9. Positive definite mappings

First, we'll introduce the tensor product, which we'll use later in this section.

Given two vector spaces $U, V$ over $K$, their tensor product, denoted $U \otimes V$, is a related vector space of dimension $(\dim U)(\dim V)$.

**Def.** If $\{u_1, \ldots, u_n\}$ is a basis of $U$ and $\{v_1, \ldots, v_m\}$ a basis of $V$, then $\{u_i \otimes v_j : 1 \leq i \leq n, 1 \leq j \leq m\}$ is a basis of $U \otimes V$.

**Analogy:** $U \cong \text{Span}\{1, x, \ldots, x^{n-1}\}$, $V \cong \text{Span}\{1, y, \ldots, y^{m-1}\}$, 

$U \otimes V \cong \text{Span}\{x^iy^j : 0 \leq i < n, 0 \leq j < m\}$.

Compare to $U \times V$, which has dimension $\dim U + \dim V$, with basis $\{(x_i, 0)\} \cup \{(0, y_j)\}$.

There's a better, basis-free way to construct $U \otimes V$. 
Def 2. $U \otimes V = \{ \Sigma u_i \otimes v_i : u_i \in U, v_i \in V \} / N$

where $N = \text{Span} \left\{ (u_1 + u_2) \otimes v - u_1 \otimes v - u_2 \otimes v, \right.$
$\left. u \otimes (v_1 + v_2) - u \otimes v_1 - u \otimes v_2 \right\}$

Basically, we are forcing the distributive law, i.e.,

$(u_1 + u_2) \otimes v = u_1 \otimes v + u_2 \otimes v$, etc.

Let $\text{Hom}(X, Y)$ be the set of linear maps from $X$ to $Y$.

Theorem 9.1: There is a natural isomorphism $U \otimes V \longrightarrow \text{Hom}(U', V)$.

Proof (sketch). Define the map as follows:

$U \otimes V \longrightarrow \text{Hom}(U', V)$

$u \otimes v \longmapsto \{ l \mapsto (l, u) \cdot v \}$ and extend linearly

$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \otimes \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} \mapsto \begin{bmatrix} v_1 \end{bmatrix} \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} = \begin{bmatrix} v_{1u_1} & v_{1u_2} & \cdots & v_{1u_n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{mu_1} & v_{mu_2} & \cdots & v_{mu_n} \end{bmatrix}_{m \times n}$

Remark: We could similarly define an isomorphism $U \otimes V \longrightarrow \text{Hom}(V', U)$.

Also, $e_i \otimes e_j$ under this isomorphism is the matrix $E_{ij}$ (i.e., the

ijkl entry is 1, all others 0).
If $U, V$ are Euclidean spaces (so $U^* = U$), there is a natural way to endow $U \otimes V$ with a Euclidean structure.

For $M, L \in \mathcal{L}(U, V)$, define $(M, L) = tr(L^* M) = \sum_{ij} \bar{e}_i \cdot m_{ji} \cdot e_j$.

Note that $\|M\|^2 = (M, M) = \sum_{ij} |m_{ji}|^2$

Clearly, $\{e_i, e_j : 1 \leq i \leq n, 1 \leq j \leq m\}$ is an orthonormal basis.

Ex: $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $L = \begin{bmatrix} a' & L'_1 \\ c' & d' \end{bmatrix}$, all real.

$$(M, L) = tr \ L^* M = tr \begin{bmatrix} a' & c' \\ b' & d' \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$= tr \begin{bmatrix} aa' + cc' & a'b + c'd' \\ a'b' + cd' & bb' + dd' \end{bmatrix}$$

$$= aa' + bb' + cc' + dd'.$$

Another way to view tensor products.

Let $X$ be an $n$-dimensional real vector space.

Note that $C$ is a 2-diml $\mathbb{R}$-vector space (basis $\{e_1, i\}$).

Suppose $A : X \to X$ is a linear map with min. poly $x^2 + 1$. 
Then $i$ and $-i$ are eigenvalues of $A$, but $i \notin \mathbb{R}$.

So if $v$ is an eigenvector for $\lambda = i$, $v \notin X$.

However, $v$ should lie in some "extension" of $X$.

