Aim lecture

In first year you learnt that you can mutliply not only a (real) matrix with a (real) vector, but more generally, matrices together (of compatible sizes). Furthermore, you learnt the following matrix representation theorem which essentially states that linear maps from $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ are no more complicated than matrices.

First Year Matrix Representation Theorem

Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear map. Then there is a unique matrix $A \in M_{mn}(\mathbb{R})$ such that T is left multliplication by A, i.e. $T\mathbf{v} = A\mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^n$. In fact

$$A = (T\mathbf{e}_1 \, T\mathbf{e}_2 \ldots T\mathbf{e}_n)$$

i.e. the i-th column of A is the value of T at the i-th standard basis vector $\mathbf{e}_i \in \mathbb{R}^n$.

Aim lecture: Generalise matrix multiplication to matrices of linear maps and also the matrix representation thm above.

Matrix multiplication

Consider direct sums of $\mathbb{F}$-spaces $U = \oplus_{k=1}^{l} U_k$, $V = \oplus_{j=1}^{m} V_j$, $W = \oplus_{i=1}^{n} W_i$.

Prop-Defn

Consider matrices $T = (T_{ij})_{ij} \in (L(V_j, W_i))_{ij}$, $S = (S_{jk})_{jk} \in (L(U_k, V_j))_{jk}$. We define the *product matrix* $TS \in (L(U_k, W_i))_{ik}$ to be the one with (i,k)-th entry

$$(TS)_{ik} = \sum_{j=1}^{m} T_{ij} \circ S_{jk}.$$

The linear map assoc to TS is the composite map $T \circ S : U \longrightarrow W$.

Proof. Consider $(\mathbf{u}_k)_k \in U$. Then we calculate

$$(T \circ S)(\mathbf{u}_k)_k = T[S(\mathbf{u}_k)_k] = T[(\sum_k S_{jk}\mathbf{u}_k)_j]$$

$$= (\sum_j T_{ij}[\sum_k S_{jk}\mathbf{u}_k])_i = (\sum_j \sum_k T_{ij} \circ S_{jk}\mathbf{u}_k)_i = (TS)(\mathbf{u}_k)_k$$

Examples of matrix multiplication

E.g Compute the product

$$\begin{pmatrix} \frac{d}{dx} & 0 \\ 2\frac{d}{dx} & 3\frac{d}{dx} \end{pmatrix} \begin{pmatrix} \frac{d}{dx} & -\mathrm{id} & 0 \\ 2\,\mathrm{id} & -\frac{d}{dx} & \frac{d}{dx} + \mathrm{id} \end{pmatrix} =$$

Fact

Products of matrices of matrices are the same as the product of the big matrices of scalars. This follows since they both correspond to the same composite of linear maps.

More examples

E.g. Let I be the 2×2 identity matrix & A be the 4×4-matrix $\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$. Find all 4×4-matrices X such that $AX = XA$.

Matrix representation theorem

Consider direct sums of $\mathbb{F}$-spaces $V = \oplus_{j=1}^{n} V_j$, $W = \oplus_{i=1}^{m} W_i$. We have seen an example where we were given a linear map from $V \longrightarrow W$ and we wrote it as a matrix in $(L(V_j, W_i))_{ij}$. This is always possible by the following

Theorem

The map $\Phi : (L(V_j, W_i))_{ij} \longrightarrow L(V, W)$ which assigns to a matrix $T = (T_{ij})_{ij}$ the associated linear map $\mathbf{v} \mapsto T\mathbf{v}$ is an isomorphism of vector spaces. In particular, every linear map from $V \longrightarrow W$ can be written uniquely in the form of a matrix in $(L(V_j, W_i))_{ij}$.

Proof. Occupies next two slides & will only be sketched.

Proof of bijectivity of Φ

We first prove the map Φ is surjective, i.e. any linear map $T : V \longrightarrow W$ comes from a matrix.

We will only look at the case $m = 1$, the general case (ex) follows by noting that an $m \times n$-matrix is just a length n row matrix of length m column matrices. To keep notn simple, we will also assume $n = 2$ though the general case is just as easy & just requires lengthier notn. Hence $T : V_1 \oplus V_2 \longrightarrow W$.

