3-11-1994

Linear transformations.

Definition. Let V and W be F-vector spaces. A **linear transformation** $f : V \to W$ is a function from V to W which satisfies

$$f(au + v) = af(u) + f(v) \quad \text{for all} \quad u, v \in V, \quad a \in F.$$

In particular, f satisfies

$$f(u + v) = f(u) + f(v) \qquad u, v \in V,$$

$$f(au) = af(u) \qquad a \in F, \quad u \in V.$$

Conversely, it is clear that if these two equations are satisfied then f is a linear transformation.

The first of these two equations says that f is a group homomorphism from the additive group $(V, +)$ to the additive group $(W, +)$. *If a result is true about group maps, it is true about linear transformations.*

For example, the following results are true for homomorphisms of abelian groups:

$$f(0) = 0 \quad \text{and} \quad f(-v) = -f(v) \quad \text{for all} \quad v \in V.$$

Hence, they are true for linear transformations as well.

Linear transformations are to vector spaces as group maps are to groups, and as ring maps are to rings. In each case, the functions under consideration *preserve the structure of the objects under study.* Vector spaces are characterized by an addition and a scalar multiplication. Hence, maps of vector spaces should be functions which "commute" with addition and scalar multiplication.

Example. If V is a vector space, the identity map $\mathrm{id} : V \to V$ is a linear transformation.

If V and W are vector spaces, the zero map $0 : V \to W$ given by $0(u) = 0$ for all $u \in V$ is a linear transformation. □

Example. If W is a subspace of a vector space V, the inclusion map $i : W \hookrightarrow V$ is a linear transformation. □

Example. The composite of linear transformations is a linear transformation. □

Example. (**Quotient maps**) If W is a subspace of V, then $(W, +)$ is a subgroup of $(V, +)$. Form the quotient group

$$V/W = \{v + W \mid v \in V\}.$$

V/W is an abelian group under coset multiplication. It becomes an F-vector space if scalar multiplication is defined in the obvious way:

$$a \cdot (v + W) = av + W, \quad \text{for} \quad a \in F, v \in V.$$

It is easy to check that the axioms for a vector space hold.

The quotient map $\pi : V \to V/W$ given by

$$\pi(v) = v + W$$

is a group homomorphism. This means that $\pi(u + v) = \pi(u) + \pi(v)$.

To show that π is a linear transformation, it therefore suffices to show that "constants pull out":

$$\pi(a \cdot v) = av + W = a \cdot (v + W) = a \cdot \pi(v).$$

Therefore, π is a linear transformation. □

Example. Let F be a field, and consider the n-dimensional vector space F^n. Let $A = \{a_{ij}\}$ be an $m \times n$ matrix with entries in F. A gives rise to a linear transformation $T : F^n \to F^m$ by matrix multiplication. That is, if $v \in F^n$, define

$$T(v) = Av.$$

Standard properties of matrices show that this is linear: If $u, v \in F^n$ and $c \in F$, then

$$T(au + v) = A(cu + v) = c \cdot Au + Av = c \cdot T(u) + T(v).$$

It works the other way as well. A linear transformation of finite dimensional vector spaces gives rise to matrices — one matrix for each choice of bases for the domain and range. □

Example. Let $f : \mathbb{R}^n \to \mathbb{R}^m$. f is **differentiable** at p if there is a linear transformation $Df(p) : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h \to 0} \frac{|f(p + h) - f(p) - Df(p) \cdot h|}{|h|} = 0.$$

$Df(p)$ is the derivative of f at p.

If f is differentiable at p, the partials of f exist, and

$$Df(p) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix}.$$

For example, consider $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by $f(x, y) = (x^2 - y^2, 3xy)$. Then

$$Df = \begin{bmatrix} 2x & -2y \\ 3y & 3x \end{bmatrix}.$$

Notice that you don't get a linear transformation here until you plug in a point (x, y).