In this bigger vector space, we want to have vectors like $zv$, $z \in \mathbb{C}$, $v \in X$.

What we really want is $\mathbb{C} \otimes X$.

This has basis $\{x_1, \ldots, x_n, ix_1, \ldots, ix_n\}$, where

$x_1, \ldots, x_n$ is a basis of $X$ [and here, $ix_j \leftrightarrow i \otimes x_j$].

Note that we need certain associativity and distributivity,
like $(3i)v = (i3)v = i(3v)$
i.e., $3i \otimes v = i3 \otimes v = i \otimes 3v$.

But this comes for free with the construction!

Similarly, compare this to polynomials $i$ matrices:

$(3x)i_{ij} = x_i(3y_j)$ and $(3u) V^T = u (3V^T)$

$3x_i \otimes y_j = x_i \otimes 3y_j$

$3u \otimes v = u \otimes 3v$. 
Def: A self-adjoint map $H: X \to X$ is positive (or positive definite) if $(x, Hx) > 0$ for all $x \neq 0$. It is nonnegative (or positive semidefinite) if $(x, Hx) \geq 0$.

We denote these as e.g., $H > 0$ and $H \geq 0$, resp.

Theorem 9.2. Let $X$ be a Euclidean space.

(i) The identity map $I$ is positive.

(ii) If $M, N > 0$ then $M+N > 0$ and $aM > 0$ for $a > 0$.

(iii) If $H > 0$ and $Q$ normal, then $Q^*HQ > 0$.

(iv) $H > 0$ iff all eigenvalues are positive.

(v) If $H > 0$ then $H$ is invertible.

(vi) Every positive map has a unique positive square root.

(vii) The set of positive maps is an open subset of the space of self-adjoint maps.

(viii) The boundary points of the set of positive maps are the nonnegative maps that are not positive.
Proof:  (i) \((x, Ix) = (x, x) > 0\) if \(x \neq 0\). ✓

(ii) \((x, (M+N)x) = (x, Mx) + (x, Nx) > 0\) if \(x \neq 0\). ✓

and \((x, aMx) = a(x, Mx) > 0\) ✓

(iii) \((x, Q^*HQx) = (Qx, HQx)\)

\[= (y, Hy) > 0\] where \(y = Qx\) ✓

(iv) \((\exists)\) \((x, Hx) = (x, \lambda x) = \lambda (x, x) = 0 \Rightarrow \lambda = 0\)

\((\Leftarrow)\) \(\frac{(x, Hx)}{(x, x)} \geq \lambda_{\min} > 0 \Rightarrow (x, Hx) \geq \lambda_{\min} \|x\| > 0\). ✓

(v) IF \(H\) is singular, then \(Hx = 0\) for some \(x \neq 0 \Rightarrow \lambda = 0\). ✓

(vi) Write \(\sqrt{H} x = \sum_{i=1}^{n} a_i \sqrt{\lambda_i} x_i\) \((x_i's\ e-vectors)\) ✓

(vii) Fix \(H > 0\), let \(N\) be any self-adjoint map such that \(\|N-H\| < \lambda_{\min}\).

Claim: \(N > 0\).

Put \(M = N-H\). Since \(\|M\| < \lambda_{\min}\), \(\|Mx\| < \lambda_{\min} \|x\|\) for all \(x \neq 0\).

Cauchy-Schwarz \(\Rightarrow\) \(\|(x, Mx)\| \leq \|x\| \cdot \|Mx\| < \lambda_{\min} \|x\|^2\).

Together we get for \(x \neq 0\):

\((x, Nx) = (x, (H+M)x) = (x, Hy) + (x, Mx) > \lambda_{\min} \|x\|^2 - \lambda_{\min} \|x\|^2 = 0.

\[\Rightarrow \lambda_{\min} \|x\|^2 > -\lambda_{\min} \|x\|^2\]

Thus \(N > 0\). ✓
(viii) By definition, if $K$ is on the boundary of positive maps, then there exists a sequence $H_n \to K$ of positive maps such that $\|H_n - K\| \to 0$.

By Cauchy-Schwarz, $\lim_{n \to \infty} (x, H_n x) = (x, K x)$, so $(x, K x) \geq 0$.

