Lecture 8: The rank-nullity theorem; isomorphisms; matrices vs linear transformations

Travis Schedler

Tue, Oct 4, 2011 (version: Wed, Oct 19, 8:37 PM)

Goals (2)
• Prove the rank-nullity theorem
• Isomorphisms
• Isomorphism between matrices and linear transformations
• Preview: Gaussian Elimination

Reading for next time: Finish Chapter 3, start reading the Gaussian Elimination handout (on website).

Warm-up exercise (3)
Use the rank-nullity theorem to reprove the difficult previous warm-up problem differently:
Prove that the dimension of \(U = \{ f \in P_n(F) : f(1) = 0, f'(2) = 0 \} \subseteq P_n(F)\) is \(n - 1\).

(a) Prove that the map \(T : P_n(F) \to F^2, \ f \mapsto (f(1), f'(2))\) is linear.
(b) Show that \(\text{null}(T) = U\).
(c) Prove that \(T\) is surjective (i.e., that \(\text{rk}(T) = 2\)).
(d) Using the rank-nullity theorem, prove that \(\dim \text{null}(T) = n - 1\).

Solution to warm-up exercise (4)

(a) \(T(f) + T(g) = (f(1) + g(1), f'(2) + g'(2)) = T(f + g)\). Similarly \(T(af) = (af(1), af'(2)) = aT(f)\).

(b) \(f \in \text{null}(T)\) exactly means \((f(1), f'(2)) = (0, 0)\), i.e., \(f \in U\).
(c) For every \((a, b) \in \mathbf{F}\), there is a polynomial \(f\) such that \((f(1), f'(2)) = (a, b)\): the degree-one such polynomial is \(f(x) = bx + (a - b)\).

(d) By the rank-nullity theorem, \(n + 1 = \dim P_1(\mathbf{F}) = \dim \text{null}(T) + \text{rk}(T) = \dim U + 2\), so \(\dim U = n - 1\).

**Rank-nullity theorem (5)**

**Theorem 1** (Theorem 3.4; “Rank-nullity theorem”). If \(V\) is f.d. and \(T : V \to W\) is linear, then
\[
\dim V = \dim \text{null}(T) + \dim \text{range}(T).
\]

Definition: the **rank** of \(T\) is \(\text{rk} T := \dim \text{range}(T)\).

**Corollary 2.** \(\text{rk} T \leq \min\{\dim V, \dim W\}\). Moreover, \(\text{rk} T = \dim V\) if and only if \(T\) is injective, and \(\text{rk} T = \dim W\) if and only if \(T\) is surjective.

**Corollary 3** (Corollaries 3.5 and 3.6). Let \(T \in \mathcal{L}(V, W)\) with \(V\) and \(W\) finite-dimensional.

(i) If \(\dim V > \dim W\), then \(T\) is not injective.

(ii) If \(\dim V < \dim W\), then \(T\) is not surjective.

**Proof of rank-nullity theorem (on board) (6)**

Recall: \(T : V \to W\) is linear, and \(V\) is f.d.

- Let \((u_1, \ldots, u_m)\) be a basis of \(\text{null} T\).
- Extend \((u_1, \ldots, u_m)\) to a basis \((u_1, \ldots, u_m, v_1, \ldots, v_n)\) of \(V\).
- Then, \(\dim V = m + n = \dim \text{null} T + n\). We need: \(\dim \text{range} T = n\).
- Claim: \((T(v_1), \ldots, T(v_n))\) is a basis for \(\text{range} T\). This proves the theorem.

**Proof of Claim (on board) (7)**

Claim: \((T(v_1), \ldots, T(v_n))\) is a basis for \(\text{range} T\).

