Matrix Algebra

Ming-Ching Luoh

2005.9.12

Some Terminology

A matrix is a rectangular array of numbers, denoted as

$$
\begin{aligned}
A = [a_{ik}] = [A]_{ik} &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nK} \end{bmatrix} \\
&= \begin{bmatrix} a^1 \\ a^2 \\ \vdots \\ a^n \end{bmatrix} = [a_1, a_2, \cdots, a_K]
\end{aligned}
$$

where

$$a^i = [a_{i1}, a_{i2}, \cdots, a_{iK}]$$

is a row vector, and

$$a_k = \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{bmatrix}$$

is a column vector.

A couple of notations we need to know.

- dimemsions:
  The dimensions of matrix A is $n \times K$.
- square matrix:
  A is a square matrix if $n = K$.
- symmetric matrix:
  Square matrix A is symmetric if $a_{ik} = a_{ki}$ $\forall$ i and k.

- diagonal matrix:
  Square matrix A is diagonal if $a_{ik} = 0$ for $i \neq k$.
- scalar matrix:
  Diagonal matrix A is a scalar matrix if $a_{ii} = a_{kk}$ for $i \neq k$.
- identity matrix:
  Scalar matrix A is a identity matrix if $a_{ii} = 1 \, \forall \, i$, denoted as I.

- triangular matrix:
  Square matrix A is an upper triangular matrix if $a_{ik} = 0$ for $i > k$.
  Square matrix A is an lower triangular matrix if $a_{ik} = 0$ for $i < k$.
- zero matrix:
  Matrix A is a zero matrix if $a_{ik} = 0 \; \forall \; i$ and k.

Algebraic Manupulation of Matrices

Equality

If A and B have the same dimensions, then

$$A = B \text{ if } a_{ik} = b_{ik}, \forall i \text{ and } k.$$

Transposition

The transposition of matrix A, denoted as A', is

$$B = A' \Leftrightarrow b_{ik} = a_{ki}, \forall i \text{ and } k.$$

It follows that

$$(A')' = A$$

Therefore, if A is symmetric, then $A' = A$.

Addition

If A and B have the same dimensions, then the sum of A and B is matrix

$$C = A + B = [a_{ik} + b_{ik}], \text{ or } c_{ik} = a_{ik} + b_{ik}$$

Obviously, the sum of A and a zero matrix is $A + O = A$.

Substraction

Substraction of matrices is defined similar to addition.

$$C = A - B \Rightarrow c_{ik} = a_{ik} - b_{ik}$$

It follows that

$$
\begin{aligned}
A + B &= B + A \\
(A + B) + C &= A + (B + C) \\
(A + B)' &= [a_{ik} + b_{ik}]' = [a_{ki} + b_{ki}] \\
&= [a_{ki}] + [b_{ki}] = [a_{ik}]' + [b_{ik}]' \\
&= A' + B'
\end{aligned}
$$

Multiplication

Inner product of two vectors a and b is

$$\begin{aligned} a'b &= [a_1, a_2, \cdots, a_n] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \\ &= a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i \end{aligned}$$

This implies that $a'b = b'a$.

The multiplication of matrices is defined as

$$\underset{n\times T}{C} = \underset{n\times K}{A} \; \underset{K\times T}{B}$$

$$c_{ij} = a^i b_j = [a_{i1}, a_{i2}, \cdots, a_{ik}] \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = \sum_{k=1}^{K} a_{ik} b_{kj}$$

The number of columns of matrix A has to be equal to the number of rows of matrix B, i.e., A and B has to be **conformable.**

For example,

$$\underset{2\times 2}{AB} = \underset{2\times 3}{\begin{bmatrix} 1 & 3 & 2 \\ 4 & 5 & -1 \end{bmatrix}} \underset{3\times 2}{\begin{bmatrix} 2 & 4 \\ 1 & 6 \\ 0 & 5 \end{bmatrix}} = \begin{bmatrix} 5 & 32 \\ 13 & 41 \end{bmatrix}$$

It can be easily shown that the dimensions of BA is 3×3.