Since $K \geq 0$, then $K > 0$. 

Put a partial order on the set of self-adjoint maps:

Say $M \leq N$ if and only if $N - M > 0$.

We get the following properties (almost) for free:

Additive: $M_1 < N_1$ and $M_2 < N_2 \Rightarrow M_1 + M_2 < N_1 + N_2$.

Transitive: $L < M < N \Rightarrow L \leq N$.

Multiplicative: $M \leq N$, $\alpha \in \mathbb{R}$, $\Rightarrow Q^* M Q < Q^* N Q$.

Theorem 9.3: Suppose $0 < M < N$. Then $N^{-1} > M^{-1} > 0$.

Proof: First, suppose $N = I$.

Then $M < I \Rightarrow I - M > 0 \Rightarrow$ Eigenvalues of $I - M$ are positive.

Say $(I - M)x = \lambda x$ for $x \neq 0$.

Then $Mx = x - \lambda x = (1 - \lambda)x$. Since $1 - \lambda > 0$, $0 < \lambda < 1$. 

In summary: \( \lambda \) eigenvalue of \( I - M \) \( \Rightarrow \) \( 0 < \lambda < 1 \)
\[ \Rightarrow (1 - \lambda) \text{ eigenvalue of } M, \text{ and } 0 < 1 - \lambda < 1 \]
\[ \Rightarrow \frac{1}{1 - \lambda} \text{ eigenvalue of } M^{-1}, \text{ and } \frac{1}{1 - \lambda} > 1 \]
\[ \Rightarrow \text{Eigenvalues of } M^{-1} - I > 0 \]
\[ \Rightarrow M^{-1} - I > 0 \Rightarrow M^{-1} > I. \]

Now, consider arbitrary \( N > M > 0 \).

Factor \( N = R^k, \ R > 0 \) and invertible.

Put \( Q = R^{-1} \) (so \( Q^T = R^{-1} \))

\[ 0 < M < N \Rightarrow 0 < R^{-1} M R^{-1} < R^{-1} N R^{-1} = I \]

Take inverses:
\[ 0 > R M^{-1} R > I \]
\[ \Rightarrow 0 > R^T (R M^{-1} R) R^{-1} > R^T I R^{-1} = R^{-1} = N^{-1} \]
\[ \Rightarrow 0 > M^{-1} > N^{-1}. \]

Caveat: The product of self-adjoint maps is in general, not self-adjoint.

Example: Let \( A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}, \ B = \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}, \ x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \).
\[ A x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \ B x = \begin{pmatrix} -2 \\ 5 \end{pmatrix}, \ (x, A B x) = (A x, B x) = 3. \]
Def: If $A$, $B$ are self-adjoint, define their symmetrized product as $S = AB + BA$.

Note that $(x, Sx) = (x, ABx) + (x, BAx) = (Ax, Bx) + (Bx, Ax)$.

In the real case, $(x, Sx) = 2 (Ax, Bx)$.

In the example above, $AB + BA = \begin{pmatrix} -6 & 0 \\ 0 & 42 \end{pmatrix}$.

That is, it's false that $A > 0$, $B > 0 \Rightarrow AB + BA > 0$.

But a similar statement is true:

**Theorem 9.4:** Let $A, B$ be self-adjoint. If $A > 0$ and $AB + BA > 0$, then $B > 0$.

**Proof:** Define $B(t) = B + tA$. (Note: We must show $B(0) > 0$.)

**Claim 1:** The symmetrized product of $A$ and $B(t)$ is positive for $t > 0$.

$$S(t) = A B(t) + B(t) A = A(B + tA) + (B + tA) A = AB + BA + 2tA^2 = S + 2tA^2 > 0$$

**Claim 2:** $\exists T$ s.t. if $t > T$, then $B(t) > 0$.

$$(x, B(t)x) = (x, (B + tA)x) = (x, Bx) + t(x, Ax) \quad [\text{assume } \|x\| = 1]$$

Recall: $\cdot (x, Ax) = \lambda_{\text{min}} \|x\|^2 = \lambda_{\text{min}}$

$\cdot \|x\|^2 = \|Bx\| \leq \|B\| \|x\|^2 = \|B\|$

Together, if \( \|x\| = 1 \), then

\[
(x, B(t)x) = t(x, Ax) + (x, Bx) \\
\geq t \lambda_{\text{min}} - \|B\| > 0 \quad \text{if} \quad t > \frac{\|B\|}{\lambda_{\text{min}}} := T
\]