- Spanning: First, \(\text{Span}(T(u_1), \ldots, T(u_m), T(v_1), \ldots, T(v_n)) = T(\text{Span}(u_1, \ldots, u_m, v_1, \ldots, v_n)) = T(V)\), by linearity of \(T\).
- But, \(T(u_1) = \cdots = T(u_m) = 0\) since \(u_1, \ldots, u_m \in \text{null}(T)\).
- Hence, \((T(v_1), \ldots, T(v_n))\) span \(T(V) = \text{range}(T)\).
- Lin. ind.: Suppose \(a_1 T(v_1) + \cdots + a_n T(v_n) = 0\).
- Then, \(T(a_1 v_1 + \cdots + a_n v_n) = 0\).
- So \(a_1 v_1 + \cdots + a_n v_n \in \text{null} T\). We can therefore write
  \[a_1 v_1 + \cdots + a_n v_n = b_1 u_1 + \cdots + b_m u_m.\]
- By linear independence of \((u_1, \ldots, u_m, v_1, \ldots, v_n)\), this implies that \(a_1 = \cdots = a_n = 0 = b_1 = \cdots = b_m\). So, \((T(v_1), \ldots, T(v_n))\) is lin. ind.
Isomorphism and equivalence (8)

Definition 4. A function is called \textit{bijective} if it is both injective and surjective.

A function \( f : X \to Y \) is bijective if and only if there is an inverse function \( f^{-1} : Y \to X \), such that \( f^{-1} \circ f = \text{Id}_X \) and \( f \circ f^{-1} = \text{Id}_Y \). In this case, we say \( X \cong Y \).

In fact, the same fact holds for linear transformations:

Proposition 0.1. Let \( T : V \to W \) be a linear transformation. The following are equivalent:

(a) \( T \) is invertible: i.e., it admits a linear inverse \( T^{-1} : W \to V \);

(b) \( T \) is bijective;

(c) null \( T = 0 \) and range \( T = W \).

Definition 5. \( T \) is called an \textit{isomorphism} if the equivalent conditions above are satisfied. If such a \( T \) exists, we say that \( V \cong W \).

Examples of isomorphisms (9)

- Rotations or reflections of \( \mathbb{R}^2 \).
- \( \{ \text{functions} \{1, 2, \ldots, n\} \to \mathbb{F} \} \cong \mathbb{F}^n \) by the map \( f \mapsto (f(1), f(2), \ldots, f(n)) \).
- The same holds for functions \( X \to \mathbb{F} \) if \( X \) is any (ordered) set with \( n \) elements.
- As real vector spaces, \( \mathbb{R}^2 \cong \mathbb{C} \cong \{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mid a, b \in \mathbb{R} \} \).
- More generally, if \( V \) and \( W \) are any two vector spaces over \( \mathbb{F} \) of the same dimension, then \( V \cong W \).
  - Why? Pick bases \( v_1, \ldots, v_n \) of \( V \) and \( w_1, \ldots, w_n \) of \( W \). Then we have the isomorphism \( v_i \mapsto w_i \), which extends linearly to \( \lambda_1 v_1 + \cdots + \lambda_n v_n \mapsto \lambda_1 w_1 + \cdots + \lambda_n w_n, \forall \lambda_i \in \mathbb{F} \).
- Let \( \text{Mat}_{m,n}(\mathbb{F}) := \) the space of \( m \) by \( n \) matrices. Then \( \text{Mat}_{m,n}(\mathbb{F}) \cong \mathbb{F}^{mn} \) by entries of the matrix.

Properties of linear maps (10)

- We have a zero linear map: \( 0 : V \to W \), such that \( 0(v) = 0 \) for all \( v \in V \).
- Given \( S, T : V \to W \) linear, then \( S + T \) is linear.
- Given \( T : V \to W \) linear, and \( \lambda \in \mathbb{F} \), then \( \lambda T \) is linear.

We deduce: \( \mathcal{L}(V, W) \) is a \textit{vector space}. In fact, a subspace of the vector space of all functions \( V \to W \).
More properties (11)

• Given $T : U \to V$ and $S : V \to W$ linear, then $S \circ T : U \to W$ is linear (composition). This satisfies:
  
  • Associative: $T_1, T_2, T_3$ linear, $T_1(T_2T_3) = (T_1T_2)T_3$.
  
