Lecture 11:
Isomorphisms
Dot product
Matrices as linear functions

Warmup.

1. Let $f : X \to Y$ be a function. Recall/prove each of the following statements.
   (a) The function f is injective if and only if it has a left inverse, i.e. a function $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$.
   (b) The function f is surjective if and only if is has a right inverse, i.e. a function $h : Y \to X$ such that $f \circ h = \mathrm{id}_Y$.
   (c) If f has both a left inverse $g : Y \to X$ and a right inverse $h : Y \to X$, then $g = h$. (In this case, we say f is invertible and write $g = h = f^{-1}$. From (a) and (b), we know that f is invertible if and only if it's bijective.)
   (d) Suppose $f : X \to Y$ and $g : Y \to Z$ are both bijective functions. Then $g \circ f$ is also bijective.
2. Let $f : U \to V$ and $g : V \to W$ be linear functions (where U, V, W are all vector spaces over F). Prove that $g \circ f : U \to W$ is also linear.

Injective linear functions (Let U and V denote vector spaces over F.)

Very Useful Theorem 1.
A linear function $h : U \to V$ is injective if and only if $\mathcal{N}(h) = 0$.

Proof.
($\Rightarrow$) Suppose h is injective. Compute $\mathcal{N}(h)$.

($\Leftarrow$) Suppose $\mathcal{N}(h) = 0$.
Suppose $h(\mathbf{x}) = h(\mathbf{y})$ for some $\mathbf{x}, \mathbf{y} \in U$.

Corollary 2. If $h : U \to V$ is linear and V is finite-dimensional, then the following are equivalent:

1. h is injective;
2. $\text{nullity}(h) = 0$;
3. $\text{rank}(h) = \dim(U)$;
4. If $\mathcal{B}$ is a basis for V, then $h(\mathcal{B})$ is a basis for $\mathcal{R}(h)$ (as a multiset).

(1 ⇔ 2: Thm. 1)
(2 ⇔ 3: Rank-nullity: $\dim(U) = \text{rank}(h) + \text{nullity}(h)$)
(See our proof of Rank-Nullity/homework.)

Isomorphisms (Let U and V denote vector spaces over F.)

We call a bijective linear function an isomorphism.

Example. Given an ordered basis $B = \langle \mathbf{b}_1, \ldots, \mathbf{b}_n \rangle$ of a vector space V, the representation $\text{Rep}_B : V \to F^n$, given by
$$c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n \longmapsto (c_1, \ldots, c_n),$$
is an isomorphism.
For example, using the standard ordered bases, we have the isomorphisms
$$\mathcal{P}_n(F) \to F^{n+1} \quad \text{defined by} \quad c_0 + c_1x + \cdots + c_nx^n \mapsto (c_0, c_1, \ldots, c_n);$$
and
$$M_2(F) \to F^4 \quad \text{defined by} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a, b, c, d).$$

Lemma 3.
If $h : U \to V$ is an isomorphism, then $h^{-1} : V \to U$ is also an isomorphism.

Proof: Exercise! *Hint:* Similarly to the warmup,
$$f^{-1}(\mathbf{v} + \mathbf{w}) = f^{-1}(f(f^{-1}(\mathbf{v})) + f(f^{-1}(\mathbf{w})))\ldots$$

Isomorphisms

(Let U and V denote vector spaces over F.)

We say that U is isomorphic to V if there exists an isomorphism $h : U \to V$. If so, we write $\boxed{U \cong V}$. (There might be lots of iso's!)

Examples: We just saw that $P_n(F) \cong F^{n+1}$ and $M_2(F) \cong F^4$.

Theorem 4. Isomorphism gives an equivalence relation on the set of vector spaces over a field F. Namely, the relation

$$U \sim V \quad \text{whenever} \quad U \cong V,$$

is reflexive, symmetric, and transitive. (See warmup & Lemma 3.)

