MA1101R Cheat Sheet

Austin Santoso

Chapter 2 – Matrices

Theorem 2.2.22

Let A be an m x n matrix

1. $(A^T)^T = A$
2. if be is also m x n matrix, then $(A + B)^T = A^T + B^T$
3. if c is a scalar, then $(cA)^T = cA^T$
4. if B is n x p matrix, then $(AB)^T = B^TA^T$

Matrix Inverse -

Let A be a square matrix of order n.
Then A is said to be invertible if there exists a square matrix B of order n such that
AB= I and **BA**= I.
The matrix **B** here is called an inverse of **A**.
A square matrix is called singular if it has no inverse.

Remark 2.3.4 – Cancellation laws

1. Cancellation laws for matrix multiplication:
   Let $\boldsymbol{A}$ be an invertible $m \times m$ matrix.
   (a) If $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ are $m \times n$ matrices such that $\boldsymbol{AB}_1 = \boldsymbol{AB}_2$, then $\boldsymbol{B}_1 = \boldsymbol{B}_2$.
   (b) If $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ are $n \times m$ matrices such that $\boldsymbol{C}_1\boldsymbol{A} = \boldsymbol{C}_2\boldsymbol{A}$, then $\boldsymbol{C}_1 = \boldsymbol{C}_2$.
2. If $\boldsymbol{A}$ is not invertible, the cancellation laws may not hold.
   For example, let $\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$, $\boldsymbol{B}_1 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ and $\boldsymbol{B}_2 = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}$. Then $\boldsymbol{AB}_1 = \boldsymbol{AB}_2$ but $\boldsymbol{B}_1 \neq \boldsymbol{B}_2$.

Theorem 2.3.9 – properties of inverse

Let $\boldsymbol{A}$, $\boldsymbol{B}$ be two invertible matrices and c a nonzero scalar.

1. $c\boldsymbol{A}$ is invertible and $(c\boldsymbol{A})^{-1} = \frac{1}{c}\boldsymbol{A}^{-1}$.
2. $\boldsymbol{A}^T$ is invertible and $(\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T$.
3. $\boldsymbol{A}^{-1}$ is invertible and $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$.
4. $\boldsymbol{AB}$ is invertible and $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

(Remark 2.3.10)
By Part 4, if $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_k$ are invertible matrices, then $\boldsymbol{A}_1\boldsymbol{A}_2 \cdots \boldsymbol{A}_k$ is invertible and
$$(\boldsymbol{A}_1\boldsymbol{A}_2 \cdots \boldsymbol{A}_k)^{-1} = \boldsymbol{A}_k^{-1} \cdots \boldsymbol{A}_2^{-1}\boldsymbol{A}_1^{-1}.$$

Theorem 2.4.7 – Invertible matrices equivalence

Let A be a n×n matrix. The following statements are equivalent:
1. A is invertible.
2. The linear system Ax = 0 has only the trivial solution.
3. The reduced row-echelon form of A is an identity matrix. -> No zero rows
4. A can be expressed as a product of elementary matrices.
5. det(A) $\neq$ 0.
6. The rows of A form a basis for n.
7. The columns of A form a basis for n.
8. Rank(A) = n
9. 0 is not an eigenvalue of A

Theorem 2.5.10

det($\mathbf{A}^T$) = det($\mathbf{A}$)

Theorem 2.5.15 - Effect of elementary row operations on the determinant

Let $\boldsymbol{A} = (a_{ij})$ be an $n \times n$ matrix.

$\boldsymbol{A} \xrightarrow{kR_i} \boldsymbol{B}_1$ $\det(\boldsymbol{B}_1) = k\det(\boldsymbol{A})$

$\boldsymbol{A} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{B}_2$ $\det(\boldsymbol{B}_2) = -\det(\boldsymbol{A})$

$\boldsymbol{A} \xrightarrow{R_j + kR_i} \boldsymbol{B}_3$ $\det(\boldsymbol{B}_3) = \det(\boldsymbol{A})$

Furthermore, if $\boldsymbol{E}$ is an elementary matrix of the sa size as $\boldsymbol{A}$, then $\det(\boldsymbol{EA}) = \det(\boldsymbol{E})\det(\boldsymbol{A})$.

Theorem 2.5.22.3

$\det(\mathbf{A}^{-1}) = \frac{1}{\det(A)}$

Adjoints 2.5.24

Let $\boldsymbol{A}$ be a square matrix of order n.
The (classical) adjoint of $\boldsymbol{A}$ is the $n \times n$ matrix

$$\mathrm{adj}(\boldsymbol{A}) = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}^T = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}$$

where A_{ij} which is the (i, j)-cofactor of $\boldsymbol{A}$.

