Lecture 8 – ECE 275A

Moore-Penrose Conditions & SVD
**Four Moore-Penrose Pseudoinverse Conditions**

**MOORE-PENROSE THEOREM**

Consider a linear operator $A : \mathcal{X} \to \mathcal{Y}$.

A linear operator $M : \mathcal{Y} \to \mathcal{X}$ is the unique pseudoinverse of $A$, $M = A^+$, if and only if it satisfies the

**Four Moore-Penrose Conditions:**

1. $(AM)^* = AM$
2. $(MA)^* = MA$
3. $AMA = A$
4. $MAM = M$

More simply we usually say that $A^+$ is the unique p-inv of $A$ iff

1. $(AA^+)^* = AA^+$
2. $(A^+A)^* = A^+A$
3. $AA^+A = A$
4. $A^+AA^+ = A^+$

- The theorem statement provides greater clarity because there we distinguish between a *candidate* p-inv $M$ and the *true* p-inv $A^+$.
- If and only if the candidate p-inv $M$ satisfies the four M-P conditions can we claim that indeed $A^+ = M$. 
Proof of the M-P Theorem

First we reprise some basic facts that are consequences of the definitional properties of the pseudoinverse.

FACT 1: \( \mathcal{N}(A^+) = \mathcal{N}(A^*) \)

FACT 2: \( \mathcal{R}(A^+) = \mathcal{R}(A^*) \)

FACT 3: \( P_{\mathcal{R}(A)} = AA^+ \)

FACT 4: \( P_{\mathcal{R}(A^*)} = A^+A \)

We now proceed to prove two auxiliary theorems (Theorems A and B).
Proof of the M-P Theorem – Cont.

THEOREM A

Let $C : X \rightarrow Y$ and $B : Y \rightarrow Z$ be linear mappings. It is readily shown that the composite mapping $BC : X \rightarrow Z$ is a linear mapping where $BC$ is defined by

$$(BC)x \triangleq B(Cx) \quad \forall x \in X$$

Then

$$\mathcal{N}(B) \cap \mathcal{R}(C) = \{0\} \quad \Rightarrow \quad BC = 0 \text{ iff } C = 0$$

Proof

$$BC = 0 \iff (BC)x = 0 \quad \forall x \quad \text{(definition of zero operator)}$$

$$\iff B(Cx) = 0 \quad \forall x \quad \text{(definition of composition)}$$

$$\iff Cx = 0 \quad \forall x \quad \text{(because } Cx \in \mathcal{R}(C) \cap \mathcal{N}(B) = \{0\}, \ \forall x)$$

$$\iff C = 0 \quad \text{(definition of zero operator)}$$

QED
Theorem B covers the uniqueness part of the M-P Theorem.

THEOREM B. The pseudoinverse of $A$ is unique. □

Proof. Suppose that $A^+$ and $M$ are both p-inv's of $A$. Then Fact 3 gives

\[ P_{\mathcal{R}(A)} = AA^+ = AM \]

or

\[ A(A^+ - M) = 0 \]

From Fact 2, $\mathcal{R}(A^*) = \mathcal{R}(A^+) = \mathcal{R}(M)$ and as a consequence

\[ \mathcal{R}(A^+ - M) \subseteq \mathcal{R}(A^*) \]

But $\mathcal{R}(A^*) \perp \mathcal{N}(A)$ and therefore

\[ \mathcal{R}(A^+ - M) \subseteq \mathcal{R}(A^*) = \mathcal{N}(A)^\perp \]

so that

\[ \mathcal{N}(A) \cap \mathcal{R}(A^+ - M) = \{0\} \]

Therefore from Theorem A,

\[ A^+ - M = 0 \Rightarrow A^+ = M \]

QED
Proof of the M-P Theorem – Cont.

Necessity (‘only if’ part) of the M-P Conditions. Assume that $M = A^+$.

Necessity of M-P Conditions I & II. Easy consequences of Facts 3 and 4.

Necessity of M-P Condition III. Note that Fact 4 and indempotency of a projection operator implies $(A^+ A)(A^+ A) = A^+ A$, or

$$A^+ \ (AA^+ A - A) = AC = 0$$

$$\triangleq C = A(A^+ A - I)$$

We have $\mathcal{N}(A^+) = \mathcal{N}(A^*)$ (Fact 1) and $\mathcal{R}(C) \subset \mathcal{R}(A) = \mathcal{N}(A^*)^\perp$. Therefore $\mathcal{N}(A^+) \cap \mathcal{R}(C) = \mathcal{N}(A^*) \cap \mathcal{R}(C) = \{0\}$ so that by Theorem A, $C = 0$.