In general, $AB \neq BA$.
Therefore, the order of the multiplication is important.
It the case of AB, we say that B is **pre**-multiplied by A while A is **post**-multiplied by B.

It also can be easily seen that any matrix multiplied by identity matrix is still the matrix itself.

$$
\begin{aligned}
AI_K &= \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nK} \end{bmatrix}}_{n\times K} \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}}_{K\times K} \\
&= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nK} \end{bmatrix} = A
\end{aligned}
$$

And any matrix multiplied by a zero matrix is a zero matrix, $AO = O$.

There are a couple of important properties of multiplication.

- **associative law:** $(AB)C = A(BC)$.
  proof:
  Let $A : n \times K$, $B : K \times T$, $C : T \times Q$ and let $\underset{n\times T}{D} = AB$, then

$$d_{ij} = [AB]_{ij} = \sum_{k=1}^{K} a_{ik} b_{kj}.$$

Let $E = DC$, then

$$\begin{aligned}
e_{ij} &= \sum_{t=1}^{T} d_{it} c_{tq} = \sum_{t=1}^{T} \left(\sum_{k=1}^{K} a_{ik} b_{kt} \right) c_{tq} \\
&= \sum_{t=1}^{T} \sum_{k=1}^{K} a_{ik} b_{kt} c_{tq} = \sum_{k=1}^{K} \sum_{t=1}^{T} a_{ik} b_{kt} c_{tq} \\
&= \sum_{k=1}^{K} \left(\sum_{t=1}^{T} b_{kt} c_{tq} \right) a_{ik} = \sum_{k=1}^{K} f_{kq} a_{ik}, \;\; F \equiv BC \\
&= \sum_{k=1}^{K} a_{ik} f_{kq} = g_{iq}, \; G \equiv AF \\
DC &= AF \\
\therefore (AB)C &= A(BC)
\end{aligned}$$

- **distributive law:** $A(B+C) = AB + AC$.
  proof:
  Let $A : n \times K, B : K \times T, C : K \times T$ and $D = A(B+C)$, let $E = AB$, $F = AC$, then

$$\begin{aligned} d_{ij} &= \sum_{k=1}^{K} a_{ik}(b_{kj} + c_{kj}) \\ &= \sum_{k=1}^{K} (a_{ik}b_{kj} + a_{ik}c_{kj}) \\ &= e_{ij} + f_{ij} \\ \therefore A(B+C) &= AB + AC \end{aligned}$$

- **transpose of a product:** $(AB)' = B'A'$.
  proof: First thing is to check the dimensions.
  Let $C = (AB)'$ and $D = B'A'$. $A : n \times K$, $B : K \times T$, then C will be $T \times n$ which is the same as the dimensions of D. Then

$$
\begin{aligned}
c_{ij} &= [AB]_{ji} = \sum_{k=1}^{K} a_{jk} b_{ki} \\
d_{ij} &= (i\text{th row of } B') \cdot (j\text{th column of } A') \\
&= (i\text{th column of } B)'(j\text{th row of } A)'
\end{aligned}
$$

$$
\begin{aligned}
d_{ij} &= \begin{bmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{ki} \end{bmatrix}' [a_{j1}, a_{j2}, \cdots, a_{jk}]' \\
&= \sum_{k=1}^{K} b_{ki} a_{jk} = \sum_{k=1}^{K} a_{jk} b_{ki} \\
&= [AB]_{ji} = [(AB)']_{ij} = c_{ij}. \\
D = B'A' &= (AB)'
\end{aligned}
$$

- **idempotent matrix**
  Matrix M is idempotent if and only if
  $M^2 = MM = M$.
  If M is a symmetric idempotent matrix, then
  $M'M = MM = M$.

- **sum of values.**
  We can write sum of value in terms of matrix using a vector i which consists of all one's, i.e., $i = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$.

Suppose x is vector of n elements, or a $n \times 1$ matrix,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

$$\text{then } i'x = \sum_{i=1}^{n} x_i$$

Similarly,

$$\begin{aligned}
\bar{x} &= \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n} i' x \\
\sum_{i=1}^{n} x_i^2 &= x'x \\
\sum_{i=1}^{n} x_i y_i &= x'y = y'x
\end{aligned}$$

Vecotr Spaces

The K elements of a column vector $a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_K \end{bmatrix}$ can be viewed as the corridinates of a point in a K-dimensional space.

