Math 672 Lecture 10

<u>Definition:</u> Let X, Y be sets. Then $f: X \to Y$ is said to be <u>invertible</u> iff there exists $g: Y \to X$ such that

$g \circ f$ is the identity on X (that is, $g(f(x)) = x \ \forall x \in X$)

and

$f \circ g$ is the identity on Y (that is, $f(g(y)) = y \ \forall y \in Y$).

The function g, if it exists, is called an <u>inverse</u> of f.

<u>Theorem</u>: Let X, Y be sets and consider $f: X \to Y$.

1. If f is invertible, then it has a unique inverse (which we will denote by f^{-1}).
2. f is invertible iff it is bijective.

<u>Proof:</u> 1. Suppose f is invertible and suppose $g_1: Y \to X$, $g_2: Y \to X$ are inverses of f. Then, if $y \in Y$, then

$$\begin{aligned} g_1(y) &= g_1(f(g_2(y))) \quad (\text{since } f(g_2(y)) = y \ \forall y \in Y) \\ &= (g_1 \circ f)(g_2(y)) \\ &= g_2(y). \end{aligned}$$

Thus $g_1 = g_2$, that is f has a unique inverse.

2. Suppose first that f is invertible. We must show that f is injective and surjective:

$$x_1, x_2 \in X \text{ and } f(x_1) = f(x_2) \Rightarrow f^{-1}(f(x_1)) = f^{-1}(f(x_2))$$
$$\Rightarrow x_1 = x_2 \quad (\text{thus } f \text{ is injective}),$$

$$y \in Y \Rightarrow y = f(f^{-1}(y)) \Rightarrow y = f(x), \text{ where } x = f^{-1}(y)$$
$$(\text{thus } f \text{ is surjective}).$$

Conversely, suppose f is bijective. Then, for all $y \in Y$, there exists a unique $x \in X$ such that $f(x) = y$. Hence, we can define $g: Y \to X$ by the condition that

$$g(y) = x \iff f(x) = y,$$

and g is well defined. It is then easy to show that g is an inverse (and hence the inverse) of f:

$$x \in X \Rightarrow g(f(x)) = g(y), \text{ where } y = f(x) \text{ and } g(y) = x$$
$$\Rightarrow g(f(x)) = x,$$
$$y \in Y \Rightarrow f(g(y)) = f(x), \text{ where } x = g(y) \text{ and } y = f(x)$$
$$\Rightarrow f(g(y)) = y.$$

This completes the proof. //

All of the above applies to functions in general. Here is a critical fact about linear maps in particular.

<u>Theorem</u>: Let $T \in \mathcal{L}(V,W)$. If T is invertible, then T^{-1} is linear ($T^{-1} \in \mathcal{L}(W,V)$).

<u>Proof:</u> Assume T is invertible, let $w_1, w_2 \in W$, and let $\alpha_1, \alpha_2 \in F$.

Then

$$\begin{aligned}
T^{-1}(\alpha_1 w_1 + \alpha_2 w_2) &= T^{-1}(\alpha_1 T(T^{-1}(w_1)) + \alpha_2 T(T^{-1}(w_2))) \\
&= T^{-1}(T(\alpha_1 T^{-1}(w_1) + \alpha_2 T^{-1}(w_2))) \quad (\text{since } T \text{ is linear}) \\
&= \alpha_1 T^{-1}(w_1) + \alpha_2 T^{-1}(w_2) \quad (\text{since } T^{-1}(T(w)) = w \ \forall w \in W).
\end{aligned}$$

This proves that T^{-1} is linear. //

<u>Definition</u>: Let V, W be vector spaces over a field F. We say that V and W are <u>isomorphic</u> iff there exists an invertible linear map $T: V \to W$. If such a map exists, it is called an <u>isomorphism</u> from V to W.

Theorem:

1. If $T \in \mathcal{L}(V,W)$ is an isomorphism from V to W, then T^{-1} is an isomorphism from W to V.
2. If $T \in \mathcal{L}(V,W)$ and $S \in \mathcal{L}(W,Z)$ are isomorphisms, then $ST \in \mathcal{L}(V,Z)$ is an isomorphism.
3. "Is isomorphic to" defines an equivalence relation on the set of all vector spaces over a given field F.

Proof: 1. This follows immediately from the previous theorem and the fact that if T is invertible, then T^{-1} is invertible (and $(T^{-1})^{-1} = T$).