AMAZING Thm 5. For vector spaces U and V over a field F,

$$U \cong V \quad \text{if and only if} \quad \dim(U) = \dim(V).$$

Proof. ($\Rightarrow$) Suppose there is an isomorphism $h : U \to V$. Let $\mathcal{B}$ be a basis of U. Then by Cor. 2 #4, $h(\mathcal{B})$ is a basis of V (as a multiset), and

$$\dim(U) = |\mathcal{B}| = |h(\mathcal{B})| = \dim(V).$$

($\Leftarrow$) Suppose $\dim(U) = \dim(V)$. Let $\mathcal{A}$ and $\mathcal{B}$ be bases of U and V, respectively. Namely, $|\mathcal{A}| = |\mathcal{B}|$ (i.e. there's a bijective function $f : \mathcal{A} \to \mathcal{B}$).

Pf 1: If $|\mathcal{A}| = |\mathcal{B}| = n$, then both U and V are isomorphic to F^n by $\mathrm{Rep}_{\mathcal{A}}$ and $\mathrm{Rep}_{\mathcal{B}}$, respectively. Hence $U \cong V$ (by Thm. 4).

Pf 2: Since $\mathcal{B} \subseteq V$, we can linearly extend $f : \mathcal{A} \to \mathcal{B} \subseteq V$ to a linear function $h : U \to V$ that has $h(\mathbf{a}) = f(\mathbf{a}) \in \mathcal{B}$ for all $\mathbf{a} \in \mathcal{A}$. By Cor. 2, h is an isomorphism. Hence $U \cong V$. $\square$

Isomorphisms

(Let U and V denote vector spaces over F.)

Corollary 6. If V is a **finite-dimensional** vector space, and $h : V \to V$ is linear map from V **to itself**, then the following are equivalent:

1. h is injective;
2. h is surjective;
3. h is an isomorphism.

We call a homomorphism of the form $h : V \to V$ an endomorphism.

Proof. Use dimension! Three BIG facts:

(I) The null space $\mathcal{N}(h)$ and the range $\mathcal{R}(h)$ are both subspaces of V.

(II) If Y is a vector space and $X \subseteq Y$ is a subspace, then $\dim(X) \leqslant \dim(Y)$. And if Y is *finite dimensional*, then

$$\dim(X) = \dim(Y) \quad \text{if and only if} \quad X = Y.$$

(III) Thm. 5: $U \cong V$ if and only if $\dim(U) = \dim(V)$.

Caution: If V is infinite-dimensional, there are linear maps that are injective but not surjective, and vice versa.

Examples:

$$f : \mathbb{R}[x] \to \mathbb{R}[x] \text{ by } p(x) \mapsto p(x^2)$$

$$d : \mathbb{R}[x] \to \mathbb{R}[x] \text{ by } p(x) \mapsto \tfrac{d}{dx} p(x)$$

$$\iota : \mathbb{R}[x] \to \mathbb{R}[x] \text{ by } p(x) \mapsto \textstyle\int_0^x p(t)\, dt$$

Back to matrices!

The augmented matrix

$$\left(\begin{array}{cccc|c} 3 & 2 & -1 & 0 & b_1 \\ 4 & 0 & -5 & 1 & b_2 \\ 0 & 1 & 2 & 3 & b_3 \end{array}\right)$$

encodes the linear system

$$\begin{pmatrix} 3x_1 & + & 2\ x_2 & + & (-1)x_3 & + & 0\ x_4 \\ 4x_1 & + & 0\ x_2 & + & (-5)x_3 & + & (1)x_4 \\ 0x_1 & + & (1)x_2 & + & 2\ \ x_3 & + & 3\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

Define the dot product of two vectors in F^n by $\cdot : F^n \times F^n \to F$

$$(c_1, c_2, \ldots, c_n) \cdot (d_1, d_2, \ldots, d_n) := c_1 d_1 + c_2 d_2 + \cdots + c_n d_n.$$

