\perp)^\perp) = dim(\mathcal{V}) \rightarrow \mathcal{V} = (\mathcal{V}^\perp)^\perp$.

(d) Using section (b), we get that
$dim(\mathcal{V}) = dim(\mathcal{R}(\mathbf{A})) = rank(\mathbf{A}) = rank(\mathbf{A}^T)$. And,
$dim(\mathcal{V}^\perp) = dim(\mathcal{N}(\mathbf{A}^T))$.
Using the rank nullity theorem and the fact that $\mathbf{A}^T \in \mathbb{R}^{k\times n}$:
$n = rank(\mathbf{A}^T) + dim(\mathcal{N}(\mathbf{A}^T)) = dim(\mathcal{V}) + dim(\mathcal{V}^\perp)$.

(e) Assuming $\mathcal{V} \subseteq \mathcal{W}$ implies:
$\forall \mathbf{v} \in \mathcal{V} \rightarrow \mathbf{v} \in \mathcal{W}$, so taking $\mathbf{u} \in \mathcal{W}^\perp$:
$\mathbf{u}^T\mathbf{w} = 0, \forall \mathbf{w} \in \mathcal{W} \rightarrow \mathbf{u}^T\mathbf{w} = 0, \forall \mathbf{w} \in \mathcal{V} \rightarrow \mathbf{u} \in \mathcal{V}^\perp \rightarrow \mathcal{W}^\perp \subseteq \mathcal{V}^\perp$.

(f) We know that if $\mathbf{v}$ is a projection of $\mathbf{x}$ on $\mathcal{V}$, the $\mathbf{v}^\perp = \mathbf{x} - \mathbf{v}$ is orthogonal to all vectors in $\mathcal{V}$, meaning $\mathbf{v}^\perp \in \mathcal{V}^\perp$. Let us assume $\exists \mathbf{v} \neq \mathbf{u}, \mathbf{v}, \mathbf{u} \in \mathcal{V}$ and $\mathbf{v}^\perp, \mathbf{u}^\perp \in \mathcal{V}^\perp$ such that $\mathbf{x} = \mathbf{v} + \mathbf{v}^\perp = \mathbf{u} + \mathbf{u}^\perp$. So $\mathbf{v} - \mathbf{u} = \mathbf{u}^\perp - \mathbf{v}^\perp$. However, since $\mathcal{V}, \mathcal{V}^\perp$ are subspaces, we get that $\mathbf{v}-\mathbf{u} \in \mathcal{V}, \mathbf{u}^{\perp}-\mathbf{v}^{\perp} \in \mathcal{V}^{\perp}$, so $\mathbf{v}-\mathbf{u} \in \mathcal{V} \cap \mathcal{V}^{\perp} = \mathbf{0} \rightarrow \mathbf{v} = \mathbf{u}$ in contradiction to the assumption. Thus, the representation is unique.

Problem 2 A symmetric matrix $\mathbf{P} = \mathbf{P}^T \in \mathbb{R}^{n\times n}$ is said to be a projection matrix if $\mathbf{P} = \mathbf{P}^2$.

(a) $(\mathbf{I}-\mathbf{P})^T = \mathbf{I}-\mathbf{P}^T = \mathbf{I}-\mathbf{P}$ and so $\mathbf{I}-\mathbf{P}$ is symmetric. Also, $(\mathbf{I}-\mathbf{P})^2 = (\mathbf{I}-\mathbf{P})(\mathbf{I}-\mathbf{P}) = \mathbf{I}-\mathbf{P}-\mathbf{P}+\mathbf{P}^2 = \mathbf{I}-2\mathbf{P}+\mathbf{P} = \mathbf{I}-\mathbf{P}$. Hence, $\mathbf{I}-\mathbf{P}$ is symmetric and $\mathbf{I}-\mathbf{P} = (\mathbf{I}-\mathbf{P})^2$, and thus it is a projection matrix.

(b) Notice that $(\mathbf{U}\mathbf{U}^T)^T = (\mathbf{U}^T)^T\mathbf{U}^T = \mathbf{U}\mathbf{U}^T$ and $(\mathbf{U}\mathbf{U}^T)^2 = \mathbf{U}\mathbf{U}^T\mathbf{U}\mathbf{U}^T = \mathbf{U}\mathbf{I}\mathbf{U}^T = \mathbf{U}\mathbf{U}^T$.

