Math 507 | Review of theory of functions between sets: 31

Let S and T be any sets. Say $f: S \to T$ is a function from **domain** S to **codomain** T when

$\forall s \in S$, $\exists t \in T$ such that $f(s) = t$.

Def. $\text{Range}(f) = \text{Image}(f) = \{f(s) \in T \mid s \in S\}$

$= \{t \in T \mid \exists s \in S \text{ s.t. } f(s) = t\} \subseteq T$.

Def. Say $f: S \to T$ is **surjective** (onto) if $\text{Range}(f) = T$.

Def. Say $f: S \to T$ is **injective** (one-to-one) when

$\forall s_1, s_2 \in S$, $f(s_1) = f(s_2)$ implies $s_1 = s_2$.

This is logically equivalent to its **contrapositive**:

$\forall s_1, s_2 \in S$, $s_1 \neq s_2$ implies $f(s_1) \neq f(s_2)$.

Def. Say $f: S \to T$ is <u>bijective</u> (one-to-one correspondence) when f is both injective and surjective.

Def. Say $f: S \to T$ is <u>invertible</u> when $\exists g: T \to S$ such that ① $\forall s \in S$, $g(f(s)) = s$, and ② $\forall t \in T$, $f(g(t)) = t$.

Th. If $f: S \to T$ is invertible then there is only one $g: T \to S$ satisfying ① and ② above, so we can denote that <u>unique</u> inverse of f by f^{-1}.

Pf: Given $f: S \to T$ suppose $\exists g_1, g_2: T \to S$ such that ① $\forall s \in S$, $g_i(f(s)) = s$ for $i = 1, 2$, and ② $\forall t \in T$, $f(g_i(t)) = t$ for $i = 1, 2$.

By ① with $i=1$ and $s = g_2(t)$ we have, $\forall t \in T$, $g_1(f(g_2(t))) = g_2(t)$, and by ② with $i=2$, $g_1(f(g_2(t))) = g_1(t)$, so $g_1(t) = g_2(t)$ so $g_1 = g_2$ $\square$

__Def.__ Let $f : S \to T$ and $g : R \to S$ for sets R, S and T. Define the __composition__ function $(f \circ g) : R \to T$ by $(f \circ g)(r) = f(g(r))$, $\forall r \in R$.

This is summarized by the __diagram__:

$$
\begin{tikzcd}
R \arrow[r, "g"] \arrow[dr, "f \circ g"'] & S \arrow[d, "f"] \\
 & T
\end{tikzcd}
$$

Def: For any set S, the identity function (map) is $I_S : S \to S$ where $\forall s \in S$, $I_S(s) = s$.

Note: $f: S \to T$ is invertible when $\exists g : T \to S$ s.t. ① $g \circ f = I_S$ and ② $f \circ g = I_T$.

Also, $I_T \circ f = f$ for <u>any</u> $f : S \to T$ and

$g \circ I_T = g$ for <u>any</u> $g : T \to S$.

Th: $f: S \to T$ is invertible iff f is bijective.

Th: Composition of functions is associative.

Pf. let

$$\begin{tikzcd} R \arrow[r, "g"] \arrow[dr, "(f \circ g)"'] \arrow[drr, bend left=60, "(h \circ f) \circ g"] \arrow[drr, bend right=60, "h \circ (f \circ g)"'] & S \arrow[d, "f"] \arrow[dr, "(h \circ f)"] & \\ & T \arrow[r, "h"] & U \end{tikzcd}$$

Then $\forall r \in R$,

$(h \circ (f \circ g))(r) = h(f(g(r)))$

$= ((h \circ f) \circ g)(r)$, so $h \circ (f \circ g) = (h \circ f) \circ g$. □

Back to linear algebra.

<u>Th</u>: Let $L: V \to W$ be linear. Then L is injective iff $\mathrm{Ker}(L) = \{\theta_V\}$.

<u>Pf.</u> ($\Rightarrow$) Suppose L is inj. and let $v \in K = \mathrm{Ker}(L)$, so $L(v) = \theta_W = L(\theta_V)$ so $v = \theta_V$ by inj. of L. So the only element in K is θ_V.

($\Leftarrow$) Suppose $\mathrm{Ker}(L) = \{\theta_V\}$ and that for $v_1, v_2 \in V$, we have $L(v_1) = L(v_2)$. Then $L(v_1) - L(v_2) = \theta_W$ so $L(v_1 - v_2) = \theta_W$, meaning $v_1 - v_2 \in \mathrm{Ker}(L) = \{\theta_V\}$ so $v_1 - v_2 = \theta_V$ so $v_1 = v_2$. This shows L is inj. $\square$

<u>Cor.</u> Let $A \in \mathbb{F}^m_n$ and $L_A: \mathbb{F}^n \to \mathbb{F}^m$ be the linear map $L_A(X) = AX$. Then L_A is injective iff the homogeneous lin. sys. $AX = 0^m_1$ has only the trivial sol'n.

<u>Pf.</u> $\mathrm{Ker}(L_A) = \{X \in F^n \mid AX = 0_1^m\}$ is only $\{0_1^n\}$ iff L_A is injective by the last Theorem. $\square$

<u>Th:</u> $L_A : F^n \to F^m$ is surjective iff $\forall B \in F^m$, $AX = B$ is consistent.