We can plot vectors in a two-dimensional plan.

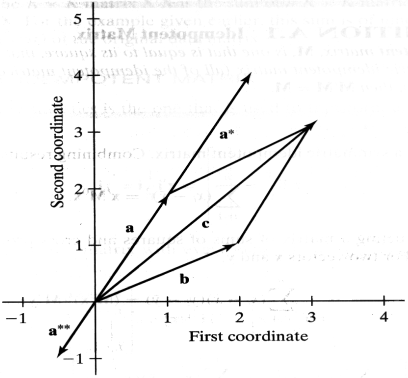

FIGURE A.1 Vector Space.

Let $a = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $a^* = \begin{bmatrix} 2 \\ 4 \end{bmatrix} = 2a$,
$a^{**} = \begin{bmatrix} -\frac{1}{2} \\ -1 \end{bmatrix} = -\frac{1}{2}a$,
and $c = a + b = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$.
We say that a^* and a^{**} are scalar multiplications of vector a. The 2-dimensional plan is denoted as R^2.

R^2 has two important properties.

- R^2 is **closed under scalar multiplication**: every scalar multiple of a vector in the plane is also in the plane.
- R^2 is **closed under addition**: the sum of any two vectors in the plane if always a vector in the plane.

Definition

Vector Space. A vector space is any set of vectors that is closed under scalar multiplication and addition.

Definition

Length of a Vector. The length, or norm, of a vector a is

$$\|a\| = \sqrt{a'a}$$

Definition

Orthogonal Vectors. Two vectors a and b are orthogonal, written as $a \perp b$, if and only if

$$a'b = b'a = 0$$

Theorem

The Cosine Law. *The angle* θ *between two vectors* a *and* b *satisfies*

$$cos\theta = \frac{a'b}{\|a\| \cdot \|b\|}$$

proof:

Note that the vector connecting the end points of a and b is $a - b$. Then

$$\begin{aligned} \|a-b\|^2 &= (a-b)'(a-b) = (a'-b')(a-b) \\ &= a'a + b'b - 2a'b = \|a\|^2 + \|b\|^2 - 2a'b \end{aligned}$$

From a simple diagram of a, b and $a - b$.

$$\begin{aligned} & \|a - b\|^2 \\ = & (\|a\|sin\theta)^2 + (\|a\|cos\theta - \|b\|)^2 \\ = & \|a\|^2 sin^2\theta + \|a\|^2 cos^2\theta + \|b\|^2 - 2\|a\|\|b\|cos\theta \\ = & \|a\|^2 + \|b\|^2 - 2\|a\|\|b\|cos\theta \end{aligned}$$

Therefore,

$$cos\theta = \frac{\|a\|^2 + \|b\|^2 - \|a - b\|^2}{2\|a\|\|b\|} = \frac{a'b}{\|a\|\|b\|}$$

- When a and b are orthogonal, $a'b = 0$, then $cos\ \theta = 0, \theta = \frac{\pi}{2}$.
- Since $-1 \leq cos\ \theta \leq 1$, it follows immediately that

$$a'b = \|a\| \cdot \|b\| cos\ \theta \leq \|a\| \cdot \|b\|$$

This is called Cauchy-Schwartz inequality.

Linear Combinations of Vectors and Basis Vectors

It can be shown that any vector in R^2 could be obtained as a linear combination of a and b, i.e.,

$$c = \alpha_1 a + \alpha_2 b, \alpha_1, \alpha_2 \in R$$

Then a and b are called **basis vectors** of R^2.

Definition

Basis Vectors. A set of vectors in a vector space is a basis for that vector if any vector in the vector space can be written as a linear combination of them.

K vectors are required to form a basis for R^K.
From the previous example of vectors, a and b can be basis vectors for R^2, but a and a^* can not be basis vectors.

