Math 672 Lecture 9

All of the following definitions should be familiar to you:

Definition:

- An $m \times n$ matrix A with entries in F is a collection of scalars

$$A_{ij}, \quad 1 \le i \le m, \; 1 \le j \le n,$$

usually arranged in a rectangular array as follows:

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix}.$$

- The collection of all $m \times n$ matrices with entries in F is denoted by $F^{m \times n}$ (or $F^{m,n}$ by our author).

- Addition and scalar multiplication of matrices in $F^{m \times n}$ are defined entrywise:

$$A, B \in F^{m \times n} \Rightarrow (A+B)_{ij} = A_{ij} + B_{ij}$$

$$\left(\text{e.g. } \begin{bmatrix} 1 & 2 & 1 \\ 2 & -1 & 3 \end{bmatrix} + \begin{bmatrix} 3 & -1 & -2 \\ -2 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 1 & -1 \\ 0 & 3 & 4 \end{bmatrix} \right)$$

$$A \in F^{m \times n}, \alpha \in F \Rightarrow (\alpha A)_{ij} = \alpha A_{ij}$$

$\left(\text{e.g. } 3\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 3 \end{bmatrix} \right)$

- Matrix-vector multiplication is defined as follows:

If $A \in F^{m\times n}$ and $x \in F^n$, then $Ax \in F^m$ is defined by

$$(Ax)_i = \sum_{j=1}^{n} A_{ij}x_j = A_{i1}x_1 + A_{i2}x_2 + \cdots + A_{in}x_n \quad \forall i=1,2,\ldots,m$$

- Matrix-matrix multiplication is defined as follows: If $A \in F^{m\times n}$ and $B \in F^{n\times p}$, then $AB \in F^{m\times p}$ and

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj} \quad \forall i=1,2,\ldots,m \; \forall j=1,2,\ldots,p.$$

Theorem

1. $F^{m\times n}$, under the addition and scalar multiplication defined above, is a vector space over F, and $\dim(F^{m\times n}) = mn$.
2. If we regard the columns of $A \in F^{m\times n}$ as defining vectors $A_1, A_2, \ldots, A_n$ in F^m, then

$$Ax = \sum_{j=1}^{n} x_j A_j \quad \forall x \in F^n.$$

(Thus Ax is a linear combination of the columns of A, where the weights in the linear combination are the components of x.)

Notation: $A = [A_1 | A_2 | \cdots | A_n]$ means that $A_1, A_2, \ldots, A_n$ are the columns of A.

3. Matrix-vector multiplication defines a linear map. That is, if $A \in F^{m \times n}$ and $T : F^n \to F^m$ is defined by $T(x) = Ax$ for all $x \in F^n$, then T is a linear map.

4. If $A \in F^{m \times n}$, $B \in F^{n \times p}$, and $B = [B_1 | B_2 | \cdots | B_p]$, then

$$AB = [AB_1 | AB_2 | \cdots | AB_p].$$

5. Matrix multiplication is associative:

$$A \in F^{m \times n}, B \in F^{n \times p}, C \in F^{p \times q} \Rightarrow (AB)C = A(BC).$$

Similarly,

$$A \in F^{m \times n}, B \in F^{n \times p}, x \in F^p \Rightarrow (AB)x = A(Bx).$$

6. Matrix multiplication distributes over addition:

$$A \in F^{m \times n}, B, C \in F^{n \times p} \Rightarrow A(B+C) = AB + AC,$$

$$B, C \in F^{m \times n}, A \in F^{n \times p} \Rightarrow (B+C)A = BA + CA.$$

Proof (partial):

1. This is straightforward but tedious to prove.

The additive identity in $F^{m \times n}$ is the matrix whose every entry is zero. The standard basis for $F^{m\times n}$ consists of the mn matrices E_{ij}, $1\le i\le m$, $1\le j\le n$, where every entry of E_{ij} is 0 except the i,j entry, which is 1.