**Claim 3:** \( B = B(0) > 0 \).

If not, then for some \( t_0 \), \( 0 \leq t_0 \leq \frac{\|B\|}{\lambda_{\text{min}}} \) such that \( B(t_0) \) is on the boundary of the set of positive maps, i.e., \( B(t_0) \geq 0 \) but \( B(t_0) 
eq 0 \).

Such a \( B(t_0) \) has \( \lambda = 0 \), so \( B(t_0)y = 0 \) for some \( y \neq 0 \).

However, \((y, S(t_0)y) = (Ay, B(t_0)y) + (B(t_0)y, Ay) = 0. \)

Thus \( B > 0 \).

**Corollary 9.5:** If \( 0 < M < N \), then \( 0 < \sqrt{M} < \sqrt{N} \).

**Proof:** Put \( A(t) = M + t(N - M) \).

For \( 0 \leq t \leq 1 \), \( A(t) = (1-t)M + tN > 0 \), \( \dot{A}(t) = N - M > 0 \).

Thus, we can define \( R(t) = \sqrt{A(t)} \) for \( 0 \leq t \leq 1 \).
Since $A = R^2$, $\dot{A} = RR + RR$ (symmetrized product of $R \in \mathbb{R}$)

[See Lax Ch. 9 for the "product rule" of linear map derivative]

We know $\dot{A} = N-M > 0 \implies R > 0$ on $[0, 1]$.

Claim: $R(t)$ is an increasing function on $[0, 1]$.

For any $x \neq 0$, $\frac{d}{dt} (x, Rx) = (x, R_x) > 0$.

By calculus (see Lax Ch. 9), $(x, R(t)x)$ is increasing.

$\implies (x, R(s)x) < (x, R(t)x)$ for $s < t$

$\implies R(t)-R(s) > 0 \implies R(t) > R(s)$. \[\square\]

In particular $R(0) < R(1)$

and $R(0) = \sqrt{A(0)} = \sqrt{M}$, $R(1) = \sqrt{A(1)} = \sqrt{N} \implies \sqrt{M} < \sqrt{N}$. \[\square\]

Def: A real-valued function $f(s), s > 0$ is a monotone matrix function (mmf) if for all self-adjoint mappings,

$0 < M < N \implies f(M) < f(N)$.

Recall that $f(L) = \sum_{i=1}^{n} f(\lambda_i) P_i$ where $L = \sum_{i=1}^{n} \lambda_i P_i$.

Examples:

(i) $f(s) = \frac{-1}{s}$ is a mmf (immediate from Theorem 9.3;

$0 < M < N \implies M^{-1} > N^{-1} > 0 \implies -M^{-1} < -N^{-1} < 0$.)
(ii) $f(s) = s^{1/2}$ is a mmf (by Corollary 10.5).

(iii) $f(s) = s^2$ is not a mmf.

Take any $A, B > 0$ with $S = AB + BA \neq 0$.

Claim: For small $t$, if $M = A$, $N = A + tB$, then $0 < M < N$ but $M^2 \neq N^2$.

Why: $N^2 = A^2 + t(AB + BA) + t^2 B^2$ with $t$ negligible for $t \approx 0$.

So $N^2 = M^2 + [\text{something non-positive}] \neq M^2$.

(iv) $f(s) = s^{-2k}$ and $f(s) = \log s$ are mmf. (HW).

Additionally, positive multiples, sums, and limits of mmf's are mmf's.

For example, $-\sum \frac{m_j}{s+t_j}$, $m_j > 0$, $t_j > 0$ is an mmf, and so is $f(s) = aS + b - \int_0^1 \frac{dM(t)}{s+t}$, $a > 0$, $b \in \mathbb{R}$, (1.9) and $M(t)$ non-negative measure for which integral converges.