Many other example follow from this one. For example, consider the vector space $\mathbb{R}[x]$ of polynomials with real coefficients. The the differentiation operator $D = \dfrac{d}{dx}$ is a linear transformation $\mathbb{R}[x] \to \mathbb{R}[x]$. E.g.,

$$D(x^2 - 2x + 1) = 2x - 2.$$

Let V be the vector space of differentiable functions $\mathbb{R} \to \mathbb{R}$. If $p(x) \in \mathbb{R}[x]$, then $p(D)$ is a linear transformation, where D^n is interpreted as the n-th derivative with respect to x. Solving a differential equation $p(D)y = 0$ amounts to finding the **kernel** of $p(D)$ — the set of functions y which are killed by the linear transformation.

For example, consider the differential equation

$$(D^2 - 2D + 1)y = 0.$$

The general solution is

$$y = c_1 e^x + c_2 x e^x.$$

The functions e^x, xe^x form a basis for the kernel of $p(D) = D^2 - 2D + 1$. $\square$

Definition. Let $f : V \to W$ be a linear transformation. f is **invertible** if there is a linear transformation $g : W \to V$ such that

$$fg = \mathrm{id}_W \quad \text{and} \quad gf = \mathrm{id}_V.$$

In this case, I'll write $g = f^{-1}$, and say that g is the **inverse** of f.

Remarks.

1. If $f : V \to W$ is a linear transformation and $g : W \to V$ is an inverse for f *as a map of sets*, then g is automatically a linear transformation. $\square$

2. A linear transformation is invertible if and only if it's bijective. For a set map is invertible if and only if it's bijective, and the inverse of a linear transformation is a linear transformation by the first remark. $\square$

Definition. Let V and W be F-vector spaces. V and W are **isomorphic** if there is an invertible linear transformation $f : V \to W$. Notation: $V \approx W$.

Lemma. Let V and W be F-vector spaces. Let $\{v_i\}$ be a basis for V, and let $\{w_i\}$ be a subset of W (with the same index set as $\{v_i\}$). There is a unique linear transformation $f : V \to W$ satisfying

$$f(v_i) = w_i.$$

Proof. First, recall that every element $v \in V$ may be uniquely written as a linear combination of vectors in $\{v_i\}$:

$$v = a_1 v_{i_1} + \cdots + a_n v_{i_n}.$$

Therefore, I may define

$$f(v) = a_1 w_{i_1} + \cdots + a_n w_{i_n}.$$

I claim f is a linear transformation. This is actually obvious, but requires some fussing to write out. The fussing proceeds below.

Let $u \in V$. I may write u as a linear combination of vectors in $\{v_i\}$:

$$u = b_1 v_{j_1} + \cdots + b_m v_{j_m}.$$

Collect all the vectors $\{v_{i_1}, \ldots, v_{i_n}, v_{j_1}, \ldots, v_{j_m}\}$. There are finitely many all together. Hence, I may reindex them and write both u and v as linear combinations of vectors in a set $\{v_{k_1}, \ldots, v_{k_p}\} \subset \{v_i\}$:

$$u = c_1 v_{k_1} + \cdots + c_p v_{k_p},$$

$$v = d_1 v_{k_1} + \cdots + d_p v_{k_p}.$$

Now let $s \in F$. Then

$$f\left(s \sum c_l v_{k_l} + \sum d_l v_{k_l}\right) = f\left(\sum (sc_l + d_l) v_{k_l}\right) =$$

$$\sum (sc_l + d_l) w_{k_l} = s \sum c_l w_{k_l} + \sum d_l w_{k_l} =$$

$$f\left(s \sum c_l v_{k_l}\right) + f\left(\sum d_l v_{k_l}\right).$$

Therefore, f is a linear transformation.

Uniqueness is clear, since if $g : V \to W$ is a linear transformation satisfying $g(v_i) = w_i$, then by linearity

$$g\left(\sum s_i v_i\right) = \sum s_i w_i = f\left(\sum s_i v_i\right). \ \square$$

In other words, a linear transformation is determined by where you send a basis.

Theorem. Let V be an n-dimensional vector space over a field F. Then $V \approx F^n$.

Proof. Since $\dim v = n$, there is a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ for V. Let $\mathcal{C} = \ e_1, \ldots, e_n\}$ denote the standard basis for F^n. Define linear transformations $f : V \to F^n$ and $g : F^n \to V$ by

$$f(v_i) = e_i \quad \text{and} \quad g(e_i) = v_i.$$

It is evident that f and g are inverses. Hence, $V \approx F^n$. $\square$

Representation of transformations by matrices.