2. We have already verified that the product of linear maps is linear. For functions in general, the composition of two bijections is bijective:

$$v_1, v_2 \in V \text{ and } (ST)(v_1) = (ST)(v_2) \Rightarrow S(T(v_1)) = S(T(v_2))$$

$$\Rightarrow T(v_1) = T(v_2) \quad (\text{since } S \text{ is injective})$$

$$\Rightarrow v_1 = v_2 \quad (\text{since } T \text{ is injective})$$

(thus ST is injective)

$$z \in Z \Rightarrow \exists\, w \in W,\ S(w) = z \quad (\text{since } S \text{ is surjective})$$

$$\Rightarrow \exists\, v \in V,\ T(v) = w \quad (\text{since } T \text{ is surjective})$$

$$\Rightarrow \exists\, v \in V,\ (ST)(v) = z \quad (\text{since } (ST)(v) = S(T(v)) = S(w) = z)$$

(thus ST is surjective).

Thus ST is an isomorphism.

3. Write $V \cong W$ to mean that vector spaces V and W are isomorphic.

   - $V \cong V$ $\forall V$ (since the identity operator $I: V \to V$ is an isomorphism)
   - $V \cong W \Rightarrow W \cong V$ (since if $T \in \mathcal{L}(V,W)$ is an isomorphism, then $T^{-1} \in \mathcal{L}(W,V)$ is an isomorphism)
   - $V \cong W$ and $W \cong Z \Rightarrow V \cong Z$ (since if $T \in \mathcal{L}(V,W)$ and $S \in \mathcal{L}(W,Z)$ are isomorphisms, then $ST \in \mathcal{L}(V,Z)$ is an isomorphism).

This completes the proof. //

Theorem: Let V and W be finite-dimensional vector spaces over F. Then $V \cong W$ iff $\dim(V) = \dim(W)$.

Proof: By the previous result, it suffices to prove that if V is an n-dimensional vector space over F, then V is isomorphic to F^n.

Let $\{v_1, v_2, \ldots, v_n\} \subseteq V$ be a basis for V, let $\{e_1, e_2, \ldots, e_n\}$ be the standard basis for F^n ($(e_i)_j = 1$ if $i=j$ and $(e_i)_j = 0$ if $i \neq j$), and define $\mathcal{M} \in \mathcal{L}(V, F^n)$ by

$$\mathcal{M}(v_j) = e_j, \quad j = 1, 2, \ldots, n.$$

By an earlier theorem, $\mathcal{M}$ is well defined (and linear). It remains only to show that $\mathcal{M}$ is bijective.

Suppose first that $\mathcal{M}(v)=0$ and $v=\sum_{j=1}^{n}\alpha_j v_j$. Then

$$\mathcal{M}(v)=0$$

$$\Rightarrow \mathcal{M}\Big(\sum_{j=1}^{n}\alpha_j v_j\Big)=0$$

$$\Rightarrow \sum_{j=1}^{n}\alpha_j e_j=0$$

$$\Rightarrow \alpha_j=0 \ \forall j=1,2,\dots,n \ \text{(since } \{e_1,e_2,\dots,e_n\} \text{ is linearly independent)}$$

$$\Rightarrow v=0.$$

Thus $\mathcal{N}(\mathcal{M})=\{0\}$, which implies that $\mathcal{M}$ is injective.

If $x\in F^n$, then

$$x=\sum_{j=1}^{n}x_j e_j=\sum_{j=1}^{n}x_j\mathcal{M}(v_j)=\mathcal{M}\Big(\sum_{j=1}^{n}x_j v_j\Big)$$

$$\Rightarrow x\in\mathcal{R}(\mathcal{M}).$$

Thus $\mathcal{M}$ is surjective, and the proof is complete. //

Definition: Let V be a finite-dimensional vector space with basis $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$. The notation $\mathcal{M}: V \to F^n$ (or $\mathcal{M}_{\mathcal{B}}: V \to F^n$ if we need to emphasize the particular basis used) denotes the isomorphism used in the previous proof:

$$\mathcal{M}\left(\sum_{j=1}^{n} \alpha_j v_j\right) = \sum_{j=1}^{n} \alpha_j e_j = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in F^n.$$

(Our author calls $\mathcal{M}(v)$ "the matrix of v with respect to the basis $\mathcal{B}$." This terminology is unusual; most authors call it the coordinate vector of v w.r.t. $\mathcal{B}$. It is often denoted $[v]_{\mathcal{B}}$ instead of $\mathcal{M}(v)$. Although I don't like Axler's terminology, I prefer his notation, which emphasizes the map $\mathcal{M}$.)