In fact, every mmf has the form of (1.9) (Theorem of C. Loewner — very non-trivial!).
Surprisingly, functions of the form (\( \ast \)) are easy to characterize:

**Theorem (Herglotz, Riesz):** Every function \( f \) which is analytic on the upper half-plane with \( \text{Im}(f) > 0 \) there, and \( \text{Im}(f) = 0 \) on the positive real axis, has the form (\( \ast \)). Conversely, every function of the form (\( \ast \)) can be extended to be analytic on the upper half-plane, with \( \text{Im}(f) > 0 \) there.

**Proof:** See Lax's book "Functional analysis."

**How to construct (all) positive matrices:**

**Def.** Let \( f_1, \ldots, f_m \) be a sequence of vectors in a Euclidean space \( X \). Define the \( m \times m \) matrix \( G \), where \( G_{ij} = (f_i, f_j) \). This is called the **Gram matrix** of \( f_1, \ldots, f_m \).

**Theorem 9.6:**

(i) Every Gram matrix is nonnegative.

(ii) The Gram matrix of a set of linearly independent vectors is positive.

(iii) Every positive matrix is a Gram matrix!
Proof: (i), (ii): $(y, Gx) = \sum_{i,j} x_i \overline{G}_{ij} \bar{y}_j = \sum_{i,j} (f_i, f_j) x_i \bar{y}_j$

$= \left( \sum_{i=1}^n x_i f_i, \sum_{j=1}^n \bar{x}_j f_j \right) = \left\| \sum_{i=1}^n x_i f_i \right\|^2$

(iii) Let $H = (h_{ij})$ be positive, and define the nonstandard inner product by $(x, y)_H = (x, H y)$.

Note that the Gram matrix of $e_1, \ldots, e_m$ has $ij$-entry

$(e_i, e_j)_H = (e_i, H e_j) = h_{ij}$. $\square$

Example:

(i) Let $X = \{ f: [0,1] \to \mathbb{R} \}$, $(f, g) = \int_0^1 f(t) g(t) \, dt$.

If $f_1 = 0$, $f_2 = t$, ..., $f_i = t^{i-1}$, then

$G = (G_{ij})$ where $G_{ij} = \frac{1}{i+j-1}$.

(ii) Define $(f, g) = \int_0^{2\pi} f(\theta) \overline{g(\theta)} w(\theta) \, d\theta$, $w: \mathbb{R} \to \mathbb{R}^n$.

If $f_j = e^{ij \theta}$, $j = -n, \ldots, n$, then the $(2n+1) \times (2n+1)$ corresponding Gram matrix is $G_{kj} = c_{k-j}$, where

$c_0 = \int w(\theta) e^{-i p \theta} \, d\theta$. 
Theorem 9.7 (Schur): Let \( A = (A_{ij}) \) and \( B = (B_{ij}) \) be positive matrices. Then \( M = (M_{ij}) := (A_{ij}B_{ij}) \) is positive.

Proof: Since \( A \) and \( B \) are Gram matrices (Thm 9.6)

write \( A_{ij} = (u_i, u_j) \), \( B_{ij} = (v_i, v_j) \) where

\( u_1, \ldots, u_n \) and \( v_1, \ldots, v_n \) are linearly independent.

Define \( g_i \in U \otimes V \) as \( g_i := u_i \otimes v_i \).

Note that \( (g_i, g_j) = (u_i, u_j)(v_i, v_j) = A_{ij}B_{ij} \) (Exercise)

Thus \( M \) is a Gram matrix, and so \( M > 0 \) by Thm 9.6. \( \Box \)

Singular value decomposition (SVD)

Big idea: If \( A : X \to Y \) is linear, write \( A = U \Sigma V^* \),

where \( U, V \) are orthogonal, \( \Sigma \) diagonal.

Special case: \( A = Q^*DQ \) if \( A \) is self-adjoint.