  • Distributive: $(S_1 + S_2)T = S_1T + S_2T$ and $S(T_1 + T_2) = ST_1 + ST_2$.
  
  • Has identity: For $T : V \to W$, $TI_V = T = I_WTI$ (identity for $V$, $I_W$ = identity for $W$).
  
  • NOT commutative: $ST \neq TS$ (doesn’t even make sense in general!)

Notice: Matrices satisfy the same properties!  (NOT a coincidence as we will see.)

When $V = W$: $L(V, V)$ has an associative, distributive multiplication operation (composition) that has a multiplicative identity. (Such a vector space is called an algebra.)

Example: Rotations (12)

• For every angle $\theta$, let $R_\theta \in L(\mathbb{R}^2, \mathbb{R}^2)$ be the counter-clockwise rotation about the origin by angle $\theta$.
  
  • Associativity says: if $\theta, \phi, \psi$ are three angles, then $R_\theta(R_\phi R_\psi) = (R_\theta R_\phi)R_\psi$.
  
  • Indeed, they are all equal to $R_{\theta+\phi+\psi}$!

Matrices (13)

Goal: If $V, W$ are f.d., $L(V, W) \cong \text{Mat}(\dim W, \dim V, \mathbb{F})$.

Note: For $V = \mathbb{F}$, deduce $W \cong L(\mathbb{F}, W) \cong \text{Mat}(\dim W, 1, \mathbb{F}) \cong \mathbb{F}^{\dim W}$!

Given bases $(v_1, \ldots, v_n)$ for $V$ and $(w_1, \ldots, w_m)$ for $W$, define

$$\mathcal{M} : L(V, W) \to \text{Mat}(\dim W, \dim V, \mathbb{F}) = \text{Mat}(m, n, \mathbb{F})$$

as follows: Let $T(v_j) = a_{1j}w_1 + a_{2j}w_2 + \cdots + a_{mj}w_m$. Then

$$T \mapsto \mathcal{M}(T) := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Theorem 6 (PS4 problems 2 and 3 (b)). $\mathcal{M}$ is an isomorphism. Moreover, $\mathcal{M}(S \circ T) = \mathcal{M}(S)\mathcal{M}(T)$!
Nullspace and column space of matrices (14)

What do null $T$ and range $T$ translate into in terms of $M(T)$? Answer:

- $M(\text{null}(T)) = \text{Span}(v : M(T)v = 0) \subseteq \text{Mat}(\text{dim}V, 1, F)$ is the nullspace, a subspace of column space.
- $M(\text{range}(T)) = \text{Span}(M(T)(v)) \subseteq \text{Mat}(\text{dim}W, 1, F)$ is the column space, the span of the columns of $M(T)$. Note: dim colspace($M(T)$) = rk $T$.
- We deduce: the dimensions of nullspace and column space of $M(T)$ depend only on $T$, not on the choice of bases.
- By rank-nullity, the sum of these dims is dim $V$.
- By taking the transpose of the matrix $M(T)$, one can say the same about row space and left null space.
- We will prove: Theorem: dim(rowspace($M(T)$)) = dim(colspace($M(T)$)) = rk($T$).

Preview: Gaussian elimination (15)

We are finally ready to compute null($T$), range($T$), and therefore rk($T$), etc. This will also prove the theorem of last slide.

The primary tool will be Gaussian elimination.

This works by reducing a matrix to row echelon form, a staircase form with zeros below the staircase and 1’s on the corners, e.g.:

$$
\begin{pmatrix}
1 & -2 & 0 & 1 \\
0 & 1 & 3/2 & -1/2 \\
0 & 0 & 0 & 1
\end{pmatrix}
$$

(0.4)

For this matrix, we can read off the column space: the span of the columns meeting a corner of the staircase (with a 1 entry). Here, the first, second, and fourth columns.

We can similarly read off the null space: set the entries of a column not corresponding to a corner (here, only the third entry) arbitrarily, and solve for the other entries.