How to use composition of linear maps to define matrix multiplication and give a conceptual proof that it is associative.

Let $A = [a_{ij}] \in F^m_n$, $B = [b_{jk}] \in F^n_p$ so get

$$\begin{tikzcd} F^p \arrow[r, "L_B"'] \arrow[dr, "L_C"'] & F^n \arrow[d, "L_A"] \\ & F^m \end{tikzcd}$$

Can we find $C = [c_{ik}] \in F^m_p$ such that

$$L_C = L_A \circ L_B \ ?$$

This would mean that $\forall\, Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_p \end{bmatrix} \in F^p$,

$CY = A(BY)$ since

$L_C(Y) = L_A(L_B(Y))$. Let

$$X = BY = \begin{bmatrix} \sum_{k=1}^{p} b_{1k} Y_k \\ \vdots \\ \sum_{k=1}^{p} b_{nk} Y_k \end{bmatrix} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} \in F^n$$ so that

$$AX = A(BY) = \begin{bmatrix} \sum_{j=1}^{n} a_{1j} X_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} X_j \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} \sum_{k=1}^{p} a_{1j} b_{jk} Y_k \\ \vdots \\ \sum_{j=1}^{n} \sum_{k=1}^{p} a_{mj} b_{jk} Y_k \end{bmatrix} =$$

$$\begin{bmatrix} \sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{1j} b_{jk}\right) y_k \\ \vdots \\ \sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{mj} b_{jk}\right) y_k \end{bmatrix} = CY \text{ iff } \boxed{c_{ik}} = \sum_{j=1}^{n} a_{ij} b_{jk}$$

for $1 \le i \le m$, $1 \le k \le p$.

<u>Th</u>: For any $A = [a_{ij}] \in F_n^m$, $B = [b_{jk}] \in F_p^n$, the matrix $C = [c_{ik}] \in F_p^m$ such that $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$ satisfies $L_C = L_A \circ L_B$.

<u>Def</u>: Let the matrix C above be called the matrix product of A and B, denoted $C = AB$.

When we defined $L_A : F^n \to F^m$ from a choice of $A \in F^m_n$, we should have considered this question: If $A, B \in F^m_n$ and $L_A = L_B$, does $A = B$ have to be true?

$L_A = L_B$ means $\forall X \in F^n$, $AX = BX$, so in particular, looking back at the definition on p.22, let $X = e_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{row } j$ be the column vector with 1 in row j, 0 in all other rows.

Then $AX = \sum_{j=1}^{n} x_j \, Col_j(A)$ says $Ae_j = Col_j(A)$ and $BX = Col_j(B)$. So $Col_j(A) = Col_j(B)$ for all $1 \le j \le n$, which means $A = B$.

We have proven:

<u>Th</u>: For $A, B \in F^m_n$, if $L_A = L_B$ then $A = B$.

We can now prove that matrix multiplication is associative, and see that it comes from the associativity of composition of functions.

<u>Th</u>: Suppose $A \in F^m_n$, $B \in F^n_p$ and $C \in F^p_q$, so $AB \in F^m_p$, $BC \in F^n_q$, $(AB)C \in F^m_q$ and $A(BC) \in F^m_q$. Then $(AB)C = A(BC)$.

<u>Pf</u>. By definition, $L_{AB} = L_A \circ L_B$, $L_{BC} = L_B \circ L_C$, $L_{(AB)C} = L_{AB} \circ L_C$ and $L_{A(BC)} = L_A \circ L_{BC}$ so

$L_{(AB)C} = (L_A \circ L_B) \circ L_C = L_A \circ (L_B \circ L_C) = L_{A(BC)}$.

The middle "=" is from assoc. of composition and the last Theorem tells us that $L_{(AB)C} = L_{A(BC)}$ implies $(AB)C = A(BC)$. $\square$

<u>Note</u>: For $A = [a_{ij}] \in F_n^m$ and $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in F_1^n = F^n$ our definition of $AX \in F^m$ on page 22 is a special case of the matrix multiplication on page 38. In fact, the direct connection is

$$\underset{(m\times n)(n\times p)}{AB} = \left[A\,Col_1(B) \middle| A\,Col_2(B) \middle| \cdots \middle| A\,Col_p(B) \right] = C$$

that is, $Col_k(AB) = A\,Col_k(B)$ for $1 \le k \le p$.