For example, for any $x_1, \ldots, x_n \in F$,

$$\begin{aligned} (3, 2, -1, 0) \cdot (x_1, x_2, x_3, x_4) &= 3x_1 + 2x_2 + (-1)x_3 + 0x_4, \\ (4, 0, -5, 1) \cdot (x_1, x_2, x_3, x_4) &= 4x_1 + 0x_2 + (-5)x_3 + (1)x_4, \text{ and} \\ (0, 1, \ \ 2, 3) \cdot (x_1, x_2, x_3, x_4) &= 0x_1 + (1)x_2 + 2x_3 \ + \ 3x_4. \end{aligned}$$

So the above linear system is *also* encoded by

$$\begin{pmatrix} (3,\ 2,\ -1,\ 0) \cdot (x_1, x_2, x_3, x_4) \\ (4,\ 0,\ -5,\ 1) \cdot (x_1, x_2, x_3, x_4) \\ (0,\ 1,\ \ \ 2,\ 3) \cdot (x_1, x_2, x_3, x_4) \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Notice

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} (3,\ 2,\ -1,\ 0) \cdot (x_1, x_2, x_3, x_4) \\ (4,\ 0,\ -5,\ 1) \cdot (x_1, x_2, x_3, x_4) \\ (0,\ 1,\ \ \ 2,\ 3) \cdot (x_1, x_2, x_3, x_4) \end{pmatrix}$$

defines a function $h : F^4 \mapsto F^3$! For example,

$$h\left(\begin{pmatrix} 1 \\ 5 \\ -2 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} (3,\ 2,\ -1,\ 0) \cdot (1, 5, -2, 0) \\ (4,\ 0,\ -5,\ 1) \cdot (1, 5, -2, 0) \\ (0,\ 1,\ \ \ 2,\ 3) \cdot (1, 5, -2, 0) \end{pmatrix}$$

$$= \begin{pmatrix} 3(1) + \ 2\ (5) + (-1)(-2) + \ 0\ (0) \\ 4(1) + \ 0\ (5) + (-5)(-2) + (1)(0) \\ 0(1) + (1)(5) + \ \ 2\ \ (-2) + \ 3\ (0) \end{pmatrix} = \begin{pmatrix} 15 \\ 14 \\ 1 \end{pmatrix}$$

It turns out that this is a linear function! [See HW 6 for properties of dot product.] Specifically, h is the linear extension of the function

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} -1 \\ -5 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}.$$

To a matrix $A \in M_{k,\ell}(F)$,

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,\ell} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,\ell} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k,1} & a_{k,2} & \cdots & a_{k,\ell} \end{pmatrix},$$

we associate a function $A : F^\ell \to F^k$ given by

$$A\colon \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_\ell \end{pmatrix} \mapsto \begin{pmatrix} (a_{1,1}, \ldots, a_{1,\ell}) \cdot (v_1, \ldots, v_\ell) \\ (a_{2,1}, \ldots, a_{2,\ell}) \cdot (v_1, \ldots, v_\ell) \\ \vdots \\ (a_{k,1}, \ldots, a_{k,\ell}) \cdot (v_1, \ldots, v_\ell) \end{pmatrix} \quad \text{i.e.} \quad A\colon \mathbf{v} \mapsto \begin{pmatrix} \mathbf{r}_1 \cdot \mathbf{v}^T \\ \mathbf{r}_2 \cdot \mathbf{v}^T \\ \vdots \\ \mathbf{r}_k \cdot \mathbf{v}^T \end{pmatrix},$$

where $\mathbf{r}_i$ is the ith row vector of A, and $\mathbf{v}^T$ is the transpose of $\mathbf{v}$ (from Homework 4). *Notation:* write $A\mathbf{v}$ to mean $A(\mathbf{v})$.

Example. Back to our example from before,

$$\left(\begin{array}{cccc|c} 3 & 2 & -1 & 0 & b_1 \\ 4 & 0 & -5 & 1 & b_2 \\ 0 & 1 & 2 & 3 & b_3 \end{array}\right) \quad \text{and} \quad \begin{pmatrix} 3 & 2 & -1 & 0 \\ 4 & 0 & -5 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

mean the same thing!