If **A** is an invertible matrix,
the $\mathbf{A}^{-1} = \frac{1}{\det(A)}$ adj(**A**)

Theorem 2.5.27 - Cramer's Rule

Suppose $\boldsymbol{Ax} = \boldsymbol{b}$ is a linear system

where $\boldsymbol{A} = (a_{ij})_{n \times n}$, $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$.

Let $\boldsymbol{A}_i$ be the $n \times n$ matrix obtained from $\boldsymbol{A}$ by replacing the i^{th} column of $\boldsymbol{A}$ by $\boldsymbol{b}$,

i.e. $\boldsymbol{A}_i = \begin{bmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,i-1} & b_2 & a_{2,i+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{bmatrix}$.

If $\boldsymbol{A}$ is invertible, then the system has only one solution

$$\boldsymbol{x} = \frac{1}{\det(\boldsymbol{A})} \begin{bmatrix} \det(\boldsymbol{A}_1) \\ \det(\boldsymbol{A}_2) \\ \vdots \\ \det(\boldsymbol{A}_n) \end{bmatrix}.$$

Proof: $\boldsymbol{Ax} = \boldsymbol{b} \Leftrightarrow \boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} = \frac{1}{\det(\boldsymbol{A})}\mathrm{adj}(\boldsymbol{A})\,\boldsymbol{b}$.

$$\mathrm{adj}(\boldsymbol{A})\,\boldsymbol{b} = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} b_1A_{11} + b_2A_{21} + \cdots + b_nA_{n1} \\ b_1A_{12} + b_2A_{22} + \cdots + b_nA_{n2} \\ \vdots \\ b_1A_{1n} + b_2A_{2n} + \cdots + b_nA_{nn} \end{bmatrix}$$

So the solution to the system is

$$\boldsymbol{x} = \frac{1}{\det(\boldsymbol{A})} \begin{bmatrix} b_1A_{11} + b_2A_{21} + \cdots + b_nA_{n1} \\ b_1A_{12} + b_2A_{22} + \cdots + b_nA_{n2} \\ \vdots \\ b_1A_{1n} + b_2A_{2n} + \cdots + b_nA_{nn} \end{bmatrix}$$

where for $i = 1, 2, \ldots, n$,

$$\det(\boldsymbol{A}_i) = \begin{vmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,i-1} & b_2 & a_{2,i+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{vmatrix}$$

$\boxed{\text{cofactor expansion along the } i^{th} \text{ column}} \longrightarrow = b_1A_{1i} + b_2A_{2i} + \cdots + b_nA_{ni}$.

Chapter 3 – Vector Spaces

Discussion 3.2.5

To prove Span(S) = R^n → Use rref

$$\text{Let } \boldsymbol{A} = \begin{array}{c} \begin{array}{cccc} \boldsymbol{u}_1 & \boldsymbol{u}_2 & & \boldsymbol{u}_k \end{array} \\ \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{kn} \end{bmatrix} \end{array}.$$

1. If a row-echelon form of $\boldsymbol{A}$ has no zero row, then the linear system is always consistent regardless the values of $v_1, v_2, \ldots, v_n$ and hence $\mathrm{span}(S) = \mathbb{R}^n$.
2. If a row-echelon form of $\boldsymbol{A}$ has at least one zero row, then the linear system is not always consistent and hence $\mathrm{span}(S) \neq \mathbb{R}^n$.

Definition 3.3.2 – Subspaces

Let V be a subset of R^n
V is called a subspace of R^n

If V = Span(S), where S = {u_1, u_2, …. u_k} for u_1, u_2, …. $u_k \in R^n$

From theorem 3.2.9 – properties of subspace

Let V be a subspace of R^n

1. 0 ∈ V (V must contain the origin)
2. For any v_1, v_2, …, v_r ∈ V and c_1, c_2, …, c_r ∈ R, $c_1v_1 + c_2v_2 \ldots + c_rv_r \in V$

Vector space

We adopt the following conventions:

1. A set V is called a vector space if either $V = \mathbb{R}^n$ or V is a subspace of $\mathbb{R}^n$.
2. Let W be a vector space, say, $W = \mathbb{R}^n$ or W is a subspace of $\mathbb{R}^n$.
   A set V is called a subspace of W if V is a vector space and $V \subseteq W$,
   i.e. V is a subspace of $\mathbb{R}^n$ which lies completely inside W.

Basis –

Let V be a vector space
Let S = {u_1, u_2, …. u_k} a subset of V
Then S is called a Basis for V if

1. S is linearly independent
2. S spans V

[Most Effective Span]

Coordinate System

You must understand yourself

Theorem 3.6.7

If we want to check that S is a basis for V, and we know the dimension of V is K.
we only need to check any two of the three conditions:

(i) S is linearly independent;
(ii) S spans V;
(iii) |S| = k.

