§ 1.6 More on linear systems and invertible matrices

Theorem 1.6.1 A system of linear equations has zero, one, or infinitely many solutions. There are no other possibilities.

proof. It is clear that there are 3 mutually exclusive possibilities

(i) zero solution

(ii) unique solution

(iii) Two or more solutions.

Obviously there are examples of (i) and (ii). So it is enough to prove that if there are two or more solutions, then there must be infinitely many solutions.

Suppose ______ are two distinct solutions, i.e.

$$Ax = b \quad \text{and} \quad AY = b.$$

let t be an arbitrary real number, we have

$$A\big(\underbrace{tx + (1-t)Y}_{Y + t(x-Y)}\big) = A(tx) + A((1-t)Y)) = tAx + (1-t)AY = tb + (1-t)b = b.$$

Thus $tx + (1-t)Y$ is a solution for any $t \in \mathbb{R}$. Since the set of such vectors is infinite(*), the proof is done. ∎

(*) This is clear from the form $Y + t(x-Y)$ since $x - Y \neq 0$.

Theorem 1.6.2 If A is invertible $n \times n$ matrix, then for each $n \times 1$ matrix b, the system $Ax = b$ has a unique solution.

proof.

Note that x $A^{-1}b$ satisfies $Ax = A(A^{-1}b) = (AA^{-1})b = b$ which means that $x = A^{-1}b$ is a solution to $Ax = b$. I

On the other hand $Ax = b$ implies $x = A^{-1}b$ which says any solution of $Ax = b$ is of the form $x = A^{-1}b$, thus unique.

It occurs often that one has to solve a <u>set</u> of linear systems with the <u>same A</u> but with different right-hand sides b.

<u>Ex.</u> Solve $Ax = b$ and $Az = c$.

Of course we can use Gauss elimination separately, e.g.

$$[A \mid b] \longrightarrow [\mathrm{rref}(A) \mid \tilde{b}]$$

$$[A \mid c] \longrightarrow [\mathrm{rref}(A) \mid \tilde{c}]$$

This would be greatly inefficient since we are repeating the work done on A.

~~Rather, it is better to add c as a second right~~ hand side, then apply Gauss-Elimination

$$[A \mid b \mid c] \longrightarrow [\mathrm{rref}(A) \mid \tilde{b} \mid \tilde{c}]$$

<u>Ex.</u> Solve
$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 4 \\ 2x_1 + 5x_2 + 3x_3 &= 5 \\ x_1 \qquad + 8x_3 &= 9 \end{aligned} \quad \text{and} \quad \begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ 2x_1 + 5x_2 + 3x_3 &= 6 \\ x_1 \qquad + 8x_3 &= -6 \end{aligned}$$

$$\Rightarrow \left[\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 1 \\ 2 & 5 & 3 & 5 & 6 \\ 1 & 0 & 8 & 9 & -1 \end{array}\right] \longrightarrow \longrightarrow \cdots \rightarrow \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & -1 \end{array}\right]$$

$$\Rightarrow x = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad z = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}.$$

<u>Remark</u> The matrix inversion algorithm is very similar to this idea

$$\left[\begin{array}{c|ccc} & 1 & 0 & 0 \\ A & 0 & 1 & \\ & & & \ddots \\ & 0 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{c|c} I & A^{-1} \end{array}\right]$$

Remark Recall that we said that a square matrix A is invertible if a square matrix B can be found satisfying

$$AB = I \text{ and } BA = I.$$

Then we called B the inverse of A and denoted it by A^{-1}.

With the increased knowledge at hand, we will show that actually it is sufficient to verify only one of the above two conditions.

Theorem 1.6.3 Let A be a square matrix and suppose B is such that $AB = I$ or $BA = I$. Then A is invertible and $B = A^{-1}$.

proof

Suppose $AB = I$. We need to show $BA = I$. Consider the 2 cases

(i) B is invertible $\Rightarrow B = E_1 E_2 \cdots E_m$. Also $I = AB = AE_1 \cdots E_m$
$\Rightarrow A = E_m^{-1} \cdots E_1^{-1} \Rightarrow BA = E_1 E_2 \cdots E_m E_m^{-1} \cdots E_1^{-1} = I$ ✓

(ii) B is not invertible. Then by Thm. 1.5.3 $\exists\, x \neq 0$ such that $Bx = 0$. But then

$$x = Ix = (AB)x = A(Bx) = A0 = 0 \quad \times. \qquad \square$$

Remark We saw in Thm. 1.4.6 that the product of two invertible matrices is also invertible. This easily generalizes ... any finite product of invertible matrices is invertible.

Now we will show that the converse of this result is also true, i.e. if the product of two or more matrices is invertible, then each matrix in the product is invertible.

