Notes on Linear Operators

Theorem 1. *Let $T : U \to V$ be a linear mapping. Then U is finite-dimensional iff* $\mathrm{Ker(T)}$ *and* $\mathrm{Im(T)}$ *are finite-dimensional in which case*

$$\dim(U) = \dim(\mathrm{Ker(T)}) + \dim(\mathrm{Im(T)}).$$

Proof. ($\Rightarrow$) If V is finite-dimensional then so is $\mathrm{Ker(T)}$ since a subspace of a finite-dimensional vector space is also finite-dimensional. Also, if $U = \mathrm{Span(e_1, \ldots, e_n)}$, then $\mathrm{Im(T)} = \mathrm{Span(T(e_1), \ldots, T(e_n))}$ so that the image of T is also finite-dimensional. Now if $f_1, \ldots, f_s$ is a basis for $\mathrm{Ker(T)}$ and we complete $f_1, \ldots, f_s$ to a basis $f_1, \ldots, f_s, f_{s+1}, \ldots f_{s+r}$ of V then we claim that $T(f_{s+1}), \ldots, T(f_{s+r})$ is a basis for $\mathrm{Im(T)}$. Indeed, they span $\mathrm{Im(T)}$ since $T(f_1) = \cdots T(f_s) = 0$, and if $a_1T(f_{s+1}) + \cdots + c_rT(f_{s+r}) = 0$ we have $T(a_1f_{s+1} + \cdots a_rf_{s+r}) = 0$ which implies $a_1f_{s+1} + \cdots a_rf_{s+r} = b_1f_1 + \cdots b_sf_s$. Bringing all terms to the left side we get a dependence relation among $f_1, \ldots, f_{r+s}$. Since $f_i's$ are linearly independent we get $a_1 = a_2 = \cdots = a_s = 0$. This yields $\dim(U) = s + r = \dim(\ker(T)) + \dim(\mathrm{Im(T)})$.

($\Leftarrow$) Now suppose $\mathrm{Ker(T)}$ and $\mathrm{Im(T)}$ are finite-dimensional. Let $f_1, \ldots, f_s$ be a basis for $\mathrm{Ker(T)}$ and let $h_1, \ldots h_r$ be a basis for $\mathrm{Im(T)}$. We have $h_i = T(f_{s+i})$ with $f_{s+1}, \ldots, f_{s+r} \in U$. We claim that $f_1, \ldots, f_{s+r}$ is a basis for U. Indeed, if $u \in U$, then $T(u) = a_{s+1}T(f_{s+1} + \cdots + a_{s+r}T(f_{s+r})$ which implies that $u - a_{s+1}f_{s+1} - \cdots - a_{s+r}f_{s+r} \in \mathrm{Ker(T)}$ and hence that

$$u - a_{s+1}f_{s+1} - \cdots - a_{s+r}f_{s+r} = a_1f_1 + \cdots + s_sf_s$$

which gives $u = a_1f_1 + \cdots a_{s+r}f_{s+r}$ and hence that $f_1, \ldots f_{s+r}$ generate U. To show linear independence of these vectors suppose that $a_1f_1 + \cdots a_{s+r}f_{s+r} = 0$. Applying T to both sides yields $a_{s_1}h_1 + \cdots a_{s+r}h_r = 0$ which gives $a_{s+1} = \cdots = a_{s+r} = 0$ since $h_1, \ldots h_r$ are linearly independent. But then $a_1f_1 + \cdots a_sf_s = 0$ which gives $a_1 = \cdots = a_s = 0$ by the fact that f_1, $dots, f_s$ are linearly independent.Thus $\dim(U) = r + s = \dim(\mathrm{Ker(T)}) + \dim(\mathrm{Im(T)})$. $\square$

Corollary 2. *Let $T : U \to V$,$S : V \to W$ be linear mappings such that* $\mathrm{Ker(S)}$ *and* $\mathrm{Ker(T)}$ *are finite-dimensional. Then*

$$\dim \mathrm{Ker(ST)} = \dim \mathrm{Ker(T)} + \dim(\mathrm{Ker(S)} \cap \mathrm{Im(T)}).$$

Proof. We first note that $\mathrm{Ker(T)} \subseteq \mathrm{ker(ST)}$ and that

$$u \in \ker(ST) \iff ST(u) = 0 \iff T(u) \in \mathrm{Ker(S)} \cap \mathrm{Im(T)}.$$

Now let $T_0 : \mathrm{Ker(ST)} \to \mathrm{W}$ be the linear mapping defined by restriction of T to $\mathrm{Ker(ST)}$. Then $\mathrm{Ker(T_0)} = \mathrm{Ker(T)}$ and $\mathrm{Im(T_0)} = \mathrm{Ker(S)} \cap \mathrm{im(T)}$ which yields the result. $\square$

Corollary 3. $\dim(\ker(ST)) \leq \dim(\ker(S)) + \dim(\ker(T))$ *with equality if* $\mathrm{Ker(S)} \subseteq \mathrm{Im(T)}$*, in particular, if T is surjective.*

Theorem 4. *Let T be a linear operator on a vector space V and let $a_1, a_2, \ldots, a_k$ be distinct scalars such that* $\dim(\mathrm{Ker(T - a_i)^{n_i}})$ *is finite-dimensional for $1 \leq i \leq k$. Then*

$$\mathrm{Ker((T - a_1)^{n_1}(T - a_2)^{n_2} \cdots (T - a_k)^{n_k})} = \mathrm{Ker(T - a_1)^{n_1}} + \mathrm{Ker(T - a_2)^{n_2}} + \cdots + \mathrm{Ker(T - a_k)^{n_k}}.$$