Transition matrix –

Convert the coordinates from one basis to another basis.
S = {u1, u2, u3 … uk}
T = {v1, v2, v3 … vk}

$[w]_T = P\,[w]_S$ -> convert a vector from basis S to basis T

$P = [[u1]_T\ [u2]_T \ldots [uk]_T]$

<u>Chapter 4 – Vector Spaces Associated with Matrices</u>

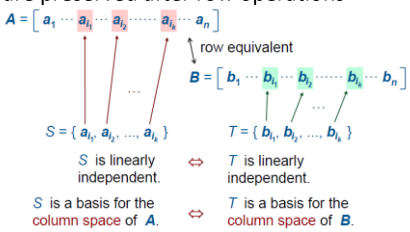
Remark 4.1.9 => Finding Basis of A using Row Vectors.

Let A be a matrix and R a row-echelon form of A.

<u>So, basis of A are the non-zero rows of R.</u>

Properties of column Vectors

Linear Independence of Column Vectors are preserved after row operations



Finding Basis of A using Column Vectors.

A -Gaussian Elimination -> R

- Column space of R =/= column space of A
- *<u>The basis for the column space of A can be obtained by taking columns of A that correspond to the pivot columns in R</u>*
- Every non pivot column is a linear combination of other columns.

Definition 4.2.3 – Rank

The rank of a matrix is the dimension of its row space or its column space

Theorem 4.2.8 – Ranks of product of matrices

Let A and B be m × n and m × p matrices respectively. Then:

rank(AB) ≤ min{ rank(A), rank(B) }.

Nullspace

The solution space of the homogeneous linear system Ax = 0is known as the nullspace of A.

The dimension of the nullspace of A is called the nullity of A and is denoted by nullity(A).

Nullity -> the non pivot column of R

The general solution to the homogeneous linear system $\mathbf{Ax} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} -s-t \\ s \\ -t \\ 0 \\ t \end{bmatrix} = s\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

where s, t are arbitrary parameters.

The reduced row-echelon form of $\mathbf{A}$:

$$\left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Thus $\{ (-1, 1, 0, 0, 0)^T, (-1, 0, -1, 0, 1)^T \}$ is a basis for the nullspace of $\mathbf{A}$.

Here nullity($\mathbf{A}$) = 2.

<u>Rank(A) + nullity(A) = no. of column of A</u>

<u>Chapter 5 – Orthogonality</u>

Dot Product is like matrix multiplication. Should know by heart.

Distance -> U1 . U1

Angle = $\cos^{-1}\left(\frac{u.v}{||u||\ ||v||}\right)$

Orthogonal – 90 degree between the two vectors (perpendicular)

Orthogonal sets are linearly independent

Orthonormal set: a set of vectors all of which are orthogonal with each other, and each have norm (length) 1.

Theorem 5.2.8.1 Orthogonal basis

It is easy to get the coordinate vector of w.

If S = {u1, u2, … uk}

W = $\frac{w.u1}{u1.u1}u1 + \frac{w.u2}{u2.u2}u2\ \ldots + \frac{w.uk}{uk.uk}uk$

$[w]_S = \left(\frac{w.u1}{u1.u1}, \frac{w.u2}{u2.u2}, \ldots\ \frac{w.uk}{uk.uk}\right)$

A vector u is said to be orthogonal to V if u is orthogonal to all vectors in V. (the Basis of V, or any spanning set). If u is orthogonal to these, then it is orthogonal to V.

Projection

Orthogonal bases & projections (Theorem 5.2.15.1)

Let V be a subspace of $\mathbb{R}^n$ and $\{ \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k \}$ an orthogonal basis for V.

Then for any $\mathbf{w} \in \mathbb{R}^n$,

$$\frac{\mathbf{w}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{w}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 + \cdots + \frac{\mathbf{w}\cdot\mathbf{u}_k}{\mathbf{u}_k\cdot\mathbf{u}_k}\mathbf{u}_k$$

is the projection of $\mathbf{w}$ onto V.

Proof: Define $\mathbf{p} = \frac{\mathbf{w}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{w}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 + \cdots + \frac{\mathbf{w}\cdot\mathbf{u}_k}{\mathbf{u}_k\cdot\mathbf{u}_k}\mathbf{u}_k \in V$ and $\mathbf{n} = \mathbf{w} - \mathbf{p}$.

Since $\mathbf{w} = \mathbf{n} + \mathbf{p}$ where $\mathbf{p}$ is a vector in V, to show that $\mathbf{p}$ is a projection of $\mathbf{w}$ onto V, it suffices to show $\mathbf{n}$ is orthogonal to V.