Let $f : V \to W$ be a linear transformation of finite dimensional vector spaces. Choose ordered bases $\mathcal{B} = \{v_1, \ldots, v_n\}$ for V and $\mathcal{C} = \{w_1, \ldots, w_m\}$ for W.

For each j, $f(v_j) \in W$. Therefore, $f(v_j)$ may be written uniquely as a linear combination of elements of $\mathcal{C}$:

$$f(v_j) = \sum_{i=0}^{m} a_{ij} w_i.$$

The numbers $\{a_{ij}\}$ are uniquely determined by f. The $m \times n$ matrix $A = (a_{ij})$ is the **matrix of f relative to the ordered bases $\mathcal{B}$ and $\mathcal{C}$.** Notation: $A = [f]_{\mathcal{B},\mathcal{C}}$.

Here is a verbal description of A:

- Left multiplication by A takes a vector written in terms of $\mathcal{B}$, applies f, and writes the result in terms of $\mathcal{C}$.

If you keep this in mind, change of coordinates will make much more sense.

I'll verify the claim above for one of the basis elements v_j. In terms of $\mathcal{B}$,

$$v_j = \langle 0, \ \ldots, \ 0, \ \underset{j}{1}, \ 0, \ \ldots, \ 0\rangle.$$

Then

$$\begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

This is correct, since $f(v_j) = \sum_{i=0}^{m} a_{ij} w_i$, and the representation of $\sum_{i=0}^{m} a_{ij} w_i$ in terms of the basis $\mathcal{C} = \{w_1, \ldots, w_m\}$ is

$$\langle a_{1j}, a_{2j}, \ldots, a_{mj}\rangle.$$

The matrix of a linear transformation is like a snapshot of a person — there are many pictures of a person, but only one person. Likewise, a given linear transformation can be represented by matrices with respect to many choices of bases for the domain and range.

In the next example, I'll illustrate how to go from one "snapshot" of a linear transformation to another.

Example. Define $f : \mathbb{R}^2 \to \mathbb{R}^3$ by

$$f \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + y \\ 2x - y \\ x - y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

The matrix above is the matrix of f relative to the standard bases of $\mathbb{R}^2$ and $\mathbb{R}^3$.

Next, consider the bases

$$\mathcal{B} = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\},$$

$$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix} \right\},$$

for $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. I'll find the matrix $[f]_{\mathcal{B},\mathcal{C}}$ of f relative to $\mathcal{B}$ and $\mathcal{C}$.

Here's the schematic picture:

$$\begin{bmatrix} \quad \text{standard} \to \mathcal{C} \quad \end{bmatrix} \begin{bmatrix} \quad f(\text{standard} \to \text{standard}) \quad \end{bmatrix} \begin{bmatrix} \quad \mathcal{B} \to \text{standard} \quad \end{bmatrix}$$

The matrix

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

translates vectors in $\mathbb{R}^2$ from $\mathcal{B}$ to the standard basis.

The matrix

$$\begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 0 \\ 1 & 1 & -2 \end{bmatrix}$$

translates vectors in $\mathbb{R}^3$ from $\mathcal{C}$ to the standard basis.

Hence, the inverse matrix

$$\begin{bmatrix} 4 & -1 & -2 \\ -2 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$$

translates vectors from the standard basis to $\mathcal{C}$.

Therefore,

$$[f]_{\mathcal{B},\mathcal{C}} = \begin{bmatrix} 4 & -1 & -2 \\ -2 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 7 \\ -2 & -3 \\ 2 & 2 \end{bmatrix}. \quad \Box$$

Rank and column rank.

Definition. Let $f : V \to W$ be a linear transformation. The **image** or **range** of f is the range of f as a set map. $\operatorname{im} f$ denotes the image of f, as usual.

Lemma. Let $f : V \to W$ be a linear transformation. $\operatorname{im} f$ is a subspace of W.

Proof. $\operatorname{im} f \neq \emptyset$, since $0 = f(0) \in \operatorname{im} f$.