The association among vectors can be classfied as linear dependence and linear indenpendence.

Definition

Linear Dependence. A set of vectors is linearly dependent if any one of the vectors in the set can be written as a linear combination of the others.

Since a^* is a multiple of a, a, and a^* are linearly dependent. Also $c = a + b$, a, b and c are linear dependent.

Definition

Linear Independence. A set of vectors $\{a_1, \cdots, a_K\}$ is linearly independent if and only if the only solution to

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_K a_K = 0$$

is $\alpha_1 = \alpha_2 = \cdots = \alpha_K = 0.$

Definition

Basis for a Vector Space. A basis for a vector space of K dimensions is any set of K **linearly independent** vectors in that space.

A set of more than K vectors in a K-dimensional space must be linearly dependent.

Rank of a Matrix

Definition

Column Space. The column space of a matrix is the vector space that is spanned by its column vectors.

Definition

Column Rank. The column rank of a matrix is the dimension of the vector space that is spanned by its columns.

Row space and row rank are defined similarly.

It follows that the column rank of a matrix is equal to the **largest** number of linearly independent column vectors it contains.

For example,

$$A = \begin{bmatrix} 1 & 5 & 6 \\ 2 & 6 & 8 \\ 7 & 1 & 8 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 1 & 5 \\ 6 & 4 & 5 \\ 3 & 1 & 4 \end{bmatrix}, C = \begin{bmatrix} 1 & 5 & 6 & 3 \\ 2 & 1 & 4 & 1 \\ 3 & 5 & 5 & 4 \end{bmatrix}$$

Since $a_3 = a_1 + a_2$, the column rank of A is 2, while the column rank of B is 3.

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 1 & 5 \\ 6 & 4 & 5 \\ 3 & 1 & 4 \end{bmatrix}, C = \begin{bmatrix} 1 & 5 & 6 & 3 \\ 2 & 1 & 4 & 1 \\ 3 & 5 & 5 & 4 \end{bmatrix}$$

- Since the column space of C is at most R^3, one of the column vector of C can be written as the linear combination of the others.
- Therefore, the column rank of C is the same as the column rank of B. But, the columns of C is in fact the rows of B. It follows that row rank of B equals the column rank of C.

Theorem

Equality of Row and Column Rank. *The column rank and row rank of a matrix are equal.*

Matrix A is said to have **full column rank** if its column rank equals to its number of column.
Similar definition applies to full row rank. Therefore, A has **short** column rank while B has **full** column rank.

Since row and column rank are the same, we can simply say rank of a matrix and have the following result,

$$\begin{aligned} \text{rank}(A) &= \text{rank}(A') \\ &\leq \min(\text{number of rows}, \ \text{number of columns}) \end{aligned}$$

In a product matrix $C = AB$, where A is $n \times K$ and B is $K \times T$, then each column of C is

$$\begin{aligned} c_i &= \begin{bmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{ni} \end{bmatrix} = \begin{bmatrix} a^1 b_i \\ a^2 b_i \\ \vdots \\ a^n b_i \end{bmatrix} = \begin{bmatrix} a^1 \\ a^2 \\ \vdots \\ a^n \end{bmatrix} b_i = A b_i \\ &= [a_1, a_2, \cdots, a_k] \begin{bmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{Ki} \end{bmatrix} = \sum_{k=1}^{K} a_k b_{ki} \end{aligned}$$

c_i is a linear combination of columns of A, so each column of C is in the column space of A.

It is possible that the set of columns in C can span this space, but it is not possible for them to span a higher-dimensional space.
At best, they could be a full set of linearly independent vectors in A's column space. We conclude that the column rank of C could not be greater than that of A.