Necessity of M-P Condition IV. Note that Fact 3 and indempotency of a projection operator implies $(AA^+)(AA^+) = AA^+$, or

$$A \ (A^+ AA^+ - A^+) = AC = 0$$

$$\triangleq C = A^+ (AA^+ - I)$$

With $\mathcal{R}(A^+) = \mathcal{R}(A^*)$ (Fact 2) we have $\mathcal{R}(C) \subset \mathcal{R}(A^+) = \mathcal{R}(A^*) = \mathcal{N}(A)^\perp$. Therefore $\mathcal{N}(A) \cap \mathcal{R}(C) = \{0\}$ so that by Theorem A, $C = 0$. 
Proof of the M-P Theorem – Cont.

Sufficiency (‘if’ part) of the M-P Conditions.

Here we assume that \( M \) satisfies all four of the M-P conditions and then show as a consequence that \( M = A^+ \).

We do this using the following steps.

1. First prove that \( P_{R(A)} = AM \) (proving that \( AM = AA^+ \) via uniqueness of projection operators).

2. Prove that \( P_{R(A^*)} = MA \) (proving that \( MA = A^+ A \)).

3. Finally, prove that as a consequence of (1) and (2), \( M = A^+ \).
Proof of the M-P Theorem – Cont.

Sufficiency – Cont.

Step (1):

From M-P conditions 1 & 3, \((AM)^* = AM\) and \(AM = (AMA)M = (AM)(AM)\), showing that \(AM\) is an orthogonal projection operator. But onto what? Obviously onto a subspace of \(\mathcal{R}(A)\) as \(\mathcal{R}(AM) \subset \mathcal{R}(A)\). However

\[
\mathcal{R}(A) = A(\mathcal{X}) = AMA(\mathcal{X}) = AM(A(\mathcal{X})) = AM(\mathcal{R}(A)) \subset AM(\mathcal{Y}) = \mathcal{R}(AM) \subset \mathcal{R}(A)
\]

yields the stronger statement that \(\mathcal{R}(AM) = \mathcal{R}(A)\). Thus \(AM\) is the orthogonal projector onto the range of \(A\), \(P_{\mathcal{R}(A)} = AM = AA^+\).

Step (2):

From M-P conditions 2 & 3, \((MA)^* = MA\) and \(MA = M(AMA) = (MA)(MA)\), showing that \(MA\) is an orthogonal projection operator. Note that M-P conditions 3 and 2 imply \(A^* = (AMA)^* = A^*M^*A^*\) and \(MA = (MA)^* = A^*M^*\). We have

\[
\mathcal{R}(A^*) = A^*(\mathcal{Y}) = (A^*M^*A^*)(\mathcal{Y}) = (A^*M^*)(\mathcal{R}(A^*)) \subset \underbrace{(A^*M^*)(\mathcal{X})}_{=\mathcal{R}(A^*)} \subset \mathcal{R}(A^*)
\]

showing that \(\mathcal{R}(MA) = \mathcal{R}(A^*)\). Thus \(P_{\mathcal{R}(A^*)} = MA = A^+A\).
Proof of the M-P Theorem – Cont.

Sufficiency – Cont.

Step (3):
Note that we have yet to use M-P condition 4, $MAM = M$. From M-P condition 4 and the result of Step (2) we have

$$MAM = P_{\mathcal{R}(A^*)}M = M$$

Obviously, then $\mathcal{R}(M) \subset \mathcal{R}(A^*)$, as can be rigorously shown via the subspace chain

$$\mathcal{R}(M) = M(\mathcal{Y}) = P_{\mathcal{R}(A^*)}M(\mathcal{Y}) = P_{\mathcal{R}(A^*)}(\mathcal{R}(M)) \subset P_{\mathcal{R}(A^*)}(\mathcal{X}) = \mathcal{R}(A^*)$$

Recalling that $\mathcal{R}(A^+) = \mathcal{R}(A^*)$ (Fact 2), it therefore must be the case that

$$\mathcal{R}(M - A^+) \subset \mathcal{R}(A^*) = \mathcal{N}(A)^\perp$$

Using the result of Step (1), $P_{\mathcal{R}(A)} = AM = AA^+$, we have

$$A(M - A^+) = 0$$

with $\mathcal{N}(A) \cap \mathcal{R}(M - A^+) = \{0\}$. Therefore Theorem A yields $M - A^+ = 0$. QED
Proof of the M-P Theorem – Cont.