Let $f(u), f(v) \in \operatorname{im} f$, where $u, v \in V$, and let $a \in F$. Then

$$af(u) + f(v) = f(au + v) \in \operatorname{im} f.$$

Hence, $\operatorname{im} f$ is a subspace of W. $\Box$

When you view a linear transformation as a matrix, there is a concrete way of describing the image of the transformation.

Definition. The **column space** of an $m \times n$ matrix A is the subspace of F^m spanned by the columns of A.

A remark before I state the next result. If A is an $m \times n$ matrix and $V \in F^n$, then the product Av is a linear combination of the columns of A.

In fact, if c_i is the i-th column of A and $v = \langle v_1, \ldots, v_n \rangle$,

$$\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ c_1 & c_2 & \cdots & c_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v_1 c_1 + v_2 c_2 + \cdots + v_n c_n.$$

Lemma. Let $f : V \to W$ be a linear transformation, and let A be a matrix which represents f. Then the image of f is equal to the column space of A.

Proof. Suppose $A = [f]_{\mathcal{B},\mathcal{C}}$, where $\mathcal{B}$ is a basis for V and $\mathcal{C}$ is a basis for W. If $V_{\mathcal{B}}$ denotes the vector $v \in V$ written in terms of $\mathcal{B}$, then

$$f(v) = Av_{\mathcal{B}}.$$

Let **Col** denote the column space of A. By the remark preceding the lemma, $Av_{\mathcal{B}}$ is a linear combination of the columns of A. Therefore, $f(v)$ is in the column space of A, i.e. $\operatorname{im} f \subset \mathbf{Col}$.

Conversely, consider a vector $w \in \mathbf{Col}$. This means that

$$w = a_1 c_1 + \cdots + a_n c_n,$$

where $a_i \in F$ and c_i denotes the i-th column of A.

Let $v = \langle a_1, \ldots, a_n \rangle_{\mathcal{B}}$. Then

$$w = A \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = f\left(\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \right).$$

Therefore, $w \in \operatorname{im} f$. Hence, $\mathbf{Col} \subset \operatorname{im} f$, so $\mathbf{Col} = \operatorname{im} f$. □

Lemma. Let A be a matrix, and let R be the row reduced echelon matrix which is row equivalent to A. Suppose the leading coefficients of R occur in columns $j_1, \ldots, j_p$, where $j_1 < \cdots < j_p$, and let c_i denote the i-th column of A. Then $\{c_{j_1}, \ldots, c_{j_p}\}$ is independent.

Proof. Suppose that

$$a_{j_1} c_{j_1} + \cdots + a_{j_p} c_{j_p} = 0, \quad \text{for} \quad a_i \in F.$$

Form the vector $v = (v_i)$, where

$$v_i = \begin{cases} 0 & \text{if } i \notin \{j_1, \ldots, j_p\} \\ a_i & \text{if } i \in \{j_1, \ldots, j_p\} \end{cases}$$

The equation above implies that $Av = 0$.

It follows that v is in the solution space of the system $Ax = 0$. Since $Rx = 0$ has the same solution space, $Rv = 0$. Let c_i' denote the i-th column of R. Then

$$0 = Rv = a_{j_1} c_{j_1}' + \cdots + a_{j_p} c_{j_p}.$$

However, since R is in row reduced echelon form, c_{j_k}' is a vector with 1 in the k-th row and 0's elsewhere. Hence, $\{c_{j_1}, \ldots, c_{j_p}\}$ is independent, and $a_{j_1} = \cdots = a_{j_p} = 0$. □

Example. The matrix

$$A = \begin{bmatrix} 1 & 1 & -2 & 3 \\ 2 & 0 & -1 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

row reduces to

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

The leading coefficients occur in the first three columns. Hence, the first three columns of A

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \right\}$$

are independent. In fact, they form a basis for the column space of A. □

Definition. Let A be a matrix. The **transpose** A^T of A is the matrix obtained by switching the rows and columns of A. That is, the i-j-th element of A^T is the j-i-th element of A.

