Solutions for Homework 5 Numerical Linear Algebra 1 Fall 2002

Problem 1

The pseudoinverse for rectangular full column-rank matrices behaves much as the inverse for non-singular matrices. To see this show the following identities are true (Stewart 73):

1. $AA^\dagger A = A$
2. $A^\dagger AA^\dagger = A^\dagger$
3. $A^\dagger A = (A^\dagger A)^T$
4. $AA^\dagger = (AA^\dagger)^T$

5. If $A \in \mathbb{R}$ has orthonormal columns then $A^\dagger = A^T$. Why is this important for consistency with simpler forms of least squares problems that we have discussed?

Solution:

\[ AA^\dagger A = A(A^T A)^{-1}A^T A \]
\[ = A(A^T A)^{-1}(A^T A) \]
\[ = AI \]
\[ = A \]

\[ A^\dagger AA^\dagger = (A^T A)^{-1}A^T A(A^T A)^{-1}A^T \]
\[ = (A^T A)^{-1}(A^T A)(A^T A)^{-1}A^T \]
\[ = I(A^T A)^{-1}A^T \]
\[ = A^\dagger \]

\[ A^\dagger A = (A^T A)^{-1}A^T A \]
\[ = (A^T A)^{-1}(A^T A) \]
\[ = I \]
\[ = A^T A(A^T A)^{-1} \]
\[ = A^T A(A^T A)^{-T} \]
\[ = (A^\dagger A)^T \]

\[ AA^\dagger = A(A^T A)^{-1}A^T \]
\[ = A(A^T A)^{-T}A^T \]
\[ = (A(A^T A)^{-1}A^T)^T \]
\[ = (AA^\dagger)^T \]
Since $A$ has orthonormal columns we have $A^T A = I$ and therefore $A^\dagger = (A^T A)^{-1} A^T = I A^T = A^T$ as desired.

Recall we have the problem of minimizing $\|b - Ax\|_2$ with $A \in \mathbb{R}^{n \times k}$ with $n \leq k$ having full column rank. This splits into several cases when we consider consistency.

If $n = k$ then $A$ is nonsingular and $x = A^{-1} b$. We have shown earlier that in this case $A^\dagger = A^{-1}$.

If $n > k$ then we want $x = A^\dagger b$ to be such that $Ax = Pb$ where $P$ is the projector from $\mathbb{R}^n$ onto $\mathcal{R}(A)$. We showed this in the notes based on the orthogonality of the residual.

This case was also discussed in simplified form earlier in the notes when first considering orthonormal bases for a space, i.e., a basis for a subspace such that $Q^T Q = I$ and the columns of $Q$ are the basis. In this case we had for the case above with $n > k$ that $x = Q^T b$. The exercise shows this is consistent with the definition of $A^\dagger$ when $A^T A = I$.

**Problem 2**

Recall, that any subspace $S$ of $\mathbb{R}^n$ of dimension $k \leq n$ must have an orthogonal matrix $Q \in \mathbb{R}^{n \times k}$ with orthonormal columns such that $\mathcal{R}(Q) = S$, i.e., the range of $Q$ is the subspace and that the matrix $P = QQ^T$ is called a projector, i.e., $Px$ is the unique component of $x$ contained in $S$.

1. $P$ is clearly symmetric, show that it is idempotent, i.e., $P^2 = P$.

2. Show that $\mathcal{R}(P) = S$.

3. Show that if $M$ is an idempotent symmetric matrix then it is a projector onto $\mathcal{R}(M)$.

4. Let $Q_1 \in \mathbb{R}^{n \times k}$ and $Q_2 \in \mathbb{R}^{n \times k}$, $k \leq n$, have orthonormal columns and be such that $\mathcal{R}(Q_1) = \mathcal{R}(Q_2)$. Show that there must exist $W \in \mathbb{R}^{k \times k}$ such that $W^T W = WW^T = I_k$ and $Q_2 = Q_1 W$. (In other words, any two orthogonal bases of the same subspace are related by a small orthogonal transformation typically called a rotation in the subspace.)

**Solution:**

$P$ is idempotent since

$$P^2 = PP = (QQ^T)(QQ^T) = QIQ^T = QQ^T = P$$

To prove that $\mathcal{R}(P) = S = \mathcal{R}(Q)$ note

$$y \in \mathcal{R}(P) \quad y = Px \quad = QQ^T x \quad = Qc \quad \in \mathcal{R}(Q)$$
To prove the other direction note $y \in \mathcal{R}(Q)$ implies $y = Qc$ and since $Q^TQ = I$ we have $c = Q^Ty$.

Now let $x = y + z$ where $z$ is any vector in the null space of $Q$ then we have

\[
\begin{align*}
y & \in \mathcal{R}(Q) \\
y &= Qc \\
&= QQ^Ty + 0 \\
&= QQ^Ty + QQ^Tz \\
&= Px \\
&\in \mathcal{R}(P)
\end{align*}
\]

If $M$ is a symmetric idempotent matrix then we know $M^T = M$ and $MM = MM^T = M^TM = M$. If $\mathcal{S} = \mathcal{R}(M)$ then we know that any vector $x$ is such that $x = y + z$ where $y \in \mathcal{S}$ and $z \in \mathcal{S}^\perp$.

Therefore, by definition we know that $y = Mc$ (possibly not uniquely since the columns of $M$ are not necessarily a basis for $\mathcal{S}$) and $M^Tz = 0$. To see $M$ is a projector we have

\[
\begin{align*}
Mx &= M(y + z) \\
&= My + Mz \\
&= My + MM^Tz \\
&= My \\
&= M(Mc) \\
&= MMc \\
&= Mc \\
&= y
\end{align*}
\]

as desired.

Finally, let $\mathcal{R}(Q_1) = \mathcal{R}(Q_2) = \mathcal{S}$. By definition, given $x \in \mathcal{S}$ we have unique $c$ and $d$ such that $x = Q_1c = Q_2d$. The columns of $Q_2$ are all in $\mathcal{S}$ therefore for $1 \leq i \leq k$ we have $Q_2e_i = Q_1w_i$ uniquely. This can be written as $Q_2 = Q_1W$ with $W = \left( \begin{array}{c} w_1 \\
\vdots \\
w_k \end{array} \right)$. We also have $Q_2^TQ_2 = I = W^TQ_1^TQ_1W = W^TW$ and since $W$ is square and nonsingular we have $W^T = W^{-1}$ and therefore $WW^T = I$ also.