Extra-Credit Problem – Four Fundamental Subspaces, Pseudoinverses, Least Squares

One fact that you have been asked to show several times since early in this course is the fact that the rank of a matrix $\mathbf{A}$ (the column rank) and the rank of its transpose $\mathbf{A}^{\mathrm{T}}$ (the row rank) are the same. There is much more that can be said, and that's what this extra-credit problem will explore.

If $\mathbf{A}$ is an $m \times n$ matrix, then it represents a linear transformation with domain $\mathbf{R}^n$ and codomain $\mathbf{R}^m$. Its transpose $\mathbf{A}^{\mathrm{T}}$ will then be an $n \times m$ matrix with domain $\mathbf{R}^m$ and codomain $\mathbf{R}^n$. Associated with these two matrices are their respective kernels and images: $\ker(\mathbf{A})$, $\operatorname{im}(\mathbf{A})$, $\ker(\mathbf{A}^{\mathrm{T}})$, and $\operatorname{im}(\mathbf{A}^{\mathrm{T}})$ – the ***Four Fundamental Subspaces*** associated with the matrix $\mathbf{A}$.

We have shown that in the domain, $\ker(\mathbf{A})$ and $\operatorname{im}(\mathbf{A}^{\mathrm{T}})$ are orthogonal complements; and in the codomain, $\operatorname{im}(\mathbf{A})$ and $\ker(\mathbf{A}^{\mathrm{T}})$ are orthogonal complements. The fact that $\operatorname{rank}(\mathbf{A}^{\mathrm{T}}) = \operatorname{rank}(\mathbf{A})$ means that $\operatorname{im}(\mathbf{A}^{\mathrm{T}})$ and $\operatorname{im}(\mathbf{A})$ have the same dimension, and are therefore ***isomorphic***. That is, when restricted to elements of $\operatorname{im}(\mathbf{A}^{\mathrm{T}})$, the matrix $\mathbf{A}$ provides a one-to-one linear correspondence onto $\operatorname{im}(\mathbf{A})$.

The **pseudoinverse** $\mathbf{A}^{+}$ of a matrix is defined as follows: To every element $\mathbf{y} \in \operatorname{im}(\mathbf{A})$, we define $\mathbf{A}^{+}\mathbf{y} = \mathbf{x}$ where $\mathbf{x} \in \operatorname{im}(\mathbf{A}^{\mathrm{T}})$ is the *unique* element such that $\mathbf{A}\mathbf{x} = \mathbf{y}$. To every element $\mathbf{y}$ in the complementary subspace $[\operatorname{im}(\mathbf{A})]^{\perp} = \ker(\mathbf{A}^{\mathrm{T}})$ we define $\mathbf{A}^{+}\mathbf{y} = \mathbf{0}$. Thus $\ker(\mathbf{A}^{+}) = \ker(\mathbf{A}^{\mathrm{T}})$ and $\operatorname{im}(\mathbf{A}^{+}) = \operatorname{im}(\mathbf{A}^{\mathrm{T}})$. In the case where $\mathbf{A}$ is an invertible $n \times n$ matrix, this pseudoinverse $\mathbf{A}^{+}$ coincides with the inverse $\mathbf{A}^{-1}$.

Problem #1: Given the matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$:

(a) Find bases for each of the Four Fundamental Subspaces of $\mathbf{A}$, i.e. $\ker(\mathbf{A})$, $\operatorname{im}(\mathbf{A})$, $\ker(\mathbf{A}^{\mathrm{T}})$, and $\operatorname{im}(\mathbf{A}^{\mathrm{T}})$.

(b) Find the matrix for the pseudoinverse $\mathbf{A}^{+}$. [This will be a 3×2 matrix.]

(c) Find the matrix for the pseudoinverse of $\mathbf{B} = \mathbf{A}^{\mathrm{T}}$. [This will be a 2×3 matrix.]

(d) What is the relationship between these two pseudoinverse matrices?

(e) Show that in general, i.e. for any matrix $\mathbf{A}$, that the matrices $\mathbf{A}\mathbf{A}^{+}$ and $\mathbf{A}^{+}\mathbf{A}$ are ***idempotent***, i.e. $(\mathbf{A}\mathbf{A}^{+})^2 = \mathbf{A}\mathbf{A}^{+}$ and $(\mathbf{A}^{+}\mathbf{A})^2 = \mathbf{A}^{+}\mathbf{A}$. [It is worth noting that this is a property shared with *projections*.]

(f) Find the matrices $\mathbf{A}\mathbf{A}^{+}$ and $\mathbf{A}^{+}\mathbf{A}$ for the given matrix $\mathbf{A}$ above, and verify that these are idempotent.

Problem #2: When solving a linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$, sometimes there is a unique solution, sometimes there can be infinitely many solutions, and sometimes the system is inconsistent and has no solutions. In the latter case, we might seek Least Squares approximate solutions, and we have seen that these are solutions of the normal equation $\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{A}^{\mathrm{T}}\mathbf{b}$. In the case where $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ is invertible, there will be a unique Least Squares solution $\mathbf{x} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b}$, but this will only be the case when $\ker(\mathbf{A}^{\mathrm{T}}\mathbf{A}) = \ker(\mathbf{A}) = \{\mathbf{0}\}$, i.e. when the columns of the matrix $\mathbf{A}$ are linearly independent.

(a) Show that in the case where there is a unique Least Squares solution, that solution is $\mathbf{x} = \mathbf{A}^{+}\mathbf{b}$ where $\mathbf{A}^{+}$ is the pseudoinverse of $\mathbf{A}$.

(b) More generally, in show that $\mathbf{x} = \mathbf{A}^{+}\mathbf{b}$ gives the Least Squares solution of minimal norm, i.e. if $\mathbf{x}_1$ is any other Least Squares solution, then $\|\mathbf{x}\| \le \|\mathbf{x}_1\|$.