

4. Matrices.

Recall the class of mappings $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$, where

$$(x_1, \dots, x_n) \mapsto (u_1, \dots, u_m), \quad u_i = \sum_{j=1}^n t_{ij} x_j \quad (*)$$

These mappings are linear. Conversely, these are the only linear maps $\mathbb{R}^n \rightarrow \mathbb{R}^m$.

Theorem 4.1: Every linear map $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ can be written in this form (i.e., as in (*)).

Proof: Write $x = \sum_{j=1}^n x_j e_j$ where e_j is the j^{th} unit basis vector.

Since T is linear, $u = Tx = \sum_{j=1}^n x_j Te_j$

Let t_{ij} be the i^{th} component of Te_j , i.e.,

$$e_j \xrightarrow{T} (t_{1j}, t_{2j}, \dots, t_{mj}) = Te_j$$

* Note: If l_1, \dots, l_m is the dual basis to e_1, \dots, e_n , (i.e., $l_i(e_j) = 1$ if $i=j$ and 0 if $i \neq j$), then

$$\boxed{t_{ij} = (l_i, Te_j)}$$

The i^{th} component of u is $u_i = \sum_{j=1}^n x_j t_{ij}$.

Remark: Recall the natural identification $\begin{cases} X \longrightarrow X'' \\ x \longmapsto \hat{x} = (l, x) \end{cases}$ (3)
 sending x to the evaluation map $\hat{x} = (l, x)$ at x ,

Now, for $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, $l = (l_1, \dots, l_n) \in (\mathbb{R}^n)'$, \hat{x} inputs

a vector l , i.e., $\hat{x}(l) = (x_1, \dots, x_n) \begin{pmatrix} l_1 \\ \vdots \\ l_n \end{pmatrix} = l_1 x_1 + \dots + l_n x_n$
 $= (l_1, \dots, l_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (l, x).$

If $S, T: \mathbb{R}^n \rightarrow \mathbb{R}^m$, it is easily checked that $(T+S)_{ij} = T_{ij} + S_{ij}$.

If $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $S: \mathbb{R}^m \rightarrow \mathbb{R}^l$, then $Te_j = c_j$, and so the j^{th} column of ST is $STe_j = Sc_j = \begin{pmatrix} s_{1j} \\ \vdots \\ s_{lj} \end{pmatrix}$, where s_i is the i^{th} row of S .

In summary, we multiply matrices as follows:

Let T be $m \times n$, and S be $l \times m$. The product ST is an $l \times n$ matrix, whose $(kj)^{\text{th}}$ entry is the product of the k^{th} row of S with the j^{th} column of T .

s_k : k^{th} row of S

c_j : j^{th} column of T

$$\begin{bmatrix} S \end{bmatrix}_{l \times m} \begin{bmatrix} T \end{bmatrix}_{m \times n} = \begin{bmatrix} ST \end{bmatrix}_{l \times n}$$

$$(ST)_{kj} = s_k c_j$$

$$= \sum_{i=1}^m s_{ki} t_{ij}$$

$$\begin{matrix} \text{--- } s_1 \text{ ---} \\ \vdots \\ \text{--- } s_m \text{ ---} \end{matrix} \begin{matrix} | c_1 | \\ \dots \\ | c_n | \end{matrix} = \begin{bmatrix} s_1 c_1 + \dots + s_k c_k + \dots + s_m c_n \\ \vdots \\ s_m c_1 + \dots + s_m c_n \end{bmatrix}$$

4

Examples:

$$1. \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix} \neq \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix} = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

$$2. \begin{pmatrix} 1 \\ 2 \end{pmatrix} \begin{pmatrix} 3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 6 & 8 \end{pmatrix} \quad \text{vs.} \quad \begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = (11)$$

$$3. \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} = (13 \ 16) \quad \text{vs.} \quad \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 11 \\ 17 \end{pmatrix}$$

$$4. \begin{pmatrix} 1 & 2 \end{pmatrix} \left(\begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right) = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 11 \\ 17 \end{pmatrix} = (45)$$

$$\text{vs.} \quad \left(\begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} \right) \begin{pmatrix} 1 \\ 2 \end{pmatrix} = (13 \ 16) \begin{pmatrix} 1 \\ 2 \end{pmatrix} = (45)$$

Remark: Let A be an $m \times n$ matrix, and $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_m \end{pmatrix}$ a diagonal matrix, that is, $D_{ij} = \begin{cases} d_j & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$.

The i^{th} row of DA is the i^{th} row of A , times d_i .

The j^{th} column of AD is the j^{th} column of A , times d_j .

An $n \times n$ matrix represents a mapping $\mathbb{R}^n \rightarrow \mathbb{R}^n$. If the mapping is invertible, then the matrix is called invertible.

Remark: Since the composition of linear maps is associative, matrix multiplication is associative as well.

* We shall identify \mathbb{R}^n with column vectors, and its dual $(\mathbb{R}^n)'$ with row vectors (with n components).