Theorem 1.6.5 Let A, B be two square matrices. If AB is invertible then A and B must be invertible.

proof

Suppose B is not invertible. Then $\exists\, x \neq 0$ such that $Bx = 0$. Now $(AB)x = A(Bx) = A0 = 0$. Since $x \neq 0$, AB cannot be invertible. $\times$ $\Rightarrow B$ is invertible. To show A is invertible, write $A = (AB)B^{-1}$. Thus A is invertible since both AB and B^{-1} are invertible. $\square$

Thus, we have shown that B is invertible. Then, we have

$A = (AB)B^{-1}$ and is invertible as the product of AB and B^{-1} which are invertible. $\boxed{\Xi}$

Theorem 1.6.4 The following statements are equivalent

as in Thm 1.5.3 $\left\{ \begin{array}{l} \text{(a) } A \text{ is invertible} \\ \text{(b) } Ax = 0 \text{ has only the trivial solution} \\ \text{(c) } \mathrm{rref}(A) = I \end{array} \right.$

(d) A is the product of elementary matrices

(e) $Ax = b$ has a unique solution for every $n \times 1$ matrix b.

Proof: It is enough to show (a) $\iff$ (e).

(a) $\Rightarrow$ (e)

A invertible means A^{-1} exists and thus $A^{-1}b$ is uniquely defined. Also, $A(A^{-1}b) = b$ Thus system has a solution. Also, if $Ax = b$, then $x = A^{-1}b$ which is unique

(e) $\Rightarrow$ (a). Consider the n systems

$$Ax^{(1)} = e_1 = \begin{bmatrix} 1 \\ 0 \\ | \\ 0 \end{bmatrix}, \quad Ax^{(2)} = e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ | \\ 0 \end{bmatrix}, \; \ldots, \; Ax^{(n)} = e_n = \begin{bmatrix} 0 \\ | \\ 0 \\ 1 \end{bmatrix}.$$

Each is uniquely solvable. Also, if we arrange the $n \times 1$ matrices $x^{(1)}, \ldots, x^{(n)}$ as columns:

$$X = \left[x^{(1)} \;\middle|\; x^{(2)} \;\middle|\; \cdots \;\middle|\; x^{(n)} \right] \quad \text{is } n \times n \text{ matrix,}$$

then

$$AX = \left[Ax^{(1)} \;\middle|\; Ax^{(2)} \;\middle|\; \cdots \;\middle|\; Ax^{(n)} \right] = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & | \\ | & 0 & & 0 \\ 0 & 0 & & 1 \end{bmatrix} = I.$$

$\boxed{\Xi}$

§1.7 Diagonal, Triangular and symmetric matrices

Defn. A square matrix A is diagonal if $a_{ij} = 0$ for $i \neq j$

Ex. $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -9 \end{bmatrix}$

Lemma A diagonal matrix $A = \begin{bmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{bmatrix}$

is invertible if and only if

$a_{ii} \neq 0$, $i = 1, \ldots, n$. Moreover, The inverse is also diagonal and

$$A^{-1} = \begin{bmatrix} 1/a_{11} & & & 0 \\ & 1/a_{22} & & \\ & & \ddots & \\ 0 & & & 1/a_{nn} \end{bmatrix}$$

Lemma Let A be diagonal. Then for any positive integer m

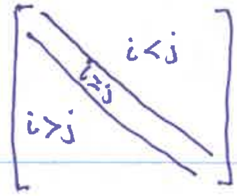
$$A^m = \begin{bmatrix} a_{11}^m & & & 0 \\ & a_{22}^m & & \\ & & \ddots & \\ 0 & & & a_{nn}^m \end{bmatrix}$$

Defn. A square matrix A is lower triangular if $a_{ij} = 0$ for $j > i$

A square matrix A is upper triangular if $a_{ij} = 0$ for $j < i$

Ex. $\begin{bmatrix} 1 & 0 & 0 \\ 4 & 0 & 0 \\ 5 & -7 & -3 \end{bmatrix}$ is lower triangular. Note: Can have zero(s) in the lower part

$\begin{bmatrix} 5 & -1 & 0 \\ 0 & 9 & 4 \\ 0 & 0 & 7 \end{bmatrix}$ is upper triangular.



Theorem 1.7.1 (a)

(a) The transpose of a <u>lower</u> triangular matrix is <u>upper triangular</u>.
The transpose of an <u>upper</u> triangular matrix is <u>lower triangular</u>.

(b) The product of two or more <u>lower</u> tri. matrices is <u>lower triangular</u>
" " " " " " <u>upper</u> tri. " " <u>upper</u> triangular

(c) A triangular (lower or upper) matrix is invertible if and only if the diagonal elements are all nonzero.