Example:

$$\underset{\substack{A \\ (2\times3)}}{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}} \underset{B\,(3\times2)}{\begin{bmatrix} -1 & 1 \\ 5 & 9 \\ 4 & 6 \end{bmatrix}} = \underset{AB \quad (2\times2)}{\begin{bmatrix} (1)(-1)+(2)(5)+(3)(4) & (1)(1)+(2)(9)+(3)(6) \\ (4)(-1)+(5)(5)+(6)(4) & (4)(1)+(5)(9)+(6)(6) \end{bmatrix}}$$

$$= \begin{bmatrix} (-1+10+12) & (1+18+18) \\ (-4+25+24) & (4+45+36) \end{bmatrix} = \begin{bmatrix} 21 & 37 \\ 45 & 85 \end{bmatrix}$$

while

$$\underset{A}{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}} \underset{Col_1(B)}{\begin{bmatrix} -1 \\ 5 \\ 4 \end{bmatrix}} = \begin{bmatrix} (1)(-1)+(2)(5)+(3)(4) \\ (4)(-1)+(5)(5)+(6)(4) \end{bmatrix} = Col_1(AB) = \begin{bmatrix} 21 \\ 45 \end{bmatrix}$$

and similarly, $A\, Col_2(B) = Col_2(AB) = \begin{bmatrix} 37 \\ 85 \end{bmatrix}$.

Matrix Algebra.

From the definitions it is easy to prove the basic laws of matrix algebra relating addition, scalar multiplication and matrix mult. Besides associativity of matrix mult. we also have: For appropriate size matrices:

Distributive laws: $A(B+C) = AB + AC$,
$$(A+B)C = AC + BC$$

For $\alpha \in F$, $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

Special matrices: Already defined the $m \times n$ "zero matrix", $0^m_n \in F^m_n$ whose entries are all $0 \in F$.

Def. Matrices in F^n_n are called "square".

Def. In F^n_n the "identity matrix" is

$I_n = [\delta_{ij}] = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$ with 1 on the "main diagonal", 0 elsewhere.

Th. ① $A I_n = A$ for any $A \in F^m_n$

② $I_m A = A$ " " " "

③ $A O^n_p = O^m_p$, $\forall A \in F^m_n$

④ $O^m_n A = O^m_p$, $\forall A \in F^n_p$.

⑤ $L_{I_n} : F^n \to F^n$ is the identity map I_{F^n}

⑥ $L_{O^m_n} : F^n \to F^m$ is the "zero map" s.t.

$L_{O^m_n}(X) = O^m_1$.

Def. For $A=[a_{ij}]\in F^m_n$ define the **transpose of A** to be $B=[b_{ji}]\in F^n_m$ s.t. $b_{ji}=a_{ij}$, and denote this $n\times m$ matrix by A^T.

Th: For appropriate size matrices we have

① $(A+B)^T=A^T+B^T$ ② $(\alpha A)^T=\alpha(A^T)$, $\alpha\in F$

③ $(A^T)^T=A$ ④ $(AB)^T=B^TA^T$

Def. Say A is **symmetric** when $A^T=A$

so $m=n$ for such a matrix, it must be square.

Say A is anti-symmetric (skew-symmetric) when $A^T=-A$. (Such an A must be square)

<u>Question:</u> For $A \in F^n_n$ (square), when is $L_A : F^n \to F^n$ invertible? What is the condition on A for this to happen?

<u>Answer:</u> There must be an inverse function $L_A^{-1} : F^n \to F^n$ such that $L_A \circ L_A^{-1} = I_{F^n} = L_A^{-1} \circ L_A$. If $L_A^{-1} = L_B$ for some $B \in F^n_n$ this would mean $L_A \circ L_B = L_{I_n} = L_B \circ L_A$ so $L_{AB} = L_{I_n} = L_{BA}$ so $AB = I_n = BA$.

<u>Def.</u> For $A \in F^n_n$ say A is <u>invertible</u> when $\exists B \in F^n_n$ such that $AB = I_n = BA$.

Problem: Given $A \in F^n_n$ determine whether or not A is invertible, and find B if it is.

Th: If $A \in F^n_n$ is invertible, there is only one (unique) matrix B such that $AB = I_n = BA$ so we can denote it by A^{-1} if it exists.

Pf. Suppose we have two candidates for an inverse of A, say $AB = I_n = BA$ and $AC = I_n = CA$. Then, by assoc. of matrix mult.,

$$C = CI_n = C(AB) = (CA)B = I_n B = B. \quad \square$$

Example: Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in F^2_2$. Compute

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} (ad-bc) & 0 \\ 0 & (ad-bc) \end{bmatrix} = (ad-bc) I_2$$

and $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} (ad-bc) & 0 \\ 0 & (ad-bc) \end{bmatrix} = (ad-bc)I_2$

So if $ad-bc \neq 0$ in F, it has a mult. inverse (reciprocal) in F denoted by $(ad-bc)^{-1} = \dfrac{1}{ad-bc}$ and if we multiply these equations by it, we get

$$A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

<u>Exercise</u>: Show that for $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in F_2^2$ if $ad-bc=0$ then A is not invertible.

<u>Def.</u> Let $\det(A) = \det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad-bc$ for $A \in F_2^2$. It "determines" if A is invertible

The discussion above leaves several open questions:

① Why should we expect that $L_A^{-1} = L_B$ for some B ?

② If $L : V \longrightarrow W$ is any invertible linear map, why should $L^{-1} : W \rightarrow V$ also be linear?

③ If $L_A : F^n \rightarrow F^m$ is invertible, why should $m = n$?

④ If $L : F^n \rightarrow F^m$ is linear why should $L = L_A$ for some $A \in F^m_n$?

⑤ How do we find $A^{-1} \in F^n_n$ if it exists?