A co-vector $l \in (\mathbb{R}^n)'$ applied to a vector $x \in \mathbb{R}^n$ is the matrix product $(l, x) = (l_1, \dots, l_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = l_1 x_1 + \dots + l_n x_n$.

Let x , T , and l be the following linear maps:

$$x: \mathbb{R} \longrightarrow \mathbb{R}, \quad T: \mathbb{R}^n \longrightarrow \mathbb{R}^m, \quad l: \mathbb{R}^m \longrightarrow \mathbb{R}.$$

Note that $l \in (\mathbb{R}^m)'$, $lT \in (\mathbb{R}^n)'$.

By associativity, $(lT)x = l(Tx)$.

We can write this as $(lT, x) = (l, Tx)$.

Recall that the transpose T' of T satisfies $(T'l, x) = (l, Tx)$.

Let's compare the matrix forms of T and T' :

Let e_1, \dots, e_n be a basis for X , and let $l_1, \dots, l_n \in X'$ be the dual basis ($l_i(e_j) = 1$ if $i=j$, 0 if $i \neq j$).

Recall that $Te_j = \sum_{k=1}^m t_{kj} e_k \Rightarrow (l_i, Te_j) = \sum_{k=1}^m t_{kj} (l_i, e_k) = t_{ij}$.

* Thus, to get t_{ij} :

- Apply T to the j^{th} basis vector of X
- Then apply the i^{th} functional in X' to this.

6

Now consider T' :

Using our recipe for t_{ij} just derived, the ij -entry t'_{ij} of T' is obtained by:

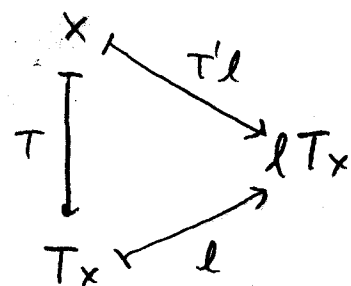
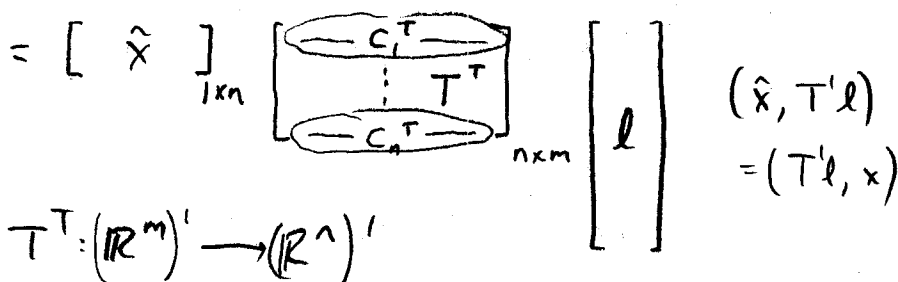
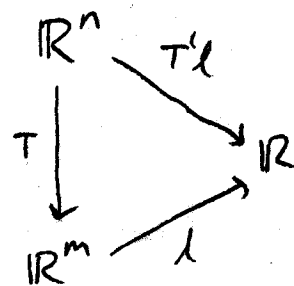
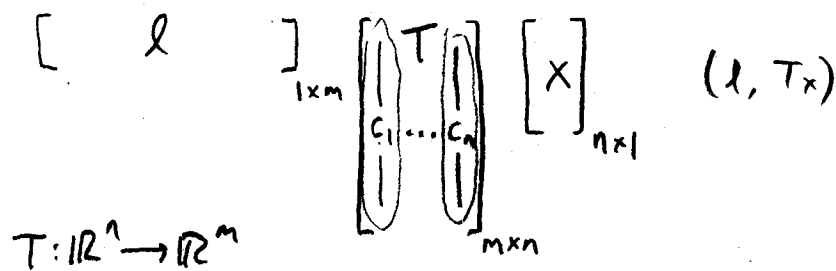
- Applying T' to the j^{th} basis vector of X' .

- Then applying the i^{th} functional in X'' to this

i.e., $(\hat{x}_i, T'l_j)$, or just $(T'l_j, x_i)$ under the natural identification $X \leftrightarrow X''$.

Note that $t'_{ij} = (T'l_j, x_i) = (l_j, Tx_i) = t_{ji}$.

* In summary, to represent the transpose T' as a matrix, we change its rows into columns, and vice-versa. Denote this matrix by T^T . We have: $(T^T)_{ij} = T_{ji}$



Now, let's express the range of T in matrix language.

Recall that $Te_j = c_j$.

Plugging this into $Tx = \sum_{j=1}^n x_j Te_j$ gives

$$u = Tx = x_1 c_1 + \dots + x_n c_n.$$

In other words:

Theorem 4.2: The range of T consists of all linear combinations of the column vectors of the matrix.

The dimension of this space is called the column rank of T , in remedial books and courses. The row rank is defined similarly.

Remarks:

- (i) The row rank of T is $\dim R_T$ (or $\dim R_{T^t}$).
- (ii) By Theorem 3.6 ($\dim R_T = \dim R_{T^t}$), the row rank and column rank of a matrix are equal.