P is clearly in V

To proof n is orthogonal to V

To show $\mathbf{n}$ is orthogonal to V:

For $i = 1, 2, \ldots, k$,

$$\begin{aligned} \mathbf{n}\cdot\mathbf{u}_i &= (\mathbf{w}-\mathbf{p})\cdot\mathbf{u}_i \\ &= \mathbf{w}\cdot\mathbf{u}_i - \mathbf{p}\cdot\mathbf{u}_i \\ &= \mathbf{w}\cdot\mathbf{u}_i - \left(\frac{\mathbf{w}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{w}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 + \cdots + \frac{\mathbf{w}\cdot\mathbf{u}_k}{\mathbf{u}_k\cdot\mathbf{u}_k}\mathbf{u}_k\right)\cdot\mathbf{u}_i \\ &= \mathbf{w}\cdot\mathbf{u}_i - \frac{\mathbf{w}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}(\mathbf{u}_1\cdot\mathbf{u}_i) - \frac{\mathbf{w}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}(\mathbf{u}_2\cdot\mathbf{u}_i) - \cdots - \frac{\mathbf{w}\cdot\mathbf{u}_k}{\mathbf{u}_k\cdot\mathbf{u}_k}(\mathbf{u}_k\cdot\mathbf{u}_i) \\ &= \mathbf{w}\cdot\mathbf{u}_i - \frac{\mathbf{w}\cdot\mathbf{u}_i}{\mathbf{u}_i\cdot\mathbf{u}_i}(\mathbf{u}_i\cdot\mathbf{u}_i) \\ &= 0. \end{aligned}$$

So $\mathbf{n}$ is orthogonal to V.

The projection of u to V is p.

The vector p is the best approximation of u in V.

Theorem 5.2.19

Gram – Schmidt process

Make any basis into an orthogonal Basis

Let $\{ \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k \}$ be a basis for a vector space V.

Let $\mathbf{v}_1 = \mathbf{u}_1$,

$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\mathbf{u}_2\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1,$$

$$\mathbf{v}_3 = \mathbf{u}_3 - \frac{\mathbf{u}_3\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{u}_3\cdot\mathbf{v}_2}{\mathbf{v}_2\cdot\mathbf{v}_2}\mathbf{v}_2,$$

$$\vdots$$

$$\mathbf{v}_k = \mathbf{u}_k - \frac{\mathbf{u}_k\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{u}_k\cdot\mathbf{v}_2}{\mathbf{v}_2\cdot\mathbf{v}_2}\mathbf{v}_2 - \cdots - \frac{\mathbf{u}_k\cdot\mathbf{v}_{k-1}}{\mathbf{v}_{k-1}\cdot\mathbf{v}_{k-1}}\mathbf{v}_{k-1}.$$

Then $\{ \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \}$ is an orthogonal basis for V.

Furthermore, $\left\{ \frac{1}{||\mathbf{v}_1||}\mathbf{v}_1, \frac{1}{||\mathbf{v}_2||}\mathbf{v}_2, \ldots, \frac{1}{||\mathbf{v}_k||}\mathbf{v}_k \right\}$ is an orthonormal basis for V.

5.3 Best Approximations

Let $\mathbf{Ax} = \mathbf{b}$ be a linear system where $\mathbf{A}$ is an $m \times n$ matrix.

A vector $\mathbf{u} \in \mathbb{R}^n$ is called a least square solution to the linear system $\mathbf{Ax} = \mathbf{b}$ if

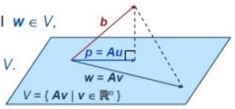
$$||\mathbf{b} - \mathbf{Au}|| \le ||\mathbf{b} - \mathbf{Av}|| \text{ for all } \mathbf{v} \in \mathbb{R}^n. \quad (\#)$$

Let $V = \{ \mathbf{Av} \mid \mathbf{v} \in \mathbb{R}^n \}$ and $\mathbf{p} = \mathbf{Au}$.

Then (#) is rewritten as

$$d(\mathbf{b}, \mathbf{p}) \le d(\mathbf{b}, \mathbf{w}) \text{ for all } \mathbf{w} \in V,$$

i.e. $\mathbf{p} = \mathbf{Au}$ is the best approximation of $\mathbf{b}$ onto V.



The best approximation to the equation Ax = b.

The least Square Solution is **u**

Where Au = p.

Where p is the span of the column space of A

To solve least square:

Alternatively, u is a solution to the equation

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}.$$

Using this method, we can find the projection of b onto A.

Since p = Ax. If x has infinitely many solution (an arbitrary parameter) take any solution to get Ax = p.

5.4 – Orthogonal Matrices

An orthogonal Matrix has the property $P^{-1} = P^T$

Given two orthonormal bases

E = {e1, e2, … ek}

S = {u1, u2, … uk}

The transition Matrix from E to S is P. The Transition matrix from S to E is Q.

P is Q^T

(By theorem 3.7.5) => $Q = P^{-1}$

So, $P^{-1} = P^T$

To find:

Theorem 5.4.6 – orthogonal matrices

1. A is orthogonal
2. The rows of A form an orthonormal basis for R^n
3. The columns of A form an orthonormal basis for R^n