No good: \( A = P^*DP \) (e-vectors might not be orthogonal)


"Cartoon" of this:

\[ X = R_{A}^{*} \oplus N_{A} \]

\[ Y = R_{A} \oplus N_{A}^{*} \]

**Goal:** Find orthogonal bases of \( R_{A}^{*} \) and \( R_{A} \) so

\[
A \begin{bmatrix} x_{1} & x_{2} & \cdots & x_{r} \end{bmatrix} = \begin{bmatrix} y_{1} & y_{2} & \cdots & y_{r} \end{bmatrix} \begin{bmatrix} \bar{v}_{1} \\ \bar{v}_{2} \\ \vdots \\ \bar{v}_{r} \end{bmatrix}
\]

Once we have this, we can extend these bases to bases of all of \( X \) and \( Y \).

\[
A \begin{bmatrix} x_{1} & \cdots & x_{r} & x_{r+1} & \cdots & x_{n} \end{bmatrix} = \begin{bmatrix} y_{1} & \cdots & y_{r} & y_{r+1} & \cdots & y_{m} \end{bmatrix} \begin{bmatrix} \bar{v}_{1} \\ \bar{v}_{2} \\ \vdots \\ \bar{v}_{r} \\ \bar{v}_{r+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}
\]

**Question:** How to find \( U, V, \Sigma \)?

(It's not even clear why there should exist!)
Recall that $A^*A$ and $AA^*$ are self-adjoint (real eigenvalues) and nonnegative (all eigenvalues $\geq 0$).

$$A^*A = (V\Sigma^*U^*)(U\Sigma V^*) = V \begin{bmatrix} \sigma_1^2 \\ & \sigma_2^2 \\ & & \ddots \\ & & & \sigma_k^2 \end{bmatrix} U^*.$$  

$$AA^* = (U\Sigma V^*)(V\Sigma^*U^*) = U \begin{bmatrix} \sigma_1^2 \\ & \sigma_2^2 \\ & & \ddots \\ & & & \sigma_k^2 \end{bmatrix} U^*.$$  

So:  
$V = \text{matrix of eigenvectors of } A^*A$  
$U = \text{matrix of eigenvectors of } AA^*$  
$\Sigma = \text{matrix of square roots of eigenvalues of } A^*A \text{ (or } AA^*).$  

**Pseudoinverse**

Let $A: X \to Y$ be linear.

Recall:  
$X = \mathbb{R}_A^* \oplus N_A \xrightarrow{A} \mathbb{R}_A \oplus N_{A^*} = Y$

and the restriction $\mathbb{R}_A^* \xrightarrow{A} \mathbb{R}_A$ is a bijection.

**Big idea:** Even if $A$ is noninvertible, it may have a left, right, or pseudoinverse.
Case 1: A has a 2-sided inverse:

\[ AA^{-1} = I = A^{-1}A \]

- full rank \( r = n = m \)
- \( N_A = \{ 0 \} \)
- \( Ax = b \) has 1 solution

Case 2: A has a left-inverse.

- \( N_A = \{ 0 \} \) "full column rank"
- \( Ax = b \) has 0 or 1 solution.
- \( A^*A \) is invertible:

\[
(A^*A)^{-1} A^* A = I_{nm}
\]

Case 3: A has a right-inverse.

- \( N_{A^*} = \{ 0 \} \) "full row rank"
- \( Ax = b \) has no solutions.
- \( AA^* \) is invertible

\[
AA^*(AA^*)^{-1} = I_{mn}
\]
Case 4: The "general case" (any $A$!)

Want: Matrix $A^+$ "the pseudoinverse"

such that

$$AA^+ = \begin{bmatrix} I_{m \times m} & 0 \\ 0 & 0 \end{bmatrix}_{n \times m}$$

and

$$A^+A = \begin{bmatrix} I_{m \times m} & 0 \\ 0 & 0 \end{bmatrix}_{m \times m}$$

How to find $A^+$:

Write $A = U \Sigma V^*$ (the SVD).

Then $A^+ = V \Sigma^+ U^*$

where $\Sigma = \begin{bmatrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_r^{-1} & \\ & & & 0 \end{bmatrix}$

Note that

$$AA^+ = (U \Sigma V^*)(V \Sigma^+ U^*) = U \Sigma \Sigma^+ U = \begin{bmatrix} I_{m \times m} & 0 \\ 0 & 0 \end{bmatrix}_{m \times m}$$

$$A^+A = (V \Sigma^+ U^*)(U \Sigma V^*) = V \Sigma \Sigma^* V^* = \begin{bmatrix} I_{n \times n} & 0 \\ 0 & 0 \end{bmatrix}_{n \times n}.$$