Recall natural isomorphisms
$\Psi_1 : V_1 \longrightarrow V(1) = V_1 \oplus \mathbf{0} : \mathbf{v}_1 \mapsto \binom{\mathbf{v}_1}{\mathbf{0}}, \Psi_2 : V_2 \longrightarrow V(2) = \mathbf{0} \oplus V_2 : \mathbf{v}_2 \mapsto \binom{\mathbf{0}}{\mathbf{v}_2}$.
Suffice prove $T = \Phi(T_1 \quad T_2)$ where

$$T_1 = T \circ \Psi_1 : V_1 \longrightarrow W, T_2 = T \circ \Psi_2 : V_2 \longrightarrow W$$

$$\begin{aligned} T\begin{pmatrix}\mathbf{v}_1\\ \mathbf{v}_2\end{pmatrix} &= T\begin{pmatrix}\mathbf{v}_1\\ \mathbf{0}\end{pmatrix} + T\begin{pmatrix}\mathbf{0}\\ \mathbf{v}_2\end{pmatrix} = T(\Psi_1\mathbf{v}_1) + T(\Psi_2\mathbf{v}_2) \\ &= T_1\mathbf{v}_1 + T_2\mathbf{v}_2 = \begin{pmatrix}T_1 & T_2\end{pmatrix}\begin{pmatrix}\mathbf{v}_1\\ \mathbf{v}_2\end{pmatrix}. \end{aligned}$$

Hence $\Phi(T_1 \quad T_2) = T$ & surjectivity is proved.

Proof injectivity of Φ

We now prove that Φ is injective. Suppose then that $\Phi(T_1 \quad T_2) = \Phi(T_1' \quad T_2')$ for some linear maps $T_1, T_1' : V_1 \longrightarrow W$ & $T_2, T_2' : V_2 \longrightarrow W$.

Suffice check $T_1 = T_1', T_2 = T_2'$ by showing they have the same outputs for every possible input. Let $\mathbf{v}_1 \in V_1$. Then

$$T_1\mathbf{v}_1 = (T_1 \quad T_2)\begin{pmatrix}\mathbf{v}_1 \\ \mathbf{0}\end{pmatrix} = (T_1' \quad T_2')\begin{pmatrix}\mathbf{v}_1 \\ \mathbf{0}\end{pmatrix} = T_1'\mathbf{v}_1.$$

Hence $T_1\mathbf{v}_1 = T_1'\mathbf{v}_1$ for all $\mathbf{v}_1 \in V_1$ which shows $T_1 = T_1'$. Sim (ex) one sees $T_2 = T_2'$.

This completes the proof of injectivity & hence bijectivity.

Proof linearity of Φ

Recall first that the vector space structure of $(L(V_j, W_i))_{ij}$ is the direct sum of the $L(V_j, W_i)$. We need to check that for any $(T_{ij}), (T'_{ij}) \in (L(V_j, W_i))_{ij}, \beta \in \mathbb{F}$ and $\mathbf{v} \in V$ we have

$$\Phi((T_{ij}) + (T'_{ij}))\mathbf{v} = [\Phi(T_{ij}) + \Phi(T'_{ij})]\mathbf{v}, \quad \Phi(\beta(T_{ij}))\mathbf{v} = [\beta\Phi(T_{ij})]\mathbf{v}.$$

This follows from the following distributive & associative laws of the matrix-vector product

Prop

With the above notn

1. $[(T_{ij}) + (T'_{ij})]\mathbf{v} = (T_{ij})\mathbf{v} + (T'_{ij})\mathbf{v}$
2. $[\beta(T_{ij})]\mathbf{v} = \beta[(T_{ij})\mathbf{v}]$

Proof. This is an easy check. We do 2) as an example. Write $\mathbf{v} = (\mathbf{v}_j)_j$. Then

$$[\beta(T_{ij})]\mathbf{v} = (\beta T_{ij})(\mathbf{v}_j)_j = (\sum_j \beta T_{ij}\mathbf{v}_j)_i = \beta(\sum_j T_{ij}\mathbf{v}_j)_i = \beta[(T_{ij})\mathbf{v}].$$

Special case of row matrices

Cor

Let $T : \mathbb{F}^n \longrightarrow V$ be a linear map of $\mathbb{F}$-spaces. Then $T\mathbf{v} = A\mathbf{v}$ where $A \in V^n$ is the row matrix

$$A = (T\mathbf{e_1} \dots T\mathbf{e_n})$$

i.e. the i-th column of A is $T\mathbf{e_i}$ where $\mathbf{e_i}$ is i-th standard basis vector.

Proof. Thm $\Longrightarrow$ T is represented by left multn by some $A \in V^n$ and we deduce the i-th column of A must be $A\mathbf{e_i} = T\mathbf{e_i}$

E.g. Represent the linear map $T : \mathbb{R}^3 \longrightarrow M_{22}(\mathbb{R})$ below by a row matrix (of matrices!).

$$T\begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix} = \begin{pmatrix}\beta_1+\beta_2 & \beta_3\\ \beta_3 & \beta_2-\beta_3\end{pmatrix}.$$