(d) The <u>inverse</u> of an <u>invertible</u> <u>lower</u> triangular matrix is <u>lower</u> tri.
The <u>inverse</u> of an <u>invertible</u> <u>upper</u> " " " <u>upper</u> tri.

<u>proof</u> (a) This is obvious from $(A^T)_{ij} = A_{ji}$

(b) Suppose A and B are lower triangular, i.e. $a_{ij} = b_{ij} = 0$, $i<j$

Note that $(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

we want to show that $(AB)_{ij} = 0$ for $\boxed{i<j}$

$$(AB)_{ij} = (a_{i1} \cdots a_{in}) \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix} = (a_{i1} \cdots a_{ii}\ 0 \text{ — } 0) \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_{jj} \\ \vdots \\ b_{nj} \end{pmatrix}$$

since $i<j$,

$$(AB)_{ij} = (a_{i1} \cdots a_{ii}\ 0 \cdots 0\ a_{ij} \cdots a_{in}) \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_{jj} \\ \vdots \\ b_{nj} \end{pmatrix} = 0 \checkmark$$

so AB is lower triangular.

Now suppose A, B are upper triangular.

$(AB)^T = B^T A^T$ = lower triangular as the product of two lower tri. matrices by the 1st part.

Thus,

$(AB)^T$ lower tri. $\Rightarrow$ AB upper tri. by (a).

(c) Suppose A is upper triangular and $a_{ii} \neq 0$ $i=1,\dots,n$.

It is clear that every column contains a leading nonzero. Using elementary row operations of type 2 and type 3, we can reduce A to rref and deduce that $\mathrm{rref}(A) = I$. Hence A is invertible.

Conversely Now suppose A is upper tri. and invertible. We want to show that $a_{ii} \neq 0$, $i=1,\dots,n$. To begin, we must have $a_{nn} \neq 0$; otherwise we will have a row of zeros (the last row) and then A cannot be invertible. Since $a_{nn} \neq 0$, we can use it to replace $a_{n-1,n}, \dots, a_{1n}$ by zeros using elem. row ops. of type 3. Note that $a_{n-1,n-1}$ has no changed. Also, it cannot be zero, since otherwise row $n-1$ would contain all zeros, i.e. A would not be invertible.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ 0 & & & a_{nn} \end{bmatrix} \xrightarrow{a_{nn} \neq 0} \begin{bmatrix} a_{11} & a_{12} & \cdots & & 0 \\ & a_{22} & & & 0 \\ & & \ddots & & \vdots \\ & & & a_{n-1,n-1} & 0 \\ 0 & & & & a_{nn} \end{bmatrix}$$

The argument can be continued to show that all the diagonal elements must be nonzero.

The argument for A lower triangular is: A^T is upper triangular. Hence by the first part

$$\begin{array}{c} A^T \text{ invertible} \iff a_{ii} \neq 0 \quad i=1,\dots,n \\ \Updownarrow \\ A \text{ invertible.} \end{array}$$

(d) We will show that the inverse of an upper triangular and invertible matrix is also upper triangular. Indeed, we know from (c) that $a_{ii} \neq 0$ $i=1,\dots,n$. Now reduce A to rref using elementary row operations. It is important to observe that we use elementary operations of type 2 or type 3. <u>Also</u>, the type 3 operations are such that we always add a multiple of a row to another row of <u>lower index</u>. What this means is that the elementary matrices used are either diagonal (type 2) or upper triangular: E^{μ}_{ij}, $i>j$

$$E_m \cdots E_1 A = I$$

$$\Updownarrow$$

$$E_m \cdots E_1 I = A^{-1}$$

$$E^{\mu}_{ij} = \left[\begin{array}{ccccc} 1 & & & & \\ & \ddots & & & \\ & & 1 & \mu & \\ & & & 1 & \\ & & & & \ddots \; 1 \end{array}\right] \begin{array}{l} \\ \\ \leftarrow j \\ \leftarrow i \\ \\ \end{array}$$

$E_m \cdots E_1$ and I are all upper triangular, so by (b) A^{-1} must be upper triangular.

Finally, if A is lower triangular and invertible

Then A^T is upper triangular & invertible. By the first part $(A^T)^{-1}$ is upper triangular. But

upper $\quad (A^T)^{-1} = (A^{-1})^T \Rightarrow A^{-1}$ is lower tri. ∎

§1.8 Matrix Transformations

R^n is the set of all ordered n-tuples

Ex. $R^3 = \{(x_1, x_2, x_3),\ x_1, x_2, x_3 \in R\}$

we may also think of R^n as the set of all $n \times 1$ matrices.