2. We have, for each $i=1,2,\dots,m$

$$\left(\sum_{j=1}^{n} x_j A_j\right)_i = \sum_{j=1}^{n} x_j A_{ij} \quad (\text{since } A_{ij} \text{ is the } i\text{th component of } A_j)$$

$$= \sum_{j=1}^{n} A_{ij} x_j = (Ax)_i .$$

Thus

$$(Ax)_i = \left(\sum_{j=1} x_j A_j\right)_i \;\; \forall i=1,2,\dots,m$$

$$\Rightarrow Ax = \sum_{j=1}^{n} x_j A_j .$$

3. Using #2, we have

$$A(\alpha x+\beta y) = \sum_{j=1}^{n} (\alpha x_j + \beta y_j) A_j \quad (\text{since } (\alpha x+\beta y)_j = \alpha x_j + \beta y_j)$$

$$= \alpha \sum_{j=1}^{n} x_j A_j + \beta \sum_{j=1}^{n} y_j A_j \quad (\text{using various vector space properties in } F^m)$$

$$= \alpha Ax + \beta Ay .$$

4. We have

$$\left([AB_1 | AB_2 | \cdots | AB_p]\right)_{ij} = (AB_j)_i \quad \text{(the ith entry of the jth column)}$$

$$= \sum_{k=1}^{n} A_{ik} B_{kj}$$

$$= (AB)_{ij},$$

Thus

$$AB = [AB_1 | AB_2 | \cdots | AB_p].$$

5. Let $A \in F^{m\times n}$, $B \in F^{n\times p}$, $x \in F^p$. Then

$$\left((AB)x\right)_i = \sum_{j=1}^{p} (AB)_{ij} x_j$$

$$= \sum_{j=1}^{p} \left(\sum_{k=1}^{n} A_{ik} B_{kj}\right) x_j$$

$$= \sum_{j=1}^{p} \sum_{k=1}^{n} A_{ik} B_{kj} x_j$$

$$= \sum_{k=1}^{n} \sum_{j=1}^{p} A_{ik} B_{kj} x_j$$

$$= \sum_{k=1}^{n} A_{ik} \left(\sum_{j=1}^{p} B_{kj} x_j\right)$$

$$= \sum_{k=1}^{n} A_{ik} (Bx)_k$$

$$= (A(Bx))_i .$$

This holds for all $i = 1, 2, \ldots, m$, so

$$(AB)x = A(Bx).$$

Given this result, let $A \in F^{m \times n}$, $B \in F^{n \times p}$, $C \in F^{p \times q}$. Then

$$\begin{aligned}(AB)C &= [(AB)C_1 | \cdots | (AB)C_q] \\ &= [A(BC_1) | \cdots | A(BC_q)] \\ &= A[BC_1 | \cdots | BC_q] \\ &= A(BC). \text{ //}\end{aligned}$$

Note: Suppose $A \in F^{m \times n}$, $B \in F^{n \times p}$, $T : F^p \to F^n$ is defined by $T(x) = Bx$ for all $x \in F^p$, and $S : F^n \to F^m$ is defined by $Sx = Ax$ for all $x \in F^n$. Then

$$(ST)(x) = S(T(x)) = S(Bx) = A(Bx) = (AB)x.$$

Thus if S is defined by multiplication by A and T is defined by

multiplication by B, then ST is defined by multiplication by AB.

Matrix multiplication is defined as it is so that this will be true.

We know that matrix-vector multiplication defines a linear map: If $A \in F^{m \times n}$, then $T: F^n \to F^m$ defined by $T(x) = Ax$ for all $x \in F^n$ is linear.

By an earlier exercise, we know that if $T: F^n \to F^m$ is linear, then there exists $A \in F^{m \times n}$ such that $T(x) = Ax$ for all $x \in F^n$.

In fact,

$$A = [T(e_1) \mid T(e_2) \mid \cdots \mid T(e_n)],$$

where $\{e_1, e_2, \dots, e_n\}$ is the standard basis for F^n:

$$e_i = (0, \dots, 0, \underset{\substack{\uparrow \\ \text{ith component.}}}{1}, 0, \dots, 0)$$

Something much stronger is true: Given any finite-dimensional vector spaces V and W and any $T \in \mathcal{L}(V, W)$, we can represent T by a matrix once we choose bases $\{v_1, v_2, \dots, v_n\}$ for V and $\{w_1, w_2, \dots, w_m\}$ for W. This is the subject of the next lecture.

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, "(*)"'] & W \\
F^n \arrow[r, "A"] & F^m \arrow[u, "(**)"']
\end{tikzcd}$$

(*) Given $v \in V$, write v as a linear combination of the basis vectors: $v = x_1 v_1 + \cdots + x_n v_n$. This yields $x = (x_1, x_2, \ldots, x_n) \in \mathbb{F}^n$.

(**) Given $y = (y_1, y_2, \ldots, y_m) \in \mathbb{F}^m$, define $w = y_1 w_1 + y_2 w_2 + \cdots + y_m w_m \in W$.