Theorem: Let V, W be finite-dimensional vector spaces over F with bases $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$, $\mathcal{C} = \{w_1, w_2, \ldots, w_m\}$, respectively, and let $T \in \mathcal{L}(V, W)$. Then there exists a unique matrix $A \in F^{m \times n}$ such that

$$\mathcal{M}_{\mathcal{C}}(T(v)) = A\mathcal{M}_{\mathcal{B}}(v) \quad \forall\, v \in V.$$

Moreover, the columns of A are $\left[\mathcal{M}_{\mathcal{C}}(T(v_1)) \,|\, \mathcal{M}_{\mathcal{C}}(T(v_2)) \,|\, \cdots \,|\, \mathcal{M}_{\mathcal{C}}(T(v_n))\right]$.

Proof: Let us define $A \in F^{m \times n}$ by

$$A = [\mathcal{M}_{\mathcal{C}}(T(v_1)) | \mathcal{M}_{\mathcal{C}}(T(v_2)) | \cdots | \mathcal{M}_{\mathcal{C}}(T(v_n))].$$

Recall that $\mathcal{M}_{\mathcal{B}}, \mathcal{M}_{\mathcal{C}}$ are linear. Let $v \in V$ and suppose $x = \mathcal{M}_{\mathcal{B}}(v)$, that is, suppose $v = \sum_{j=1}^{n} x_j v_j$. Then

$$\begin{aligned}
A\mathcal{M}_{\mathcal{B}}(v) = Ax &= \sum_{j=1}^{n} x_j \mathcal{M}_{\mathcal{C}}(T(v_j)) \\
&= \mathcal{M}_{\mathcal{C}}\left(\sum_{j=1}^{n} x_j T(v_j)\right) \quad \text{(since } \mathcal{M}_{\mathcal{C}} \text{ is linear)} \\
&= \mathcal{M}_{\mathcal{C}}\left(T\left(\sum_{j=1}^{n} x_j v_j\right)\right) \quad \text{(since } T \text{ is linear)} \\
&= \mathcal{M}_{\mathcal{C}}(T(v)).
\end{aligned}$$

Thus A satisfies

$$A\mathcal{M}_{\mathcal{B}}(v) = \mathcal{M}_{\mathcal{C}}(T(v)) \quad \forall v \in V,$$

as desired.

Now suppose $B \in F^{m \times n}$ also satisfies

$$B\mathcal{M}_{\mathcal{B}}(v) = \mathcal{M}_{\mathcal{C}}(T(v)) \quad \forall v \in V.$$

Then

$$A\,\mathcal{M}_{\mathcal{B}}(v) = B\,\mathcal{M}_{\mathcal{B}}(v) \quad \forall v \in V$$

$$\Rightarrow A\,\mathcal{M}_{\mathcal{B}}(v_j) = B\,\mathcal{M}_{\mathcal{B}}(v_j) \quad \forall j = 1,2,\dots,n$$

$$\Rightarrow Ae_j = Be_j \quad \forall j = 1,2,\dots,n \quad (\text{since } \mathcal{M}_{\mathcal{B}}(v_j) = e_j \text{ by def'n})$$

$$\Rightarrow A_j = B_j \quad \forall j = 1,2,\dots,n \quad (\text{that is, } A \text{ and } B \text{ have the same columns})$$

$$\Rightarrow A = B.$$

(Here we used the fact that for any matrix $M \in F^{m \times n}$, $Me_j = M_j$ if e_j is the jth standard basis vector for F^n.)

Thus A is unique, and the proof is complete. //

We write $\mathcal{M}(T)$ or $\mathcal{M}_{\mathcal{B},\mathcal{C}}(T)$ for the matrix A of the previous theorem, and call $\mathcal{M}_{\mathcal{B},\mathcal{C}}(T)$ the matrix of the linear map T with respect to the bases $\mathcal{B}$ of V and $\mathcal{C}$ of W. Thus

$$\mathcal{M}_{\mathcal{C}}(T(v)) = \mathcal{M}_{\mathcal{B},\mathcal{C}}(T)\,\mathcal{M}_{\mathcal{B}}(v) \quad \forall v \in V.$$

$\mathcal{M}_{\mathcal{B},\mathcal{C}}(T)$ represents T in the sense of the following commutative diagram:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\mathcal{M}_{\mathcal{B}}"'] & W \\ F^n \arrow[r, "\mathcal{M}_{\mathcal{B},\mathcal{C}}(T)"'] & F^m \arrow[u, "\mathcal{M}_{\mathcal{C}}^{-1}"'] \end{tikzcd}$$

$$\mathcal{M}_{\mathcal{C}} T = \mathcal{M}_{\mathcal{B},\mathcal{C}}(T)\mathcal{M}_{\mathcal{B}} \iff T = \mathcal{M}_{\mathcal{C}}^{-1}\mathcal{M}_{\mathcal{B},\mathcal{C}}\mathcal{M}_{\mathcal{B}}$$