Defn. If f is a function from R^n (the domain) with values in R^m (the codomain), we say that f is a transformation from R^n to R^m, or f maps R^n into R^m
we write
$$f: R^n \to R^m.$$

If $m=n$, we may f is an operator on R^n

Defn. let $T: R^n \to R^m$ be a transformation. We say that T is a linear transformation or simply linear, if

(i) $T(X+Y) = T(X) + T(Y) \quad \forall X, Y \in R^n$
(ii) $T(\alpha X) = \alpha T(X) \quad \forall \alpha \in R,\ \forall X \in R^n$.

Ex. $T: R^3 \to R^2$,
$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - 2x_2 + x_3 \\ 3x_1 + 5x_2 - x_3 \end{bmatrix}$$

show T is linear.

$$T(X+Y) = T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}\right) \underset{\text{defn. matrix add.}}{=} T\left(\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ x_3+y_3 \end{bmatrix}\right) = \begin{bmatrix} (x_1+y_1) - 2(x_2+y_2) + (x_3+y_3) \\ 3(x_1+y_1) + 5(x_2+y_2) - (x_3+y_3) \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} x_1 - 2x_2 + x_3 \\ 3x_1 + 5x_2 - x_3 \end{bmatrix}}_{T(X)} + \underbrace{\begin{bmatrix} y_1 - 2y_2 + y_3 \\ 3y_1 + 5y_2 - y_3 \end{bmatrix}}_{T(Y)} = T(X) + T(Y)$$

Similarly

$$T(\alpha X) = T\left(\alpha \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = T\left(\begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \alpha x_3 \end{bmatrix}\right) = \begin{bmatrix} \alpha x_1 - 2(\alpha x_2) + \alpha x_3 \\ 3(\alpha x_1) + 5(\alpha x_2) - (\alpha x_3) \end{bmatrix}$$

$$= \alpha \begin{bmatrix} x_1 & -2x_2 & x_3 \\ 3x_1 & 5x_2 & -x_3 \end{bmatrix} = \alpha T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right).$$

Thus T is linear.

Defn. A matrix transformation from R^n into R^m is a map defined as follows: let A be an $m \times n$ matrix. Then

$$T(x) \equiv Ax \qquad \forall x \in R^n$$

Lemma A matrix transformation is linear.

proof

$$T(x+y) = A(x+y) = Ax + Ay \quad \text{distributive property}$$
$$= T(x) + T(y) \ \checkmark$$

$$T(\alpha x) = A(\alpha x) = \alpha Ax \quad (\text{see (m) of Thm 1.4.1}) \ \blacksquare$$

This ~~lemma~~ says that matrix transformations are a way of defining linear transformations. The next result states that this is the only way

Theorem 1.8.3 Every linear transformation from R^n into R^m is a matrix transformation.

proof.

Given a linear transformation $T: R^n \to R^m$, we consistruct a, actually unique, $m \times n$ matrix A such that

$$T(x) = Ax \qquad \forall x \in R^n.$$

Indeed, let

$$A = \left[\, T(e_1) \,\middle|\, T(e_2) \,\middle|\, \cdots \; T(e_n) \,\right]$$

where $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ are the so-called canonical basis vectors of R^n.

Now for any $x \in R^n$.

$$Ax = \left[\, T(e_1) \,\middle|\, \cdots\cdots \,\middle|\, T(e_n) \,\right] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 T(e_1) + \cdots + x_n T(e_n).$$

Since T is linear,

$$x_1 T(e_1) + \cdots + x_n T(e_n) = T(x_1 e_1 + \cdots + x_n e_n) \quad (*)$$
$$= T(x). \quad \blacksquare$$

Ex. Express the linear transformation $T: R^3 \to R^2$ as a matrix transformation, i.e. find A.

$$A = \left[T\begin{pmatrix}1\\0\\0\end{pmatrix} \,\Big|\, T\begin{pmatrix}0\\1\\0\end{pmatrix} \,\Big|\, T\begin{pmatrix}0\\0\\1\end{pmatrix} \right] = \begin{bmatrix} 1 & -2 & 1 \\ 3 & 5 & -1 \end{bmatrix}.$$

(*) Here we used the following consequence of linearity of a transformation: let $\alpha_1, \ldots, \alpha$ be real numbers and $x_1, \ldots, x_k$ vectors in R^n. Then

$$T(\alpha_1 x_1 + \cdots \alpha_k x_k) = \alpha_1 T(x_1) + \cdots + \alpha_k T(x_k).$$

The $k=2$ case is the combination of the two conditions

$T(x_1 + x_2) = T(x_1) + T(x_2)$ and $T(\alpha x) = \alpha T(x)$ into

$$T(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 T(x_1) + \alpha_2 T(x_2).$$

This can be generalized to any finite $k \geq 3$ by induction.