You try: Let

$$A = \begin{pmatrix} 0 & 1 & 3 \\ -1 & 2 & -1 \end{pmatrix} : F^3 \mapsto F^2.$$

(a) For $\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in F^3$, compute $A\mathbf{v}$.

(b) Recall that the standard (ordered) basis of F^3 is

$$\mathcal{E} = \langle \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \rangle, \quad \text{where} \quad \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \ \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Compute $A\mathbf{e}$ for each $\mathbf{e} \in \mathcal{E}$.

(c) Compute $\mathcal{N}(A)$ (the vectors $\mathbf{v} \in F^3$ for which $A\mathbf{v} = \mathbf{0}_{F^2}$).

(d) What did you notice about your answer to part (b)? Can you prove your answer for a general matrix?

(e) How does the range of the function A relate to spaces we've studied before having to do with matrices?

(f) How does the rank of the function A relate to statistics we've studied before having to do with matrices?

(g) How does the nullspace of A relate spaces we've studied before having to do with matrices? nullity?

Solutions to warmup

1. Let $f : X \to Y$ be a function. Recall/prove each of the following statements.
   (a) The function f is injective if and only if it has a left inverse, i.e. a function $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$.
   ($\Rightarrow$) Build g! For each $y \in f(X)$, there is exactly on $x_y \in X$ w/ $f(x_y) = y$; define $g(y) = x_y$. And for $y \in Y - f(Y)$, define $g(y)$ arbitrarily.
   ($\Leftarrow$) If $f(x) = f(x')$, then $x = g(f(x)) = g(f(x')) = x'$.
   (b) The function f is surjective if and only if is has a right inverse, i.e. a function $h : Y \to X$ such that $f \circ h = \mathrm{id}_Y$.
   ($\Rightarrow$) Build h! For each $y \in Y$, choose one $x_y \in f^{-1}(y)$, and define $h(y) = x_y$.
   ($\Leftarrow$) If $y \in Y$, then $f(h(y)) = y$, so $y \in f(X)$.
   (c) If f has both a left inverse $g : Y \to X$ and a right inverse $h : Y \to X$, then $g = h$. (In this case, we say f is invertible and write $g = h = f^{-1}$. From (a) and (b), we know that f is invertible if and only if it's bijective.)
   Pf. We have $h = \mathrm{id}_Y \circ h = (g \circ f) \circ h = g \circ (f \circ h) = g \circ \mathrm{id}_Y = g$.
   (d) Suppose $f : X \to Y$ and $g : Y \to Z$ are both bijective functions. Then $g \circ f$ is also bijective.
   Pf. Since f and g are bijective, they have two-sided inverses f^{-1} and g^{-1}. Then $f^{-1} \circ g^{-1}$ is a two-sided inverse for $g \circ f$.

2. Let $f : U \to V$ and $g : V \to W$ be linear functions (where U, V, W are all vector spaces over F). Prove that $g \circ f : U \to W$ is also linear.
   Pf. For all $\mathbf{u}, \mathbf{v} \in U$ and $\lambda \in F$, we have

$$\begin{aligned}(g \circ f)(\mathbf{u} + \mathbf{v}) &= g(f(\mathbf{u} + \mathbf{v})) \\ &= g(f(\mathbf{u}) + f(\mathbf{v})) \\ &= g(f(\mathbf{u})) + g(f(\mathbf{v})) \\ &= (g \circ f)(\mathbf{u}) + (g \circ f)(\mathbf{v})\end{aligned}$$

   and

$$\begin{aligned}(g \circ f)(\lambda \mathbf{u}) &= g(f(\lambda \mathbf{u})) \\ &= g(\lambda f(\mathbf{u})) \\ &= \lambda g(f(\mathbf{u})) \\ &= \lambda (g \circ f)(\mathbf{u}).\end{aligned}$$