Note the similarity of the latter developments in Step 3 to the proof of Theorem B. In fact, some thought should convince yourself that the latter part of Step 3 provides justification for the claim that the pseudoinverse is unique, so that Theorem B can be viewed as redundant to the proof of the M-P Theorem.

Theorem B was stated to introduce the student to the use of Theorem A (which played a key role in the proof of the M-P Theorem) and to present the uniqueness of the pseudoinverse as a key result in its own right.
**Singular Value Decomposition (SVD)**

Henceforth, let us consider only Cartesian Hilbert spaces (i.e., spaces with identity metric matrices) and consider all finite dimensional operators to be represented as complex $m \times n$ matrices,

$$A_{m \times n} : \mathcal{X} = \mathbb{C}^n \rightarrow \mathcal{Y} = \mathbb{C}^m$$

Note that $A$ in general is non-square and therefore does not have a spectral representation (because eigenvalues and eigenvectors are not defined).

Even if $A$ is square, it will in general have complex valued eigenvalues and non-orthogonal eigenvectors. Even worse, a general $n \times n$ matrix can be defective and not have a full set of $n$ eigenvectors, in which case $A$ is not diagonalizable. In the latter case, one must use generalized eigenvectors to understand the spectral properties of the matrix (which is equivalent to placing the matrix in *Jordan Canonical Form*).

It is well known that if a square, $n \times n$ complex matrix is self-adjoint (Hermitian), $A = A^H$, then its eigenvalues are all real and it has a full complement of $n$ eigenvectors that can all be chosen to orthonormal. In this case for eigenpairs $(\lambda_i, x_i), i = 1, \cdots, n$, $A$ has a simple spectral representation given by an orthogonal transformation,

$$A = \lambda_1 x_1 x_1^H + \cdots + \lambda_n x_n x_n^H = X \Lambda X^H$$

with $\Lambda = \text{diag}(\lambda_1 \cdots \lambda_n)$, and $X$ is unitary, $X^H X = XX^H = I$, where the columns of $X$ are comprised of the orthonormal eigenvectors $x_i$. If in addition, a hermitian matrix $A$ is positive-semidefinite, denoted as $A \geq 0$, then the eigenvalues are all non-negative, and all strictly positive if the matrix $A$ is invertible (positive-definite, $A > 0$).
Singular Value Decomposition (SVD) – Cont.

Given an arbitrary (nonsquare) complex matrix operator $A \in \mathbb{C}^{m \times n}$ we can ‘regularized’ its structural properties by ‘squaring’ it to produce a hermitian, positive-semidefinite matrix, and thereby exploit the very nice properties of hermitian, positive-semidefinite matrices mentioned above.

Because matrix multiplication is noncommutative, there are two ways to ‘square’ $A$ to form a hermitian, positive-semidefinite matrix, viz

$$AA^H \quad \text{and} \quad A^HA$$

It is an easy exercise to proved that both of these forms are hermitian, positive-semidefinite, recalling that a matrix $M$ is defined to be positive-semidefinite, $M \geq 0$, if and only if the associated quadratic form $\langle x, Mx \rangle = x^H Mx$ is real and positive-semidefinite

$$\langle x, Mx \rangle = x^H Mx \geq 0 \quad \forall x$$

Note that a sufficient condition for the quadratic form to be real is that $M$ be hermitian, $M = M^H$. For the future, recall that a positive-semidefinite matrix $M$ is positive-definite, $M > 0$, if in addition to the non-negativity property of the associated quadratic form we also have

$$\langle x, Mx \rangle = x^H Mx = 0 \quad \text{if and only if} \quad x = 0$$
In Lecture 9 we will show that the eigenstructures of the well-behaved hermitian, positive-semidefinite ‘squares’ $A^H A$ and $AA^H$ are captured in the Singular value Decomposition (SVD) introduced in Example 5 of Lecture 7. As noted in that example, knowledge of the SVD enables us to compute the pseudoinverse of $A$ in the rank deficient case.

The SVD will also allow us to compute a variety of important quantities, including the rank of $A$, orthonormal bases for all four fundamental subspaces of $A$, orthogonal projection operators onto all four fundamental subspaces of the matrix operator $A$, the spectral norm of $A$, the Frobenius norm of $A$, and the condition number of $A$.

The SVD will also provide a simple geometrically intuitive understanding of the nature of $A$ as an operator based on the action of $A$ as mapping hyperspheres in $\mathcal{R}(A^*)$ to hyperellipsoids in $\mathcal{R}(A)$ in addition to the fact that $A$ maps $\mathcal{N}(A)$ to 0.