On the other hand, each row of C is a linear combination of rows of B because

$$\begin{aligned} c^i &= [c_{i1}, c_{i2}, \cdots, c_{iT}] = [a^i b_1, a^i b_2, \cdots, a^i b_K] \\ &= a^i [b_1, b_2, \cdots, b_K] = a^i B \\ &= [a_{i1}, a_{i2}, \cdots, a_{iK}] \begin{bmatrix} b^1 \\ b^2 \\ \vdots \\ b^K \end{bmatrix} = \sum_{k=1}^{K} a_{ik} b^k \end{aligned}$$

Therefore, row rank of C can not be greater than the row rank of B. Since row rank and column rank are always equal, we conclude that

$$\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B)).$$

A useful corollary is:
If A is $n \times K$ and B is a square matrix of rank K, then rank(AB)=rank(A).
proof: Note that B is of full rank and the fact that the ith column of AB is simply a linear combination of A's columns by B's column i, rank of matrixx AB is therefore equal to rank of A.

Another fact is

$$\text{rank}(A) = \text{rank}(A'A) = \text{rank}(AA').$$

proof: $\text{rank}(A'A) = \text{rank}(AA')$ since rank of matrix is equal to the rank of its transpose. And $\text{rank}(A'A) \leq \min(\text{rank}(A'), \text{rank}(A)) = \text{rank}(A)$. Similar argument can apply here. Since the ith column of AA' is a linear combination of A's columns by column i of A', or row i of A. Therefore, $\text{rank}(AA')$=$\text{rank}(A)$.

Determinant of a Matrix

The determinant of a $K \times K$ square matrix A can be written as expansion by **cofactors.**

$$\begin{aligned} |A| &= \sum_{j=1}^{K} a_{ij}(-1)^{i+j}|A_{ij}|, \; i = 1, \cdots, K \\ &= \sum_{i=1}^{K} a_{ij}(-1)^{i+j}|A_{ij}|, \; j = 1, \cdots, K \end{aligned}$$

where A_{ij} is the matrix obtained from A by deleting row i and column j. The determinant of A_{ij}, $|A_{ij}|$, is a **minor** of A. $(-1)^{i+j}|A_{ij}|$ is a **cofactor.**

Properties of determinants.

- $K = 1$, A is a scalar, the determinant of a scalar is the scalar itself.
- $K = 2$,

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} &= a_{11}(-1)^{1+1}|a_{22}| + (-1)^{2+1}a_{21}|a_{12}| \\ &= a_{11}(-1)^{1+1}|a_{22}| + (-1)^{1+2}a_{12}|a_{21}| \\ &= a_{11}a_{22} - a_{21}a_{12} \end{aligned}$$

- $K = 3$,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$\begin{aligned}
&= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\
&= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) \\
&\quad + a_{31}(a_{12}a_{23} - a_{22}a_{13}) \\
&= a_{11}a_{22}a_{33} + a_{21}a_{13}a_{32} + a_{31}a_{12}a_{23} - a_{11}a_{32}a_{23} \\
&\quad - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13}
\end{aligned}$$

- The determinant of a matrix when $K = 2$ can be regarded as the area spanned by the two column vectors.
- The area spanned by a and $r \cdot a$ is zero, where r is a scalar.
- The determinant of a matrix is nonzero if and only if it has full rank, or the columns are linearly independent.
- A square matrix with non-zero determinant is said to be non-singular. Otherwise, it is singular.

- If D is a diagonal matrix with d_i being the diagonal element, then $|D| = \prod_{i=1}^{K} d_i$.
- $|cD| = c^K |D|$, where c is a scalar.
- $|AB| = |A||B|$,

Solutions of a System of Linear Equation

Consider the set of n linear equations,

$$\underset{K\times K}{A}\ \underset{K\times 1}{x} = \underset{K\times 1}{b},$$

where A is a $K \times K$ square matrix, x is a $K \times 1$ vector, and b is also a $K \times 1$ vector. We try to solve K unknowns from K equations.

- If $b = 0$, this is called a homogeneous system. A non-zero solution exists if and only if A does not have full rank. In other words, $|A| = 0$.
- If $b \neq 0$, this is called a non-homogeneous equation system. Then a solution exist if $|A| \neq 0$.

Inverse Matrices

Solve the system of $Ax = b$, we need to find a matrix B such that $BA = I$, then $BAx = Ix = x = Bb$.
If such B exists, then it is the **inverse** of A, denoted as $B = A^{-1}$. The solution is

$$x = A^{-1}b.$$

In the case of 2×2, we can get the inverse by solving the equations directly. We have

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$
$$= \frac{1}{|A|} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

Definition

Nonsingular Matrix. A matrix whose inverse exists is nonsingular.