Example.

$$\begin{bmatrix} 1 & 1 & -2 & 3 \\ 2 & 0 & -1 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}^T = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ -2 & -1 & 1 \\ 3 & 1 & 0 \end{bmatrix}. \ \square$$

Proposition. Let A be a matrix. Then

$$\text{row rank } A = \text{column rank } A.$$

Proof. Let R be the row reduced echelon matrix which is row equivalent to A. Suppose the leading coefficients of R occur in columns $j_1, \ldots, j_p$, where $j_1 < \cdots < j_p$, and let c_i denote the i-th column of A. By the preceding lemma, $\{c_{j_1}, \ldots, c_{j_p}\}$ is independent. There is one vector in this set for each leading coefficient, and the number of leading coefficients equals the row rank. Therefore,

$$\text{row rank } A \le \text{column rank } A.$$

Now consider A^T. This is A with the rows and columns swapped, so

$$\text{row rank } A^T = \text{column rank } A,$$

$$\text{column rank } A^T = \text{row rank } A.$$

Applying the first part of the proof to A^T,

$$\text{column rank } A = \text{row rank } A^T \le \text{column rank } A^T = \text{row rank } A.$$

Therefore,

$$\text{column rank } A = \text{row rank } A. \ \square$$

Remark. The proof provides an algorithm for finding a basis for the column space of a matrix. Specifically, row reduce the matrix A to a row reduced echelon matrix R. If the leading coefficients of R occur in columns $j_1, \ldots, j_p$, then consider the columns $c_{j_1}, \ldots, c_{j_p}$ of A. These columns form a basis for the column space of A. □

Example. If A and B are row equivalent, they don't necessarily have the same column space. For example,

$$\begin{bmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix} \quad \to \quad \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

by $r_2 \to r_2 - r_1$. However, all the elements of the column space of the second matrix have their second component equal to 0; this is obviously not true of elements of the column space of the first matrix. □

Proposition. Let A, B, P and Q be matrices, where P and Q are invertible. Suppose $A = PBQ$. Then

$$\operatorname{rank} A = \operatorname{rank} B.$$

Proof. I showed earlier that $\operatorname{rank} MN \le \operatorname{rank} N$. This was row rank; a similar proof shows that

$$\text{column rank } MN \le \text{column rank } M.$$

Since row rank and column rank are the same, $\operatorname{rank} MN \le \operatorname{rank} M$.

Now

$$\operatorname{rank} A = \operatorname{rank} PBQ \le \operatorname{rank} BQ = \text{column rank } BQ \le \text{column rank } B = \operatorname{rank} B.$$

But $B = P^{-1}AQ^{-1}$, so repeating the computation gives $\operatorname{rank} B \le \operatorname{rank} A$. Therefore, $\operatorname{rank} A = \operatorname{rank} B$.
□

Of course, the relation $A = PBQ$ for P and Q invertible is the relation that obtains between two matrices for a linear transformation $f : V \to W$ when you change bases on V and W:

$$[f]_{\mathcal{B},\mathcal{C}} = [\mathcal{B}' \to \mathcal{B}][f]_{\mathcal{B}',\mathcal{C}'}[\mathcal{C} \to \mathcal{C}'].$$

Therefore, the previous result shows that *any two matrices for a linear transformation have the same rank.* The following definition now makes sense.

Definition. Let $f : V \to W$ be a linear transformation. The **rank** of f is the rank of any matrix which represents f.

The Return of the First Isomorphism Theorem.

Definition. Let $f : V \to W$ be a linear transformation. The **kernel** or **null space** of f is the kernel of f, regarded as a group map $(V, +) \to (W, +)$. As usual, the kernel is denoted by $\ker f$.

Remark. f is injective if and only if $\ker f = \{0\}$. This is true for maps of abelian groups; linear transformations are map of abelian groups, and the definition of the kernel makes no reference to the scalar multiplication. □

Lemma. Let $f : V \to W$ be a linear transformation. $\ker f$ is a subspace of V.

Proof. $\ker f \ne \emptyset$, since $f(0) = 0$ shows $0 \in \ker f$.

Let $u, v \in \ker f$, so $f(u) = f(v) = 0$, and let $a \in F$. Then

$$f(au + v) = af(u) + f(v) = a \cdot 0 + 0 = 0.$$

Therefore, $au + v \in \ker f$, and $\ker f$ is a subspace of V. □

Definition. Let $f : V \to W$ be a linear transformation. The **nullity** of f is the dimension of the null space of f.