(c) First of all, $\mathbf{A}^T\mathbf{A}$ is invertible for a full-rank tall matrix $\mathbf{A} \in \mathbb{R}^{n\times k}$ (To see this $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{0} \implies \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{0} \implies \mathbf{A}\mathbf{x} = \mathbf{0} \implies \mathbf{x} = \mathbf{0}$).
For symmetry, consider
$(\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T)^T = (\mathbf{A}^T)^T((\mathbf{A}^T\mathbf{A})^{-1})^T\mathbf{A}^T = \mathbf{A}((\mathbf{A}^T\mathbf{A})^{-1})^T\mathbf{A}^T = \mathbf{A}((\mathbf{A}^T\mathbf{A})^T)^{-1}\mathbf{A}^T = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$.
Also, consider
$(\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T)^2 = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$.

(d) It suffices to show that $\|\mathbf{x}-\mathbf{v}\|$ is minimized over all $\mathbf{v} \in \mathcal{R}(\mathbf{P})$ by $\mathbf{v}^* = \mathbf{P}\mathbf{x}$. For any $\mathbf{v} \in \mathcal{R}(\mathbf{P})$:
$\|\mathbf{x}-\mathbf{v}\|^2 = \|\mathbf{x}-\mathbf{P}\mathbf{x}+\mathbf{P}\mathbf{x}-\mathbf{v}\|^2 = \|\mathbf{x}-\mathbf{P}\mathbf{x}\|^2+\|\mathbf{P}\mathbf{x}-\mathbf{v}\|^2+2(\mathbf{x}-\mathbf{P}\mathbf{x})^T(\mathbf{P}\mathbf{x}-\mathbf{v})$
$= \|\mathbf{x}-\mathbf{P}\mathbf{x}\|^2 + \|\mathbf{P}\mathbf{x}-\mathbf{v}\|^2 + 2(\mathbf{x}^T-\mathbf{x}^T\mathbf{P})(\mathbf{P}\mathbf{x}-\mathbf{v})$
$= \|\mathbf{x}-\mathbf{P}\mathbf{x}\|^2 + \|\mathbf{P}\mathbf{x}-\mathbf{v}\|^2 + 2(\mathbf{x}^T\mathbf{P}\mathbf{x}-\mathbf{x}^T\mathbf{v}-\mathbf{x}^T\mathbf{P}^2\mathbf{x}+\mathbf{x}^T\mathbf{P}\mathbf{v})$
$= \|\mathbf{x}-\mathbf{P}\mathbf{x}\|^2 + \|\mathbf{P}\mathbf{x}-\mathbf{v}\|^2 + 2(\mathbf{x}^T\mathbf{P}\mathbf{x}-\mathbf{x}^T\mathbf{v}-\mathbf{x}^T\mathbf{P}\mathbf{x}+\mathbf{x}^T\mathbf{v})$
$= \|\mathbf{x}-\mathbf{P}\mathbf{x}\|^2 + \|\mathbf{P}\mathbf{x}-\mathbf{v}\|^2 \geq \|\mathbf{x}-\mathbf{P}\mathbf{x}\|^2$.
Where we used $\mathbf{P} = \mathbf{P^2}$ and $\mathbf{P}\mathbf{v} = \mathbf{v}, \forall \mathbf{v} \in \mathcal{R}(\mathbf{P})(\mathbf{v} = \mathbf{P}\mathbf{x} = \mathbf{P^2}\mathbf{x} = \mathbf{P}\mathbf{v}$). To achieve equality in the last inequality, we need $\|\mathbf{P}\mathbf{x}-\mathbf{v}\| = 0 \rightarrow \mathbf{v} = \mathbf{P}\mathbf{x}$. Thus, $argmin_{\mathbf{v}\in\mathcal{R}(\mathbf{P})}\|\mathbf{x}-\mathbf{v}\| = \mathbf{P}\mathbf{x}$.

(e) Consider $\mathbf{P} = \mathbf{u}\mathbf{u}^T$. Since $\mathbf{P} = \mathbf{P}^T$ and $\mathbf{P} = \mathbf{P}^2$, $\mathbf{P}$ is a projection matrix. Since $\mathcal{R}(\mathbf{P}) = span(\mathbf{u})$, using part (d) $\mathbf{P}\mathbf{x}$ is the projection of $\mathbf{x}$ onto $span(\mathbf{u})$. This can also be directly verified since $(\mathbf{u}^T\mathbf{x})\mathbf{u} = \mathbf{u}\mathbf{u}^T\mathbf{x} = \mathbf{P}\mathbf{x}$ is the component of $\mathbf{x}$ in the direction for the unit vector $\mathbf{u}$.