Let $T: X \rightarrow U$ be a linear map, and let x_1, \dots, x_n be a basis for X , and u_1, \dots, u_m a basis for U .

Since $\dim X = n$, $\dim U = m$, we have $X \cong \mathbb{R}^n$ and $U \cong \mathbb{R}^m$.

[8]

That is, we have isomorphisms:

$$B: X \longrightarrow \mathbb{R}^n \quad C: U \longrightarrow \mathbb{R}^m$$
$$x_i \longmapsto e_i \quad u_i \longmapsto e_i$$

Putting this together, we can choose isomorphisms such as these to get a linear map $CTB^{-1}: \mathbb{R}^n \longrightarrow \mathbb{R}^m$

$$\mathbb{R}^n \xrightarrow{B^{-1}} X \xrightarrow{T} U \xrightarrow{C} \mathbb{R}^m$$

If $T: X \rightarrow X$, mapping X to itself, then we can take

$$B=C, \text{ and so we get a matrix } M=BTB^{-1}.$$

Suppose we change the isomorphism B . In particular, let

$C: X \rightarrow \mathbb{R}^n$ be another isomorphism, and let N be the matrix w.r.t. this basis, i.e., $N=CTC^{-1}$.

$$\text{We have: } N=CTC^{-1}=CB^{-1}BTB^{-1}BC^{-1}=SMS^{-1},$$

where $S=CB^{-1}$, which is invertible.

$$\mathbb{R}^n \xrightarrow{C^{-1}} X \xrightarrow{B} \mathbb{R}^n \xrightarrow{B^{-1}} X \xrightarrow{T} X \xrightarrow{B} \mathbb{R}^n \xrightarrow{B^{-1}} X \xrightarrow{C} \mathbb{R}^n$$

Two square matrices N and M related by conjugation

(e.g., $N=SMS^{-1}$) are said to be similar.

* Similar matrices describe the same mapping of a space into itself, but using different bases. Thus, (as we'll see later), similar matrices share the same intrinsic properties.

Remark: We can write any $n \times n$ matrix A in block form:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{where} \quad A = \left. \begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right\} \begin{array}{l} k \text{ rows} \\ n-k \text{ rows} \end{array}$$

$\underbrace{\hspace{10em}}_{k \text{ columns}} \quad \underbrace{\hspace{10em}}_{n-k \text{ columns}}$

Addition and multiplication of block matrices "works out" just as if the blocks were entries.

Def: The square matrix $I = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$ is called the identity, or unit matrix.

Def: A square matrix $T = (t_{ij})$ for which $t_{ij} = 0$ for $i > j$ (resp. $i < j$) is called upper triangular (resp. lower triangular).

Def: A square matrix $T = (t_{ij})$ for which $t_{ij} = 0$ for $|i-j| > 1$ is called tridiagonal.

10

Systems of equations:

Matrices can be used to effectively express and solve systems of equations.

A system of equations $\sum_{i=1}^n t_{ij} x_i = u_j$ for $j=1, \dots, n$

may have a unique solution, many solutions, or no solutions.

Consider the following example, where x_1, \dots, x_4 are the unknowns:

$$x_1 + x_2 + 2x_3 + 3x_4 = u_1$$

$$x_1 + 2x_2 + 3x_3 + x_4 = u_2$$

$$2x_1 + x_2 + 2x_3 + 3x_4 = u_3$$

$$3x_1 + 4x_2 + 6x_3 + 2x_4 = u_4$$

We can write this system as $Mx = u$:

$$\begin{pmatrix} 1 & 1 & 2 & 3 \\ 1 & 2 & 3 & 1 \\ 2 & 1 & 2 & 3 \\ 3 & 4 & 6 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}$$

This system can be solved by Gaussian elimination, i.e., eliminating the variables one-by-one.

Example (cont.): Subtract multiples of E_0 from the last 3 equations to eliminate x_1 :

$$\begin{aligned}x_2 + x_3 - 2x_4 &= u_2 - u_1 \\-x_2 - 2x_3 - 3x_4 &= u_3 - 2u_1 \\x_2 - 7x_4 &= u_4 - 3u_1\end{aligned}$$

Use the 1st equation to eliminate x_2 from the last two:

$$\begin{aligned}-x_3 - 5x_4 &= u_3 + u_2 - 3u_1 \\-x_3 - 5x_4 &= u_4 - 2u_2 - 2u_1\end{aligned}$$

Subtract these two equations to eliminate x_3 (and by chance, x_4 as well):

$$0 = u_4 - u_3 - 2u_2 + u_1$$

This is a necessary and sufficient condition for our original system to have a solution. It can be

written as $l \cdot u = 0$, where $l = (1, -2, -1, 1)$,

that is, $(1, -2, -1, 1) \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = u_1 - 2u_2 - u_3 + u_4 = 0$.

Since $Mx = u$, and $lu = 0$, we must have

$lMx = lu = 0$ for all $x \in \mathbb{R}^4$, thus $lM = 0$.

* In general, if $Mx = u$ is a system of equations, then a solution is described by a linear function l (equivalently, a column vector) if $lM = 0$.