For the higher-dimensional matrix, the general formula for an inverse matrix is

$$a^{ij} = \frac{|C_{ji}|}{|A|},$$

where a^{ij} is the ijth element of A^{-1}, $|C_{ji}|$ is the jith cofactor of A. We can check the simple case of 2×2 by this general formula.
For A to be nonsingular, $|A|$ must be nonzero.

Some computational results involving inverses are

- $|A^{-1}| = \frac{1}{|A|}$,
  since $AA^{-1} = I$, then $|AA^{-1}| = |A||A^{-1}| = |I| = 1$.
- $(A^{-1})^{-1} = A$,
  since $AA^{-1} = I$, from the definition of inverse, we can find B such that $BA^{-1} = I$, then B is the inverse of A^{-1}. In this case, $B = A$. Therefore, $(A^{-1})^{-1} = A$.
- $(A^{-1})' = (A')^{-1}$,
  since $AA^{-1} = I$, then
  $(AA^{-1})' = (A^{-1})'A' = I' = I$. Therefore,
  $(A')^{-1} = (A^{-1})'$.

- If A is symmetric, then A^{-1} is symmetric.
  since A is symmetric, then $A' = A$. From the result stated above, $(A^{-1})' = (A')^{-1} = A^{-1}$. Therefore, A^{-1} is symmetric.
- When both inverses exists, $(AB)^{-1} = B^{-1}A^{-1}$.
  Let $X = (AB)^{-1}$, then $XAB = I$ from the definition of inverse. Since $XAB = (XA)B = I$, then $XA = B^{-1}$. Post-multiply both sides by A^{-1}, we have $XAA^{-1} = B^{-1}A^{-1}$, or $X = B^{-1}A^{-1}$.
- $(ABC)^{-1} = C^{-1}(AB)^{-1} = C^{-1}B^{-1}A^{-1}$.

Partitioned Matrices

$$
\begin{aligned}
A &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \\
A + B &= \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{bmatrix} \\
AB &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \\
&= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}
\end{aligned}
$$

Two cases frequently encountered are of the form

$$\begin{aligned}
\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}' \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} &= [A_1' A_2'] \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \\
&= [A_1' A_1 + A_2' A_2] \\
\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}' \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} &= \begin{bmatrix} A_{11}' & 0 \\ 0 & A_{22}' \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \\
&= \begin{bmatrix} A_{11}' A_{11} & 0 \\ 0 & A_{22}' A_{22} \end{bmatrix}
\end{aligned}$$

Determinants of Partitioned Matrices

$$\begin{aligned}
\begin{vmatrix} A_{11} & 0 \\ 0 & A_{22} \end{vmatrix} &= |A_{11}| \cdot |A_{22}| \\
\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} &= |A_{22}| \cdot |A_{11} - A_{12}A_{22}^{-1}A_{21}| \\
&= |A_{11}| \cdot |A_{22} - A_{21}A_{11}^{-1}A_{12}|
\end{aligned}$$

Inverse of Partitioned Matrices

$$\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & A_{22}^{-1} \end{bmatrix}$$

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} A_{11}^{-1}(I + A_{12}F_2A_{21}A_{11}^{-1}) & -A_{11}^{-1}A_{12}F_2 \\ -F_2A_{21}A_{11}^{-1} & F_2 \end{bmatrix}$$

where $F_2 = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}$.
This can be checked by pre-multiply the above expression by A and get an identity matrix I.