Lemma. Let W be a subspace of the finite-dimensional vector space V. Then $\dim W \le \dim V$, and $\dim V/W = \dim V - \dim W$.

Proof. If $\{v_i\}$ is a basis of W, it is independent in V and may be enlarged to a basis $\{v_i\} \cup \{w_j\}$ of V. Now $|\{v_i\} \cup \{w_j\}| = \dim V$, so $\dim W \le \dim V$.

Now let $\{v_1, \ldots, v_n\}$ be a basis of W. Enlarge it to a basis $\{v_1, \ldots, v_n, v_{n+1}, \ldots, v_m\}$ for V.

Let $v + W \in V/W$. Write

$$v = \sum_{i=1}^{m} a_i v_i = \sum_{i=1}^{n} a_i v_i + \sum_{i=n+1}^{m} a_i v_i.$$

Then

$$v + W = \pi(v) = \sum_{i=1}^{n} a_i \pi(v_i) + \sum_{i=n+1}^{m} a_i \pi(v_i) = \sum_{i=n+1}^{m} a_i (v_i + W).$$

Since $v+W$ was arbitrary, it follows that the set $\{v_{n+1}+W, \ldots, v_m+W\}$ spans V/W. Hence, $\dim V/W \le m - n$.

In fact, $\{v_{n+1} + W, \ldots, v_m + W\}$ is independent. For suppose

$$c_{n+1}(v_{n+1} + W) + \cdots + c_m(v_m + W) = 0 \in V/W,$$

where $c_{n+1}, \ldots, c_m \in F$.

Then

$$(c_{n+1} v_{n+1} + \cdots + c_m v_m) + W = 0.$$

It follows that $c_{n+1} v_{n+1} + \cdots + c_m v_m \in W$, so

$$c_{n+1} v_{n+1} + \cdots + c_m v_m = c_1 v_1 + \cdots + c_n v_n,$$

for some $c_1, \ldots, c_n \in F$.

Hence,

$$c_1 v_1 + \cdots + c_n v_n - c_{n+1} v_{n+1} - \cdots - c_m v_m = 0,$$

and $c_i = 0$ for all i by independence.

Therefore, $\dim V/W = m - n$. $\square$

Remark. The proof identifies a subspace of V complementary to W, in the sense that it is isomorphic to V/W. In the notation of the proof, it is the subspace $\langle v_{n+1}, \ldots, v_m \rangle$. $\square$

Theorem. Let $f : V \to W$ be a linear transformation. There is an isomorphism $\tilde{f} : V/\ker\ f \to \mathrm{im}\ f$ such that the following diagram commutes:

$$\begin{tikzcd} V \arrow[d, "\pi"'] \arrow[dr, "f"] & \\ V/\ker\ f \arrow[r, "\tilde{f}"'] & \mathrm{im}\ f \end{tikzcd}$$

Proof. The First Isomorphism Theorem for groups applied to the group homomorphism $f : (V, +) \to (W, +)$ yields a group isomorphism $\tilde{f}$ making the diagram commute. I only need to verify that $\tilde{f}$ commutes with scalar multiplication.

Let $v \in V$ and $a \in F$. Since $\tilde{f}\pi = f$,

$$\tilde{f}(a \cdot (v + W)) = \tilde{f}(av + W) = \tilde{f}\pi(av) = f(av) = a \cdot f(v) = a \cdot \tilde{f}\pi(v) = a \cdot \tilde{f}(v + W). \square$$

Theorem. Let $f : V \to W$ be a linear transformation of finite dimensional vector spaces. Then

$$\dim V = \operatorname{rank} f + \operatorname{nullity} f.$$

Proof. Since $\operatorname{rank} f = \dim \mathrm{im}\, f$,

$$\dim V/\ker f = \dim \mathrm{im}\, f = \operatorname{rank} f,$$
$$\dim V - \dim \ker f = \operatorname{rank} f,$$
$$\dim V = \operatorname{nullity} f + \operatorname{rank} f. \square$$