Proof. Let $W = \mathrm{Ker((T - a_1)^{n_1}(T - a_2)^{n_2} \cdots (T - a_k)^{n}_{k})}$ and let $W_i = \mathrm{Ker(T - a_i)^{n_i}}$. We first prove that $\dim(W_1 + \cdots + W_k) = \dim(W_1) + \cdots \dim(W_k)$. For this is suffices to prove that if $w_i \in W_k$ with $w_1 + w_2 + \cdots w_k = 0$ then $w_1 = w_2 = \cdots w_k = 0$. Let S_i be the product of the operators

$(T-a_j)^{n_j}$ with $j \neq i$. Then, applying S_i to both sides of $w_1 + \cdots w_k = 0$, we get $S_i(w_i) = 0$ since $S_i(w_j) = 0$ for $j \neq i$. Since $S_i(W_i) \subseteq W_i$, the restriction of S_i to W_i is a linear operator on W_i and $S_i = \prod_{j\neq i}(T_j - a_j)^{n_j}$, where T_j is the restriction of T to W_j. That $w_i = 0$ follows from the following Lemma since

Lemma 5. *If S is a linear operator on a vector space W and a is a scalar such that $(S-a)^n = 0$ then $S-b$ is invertible for every $b \neq a$.*

Proof. We first prove this in the case $a = 0$, $b = 1$. Then $S^n = 0$ so that

$$(1-S)(1+S+S^2+\cdots+S^{n-1} = 1+S+S^2+\cdots S^{n-1} - S - S^2 - \cdots - S^{k-1} = 1.$$

Since two polynomials in S commute we get that $1-S$ is invertible with inverse $1+S+S^2+\ldots+S^{k-1}$. Since $S-1 = -(S-1)$ we see that $S-1$ is also invertible. The general case follows from the identity $S-b = (a-b)(1-(b-a)^{-1}(S-a))$. □

Returning to the proof of Theorem 4, we have $Z = W_1 + W_2 + \cdots + W_k \subseteq W$ since $W_i \subseteq W$ which implies that $\dim Z \leq \dim W$. But, by Corollary 3,

$$\dim W \leq \dim W_1 + \dim W_2 + \cdots + \dim W_k = \dim Z$$

so that $\dim Z = \dim W$ and hence that $Z = W$. □

Corollary 6. *If T is a linear operator on a finite-dimensional vector space, then T is diagonalizable if and only if there are distinct scalars $a_1, a_2, \ldots a_k$ such that $(T-a_1)(T-a_2)\cdots(T-a_k) = 0$.*

We now apply these results to the case of the differential operator D and the left-shift operator L. Since $(D-a)(x^{i+1}e^{ax}) = x^i e^{ax}$, we see that $\mathrm{Span(e^{ax}, xe^{ax}, \ldots, x^{k-1}e^{ax}) \subseteq Ker(D-a)^k}$. But the functions $e^{ax}, xe^{ax}, \ldots, x^{k-1}e^{ax}$ are linearly independent and so, since $\dim\ker(D-a)^k \leq k$, they are a basis for $\ker(D-a)^k$. Hence, for example, $\mathrm{Ker(D-1)(D-2)^2Ker(D-3)^3}$ is 5-dimensional with basis $e^x, e^{2x}, xe^{2x}, e^3x, xe^{3x}, x^2e^{3x}$.

In the case of the left-shift operator L we have, in the case $a \neq 0$,

$$(L-a)(n^{i+1}a^n) \in \mathrm{Span((a^n), (na^n), \cdots, (n^ia^n))}$$

so that $\mathrm{Span((a^n), (na^n), \ldots, (n^{k-1}a^n) \subseteq Ker(L-a)^k}$. But these sequences are linearly independent and so are a basis of $\mathrm{Ker(L-a)^k}$ since $\mathrm{\dim Ker(L-a)^k \leq k}$. Hence, for example,

$$\mathrm{\dim(Ker(L-1)(L-2)^2(L-3)^3 = 5}$$

with basis $(1), (2^n), (n2^n), (3^n), (n3^n), (n^23^n)$.

As another example, consider the problem of finding a formula for $s_n = 1^3 + 2^3 + \cdots + n^3$. Let $s = (s_n)$. Then $(L-1)(s) = ((n+1)^3 - n^3) = 3(n^2) + 3(n) + (1)$ which is in the kernel of $(L-1)^3$. Hence s is in the kernel of $(L-1)^4$ so that there are constants A, B, C, D such that $s_n = A + Bn + Cn^2 + Dn^3$. The constants A, B, C, D can be found by solving the system of equations obtained by setting $n = 0, 1, 2, 3$.

Particular solutions to non-homogeneous difference and recurrence equations can be sometimes found by transforming the non-homgeneous equation into a homogeneous one. For example, the equation $(L-1)(L-2)x = (1)$ can be transformed into a homogeneous one by applying the operator $L-1$ to both sides of the equation obtaining $(L-1)^2(L-2)x = 0$. The operator $L-1$ was chosen to kill the right-hand side of the equation. Thus $x_n = a + bn + c2^n$. The term $a + c2^n$ can be omitted as it is in the kernel of $(L-1)(L-2)$. We thus look for a particular solution of the form $x_n = bn$. The constant b can be found by substitution in the given non-homogeneous equation. A similar procedure applies to differential equations.