Kronecker Product

$$\underset{n\times K}{A} \otimes \underset{m\times L}{B} = \underset{nm\times KL}{\begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1K}B \\ a_{21}B & a_{22}B & \cdots & a_{2K}B \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nK}B \end{bmatrix}}$$

Properties of Kronecker product

- $(A \otimes B)(C \otimes D) = AC \otimes BD$,
  pf: Let $A : n \times K$, $B : m \times L$, $C : K \times P$ and $D : L \times Q$.
  Therefore, $A \otimes B : nm \times KL$, $C \otimes D : KL \times PQ$, $AC : n \times P$, and $BD : m \times Q$. The dimensions of the statement is correct.

$$
\begin{aligned}
&(A \otimes B)(C \otimes D) \\
= &\begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1K}B \\ a_{21}B & a_{22}B & \cdots & a_{2K}B \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nK}B \end{bmatrix} \begin{bmatrix} c_{11}D & c_{12}D & \cdots & c_{1P}D \\ c_{21}D & c_{22}D & \cdots & c_{2P}D \\ \vdots & \vdots & \vdots & \vdots \\ a_{K1}D & c_{K2}D & \cdots & c_{KP}D \end{bmatrix} \\
= &\begin{bmatrix} a^1c_1BD & a^1c_2BD & \cdots & a^1c_PBD \\ a^2c_1BD & a^2c_2BD & \cdots & a^2c_PBD \\ \vdots & \vdots & \vdots & \vdots \\ a^nc_1BD & a^nc_2BD & \cdots & a^nc_PBD \end{bmatrix}
\end{aligned}
$$

$$(A \otimes B)(C \otimes D)$$

$$= \begin{bmatrix} a^1 c_1 BD & a^1 c_2 BD & \cdots & a^1 c_P BD \\ a^2 c_1 BD & a^2 c_2 BD & \cdots & a^2 c_P BD \\ \vdots & \vdots & \vdots & \vdots \\ a^n c_1 BD & a^n c_2 BD & \cdots & a^n c_P BD \end{bmatrix}$$

$$= \begin{bmatrix} a^1 c_1 & a^1 c_2 & \cdots & a^1 c_P \\ a^2 c_1 & a^2 c_2 & \cdots & a^2 c_P \\ \vdots & \vdots & \vdots & \vdots \\ a^n c_1 & a^n c_2 & \cdots & a^n c_P \end{bmatrix} BD = AC \otimes BD$$

- $[A \otimes B]^{-1} = [A^{-1} \otimes B^{-1}]$, since $[A \otimes B][A^{-1} \otimes B^{-1}] = I \otimes I = I$.
- If A is $M \times M$ and B is $n \times n$, then $|A \otimes B| = |A|^n |B|^M$. This can be shown by direct expansion of the definition of determinant.

- $(A \otimes B)' = A' \otimes B'$,

proof:

$$
\begin{aligned}
& (A \otimes B)' \\
= & \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1K}B \\ a_{21}B & a_{22}B & \cdots & a_{2K}B \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nK}B \end{bmatrix}' = \begin{bmatrix} a_{11}B' & a_{21}B' & \cdots & a_{K1}B' \\ a_{12}B' & a_{22}B' & \cdots & a_{K2}B' \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n}B' & a_{2n}B' & \cdots & a_{Kn}B' \end{bmatrix} \\
= & \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{K1} \\ a_{12} & a_{22} & \cdots & a_{K2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{Kn} \end{bmatrix} \otimes B' = A' \otimes B'
\end{aligned}
$$

- $\text{trace}(A \otimes B) = \text{trace}(A)\,\text{trace}(B)$,

proof:
Let $A : K \times K$, $B : L \times L$,

$$\begin{aligned}
& \text{trace}(A \otimes B) \\
= \ & \text{trace} \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1K}B \\ a_{21}B & a_{22}B & \cdots & a_{2K}B \\ \vdots & \vdots & \vdots & \vdots \\ a_{K1}B & a_{K2}B & \cdots & a_{KK}B \end{bmatrix} \\
= \ & \text{trace}(a_{11}B) + \text{trace}(a_{22}B) + \cdots + \text{trace}(a_{KK}B) \\
= \ & a_{11}\,\text{trace}(B) + a_{22}\,\text{trace}(B) + \cdots + a_{KK}\,\text{trace}(B) \\
= \ & (a_{11} + a_{22} + \cdots + a_{KK})\,\text{trace}(B) \\
= \ & \text{trace}(A)\,\text{trace}(B)
\end{aligned}$$