UIC Mtht 435 Class notes

Linear equations and linear maps

Let $A = (a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$, an $m \times n$ matrix. The entry a_{ij} is in the ith row and jth column for $1 \leq i \leq m$ and $1 \leq j \leq n$. Let $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, an $n \times 1$ matrix or a vector written vertically. The matrix product Ax is the $m \times 1$ matrix $\begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}$ with only one rather wide column. Example:

$$\begin{pmatrix} 5 & -2 & 3 \\ -4 & 3 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5x_1 - 2x_2 + 3x_3 \\ -4x_1 + 3x_2 + 2x_3 \end{pmatrix}$$

If $b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, the equation $Ax = b$ is the system of equations:

$$\begin{aligned} 5x_1 - 2x_2 + 3x_3 &= 1 \\ -4x_1 + 3x_2 + 2x_3 &= 2. \end{aligned}$$

The rule for matrix multiplication gives the two properties: $A(x + y) = Ax + Ay$ and $A(rx) = rAx$ for $r \in R$. Define $\phi : R^n \longrightarrow R^m$ by $\phi(x) = Ax$ where we think of R^n as the n-dimensional space of column vectors. ϕ is a linear map and $\phi(x) = b$ is a system of m equations in n unknowns.

Questions: If $b \in R^m$, does $b \in \operatorname{im} \phi$? If $b \in \operatorname{im} \phi$, which $x \in R^n$ have $\phi(x) = b$?

Let $\phi : V \longrightarrow W$ be a linear map. Recall $\operatorname{im} \phi = \{\phi(v) : v \in V\} \subset W$ and $\ker \phi = \{v \in V : \phi(v) = 0_W\} \subset V$. The kernel and the image are both subspaces. ϕ is *onto* if and only if $\operatorname{im} \phi = W$.

Proposition. ϕ is one-to-one $\iff$ $\ker \phi = \{0_V\}$.

Proof. $\Rightarrow$: If $v \in \ker \phi$, then $\phi(v) = 0_W$. Also $\phi(0_V) = 0_W$. If ϕ is one-to-one this implies $v = 0_V$, so $\ker \phi = \{0_V\}$.

$\Leftarrow$: If $\phi(u) = \phi(v)$ then $\phi(u - v) = 0_W$ so $u - v \in \ker \phi$. If $\ker \phi = \{0_V\}$, then this implies $u - v = 0_V$ so $u = v$ and ϕ is one-to-one.

Let $\dim V = n$, $\dim W = m$, and $\dim \ker \phi = k$. Since $\ker \phi \subset V$, by Proposition 1.1 $\ker \phi$ has a basis. Let $v_1, \ldots, v_k$ be a basis for $\ker \phi$. By Proposition 2.1 there are vectors $v_{k+1}, \ldots, v_n$ so that $v_1, \ldots, v_n$ is a basis for V. Let $U = L(v_{k+1}, \ldots, v_n)$; $\dim U = n - k$.

If $w \in \operatorname{im} \phi$ then $w = \phi(v)$ for some $v \in V$. We can write $v = \sum_{i=1}^{n} a_i v_i$ Then $w = \phi(v) = \sum_{i=1}^{n} a_i \phi(v_i)$. But $\phi(v_i) = 0_W$ for $1 \le i \le k$ so $w = \sum_{i=k+1}^{n} a_i \phi(v_i)$. therefore $L(\phi(v_{k+1}), \ldots, \phi(v_n)) = \operatorname{im} \phi$.

Lemma. $\phi(v_{k+1}), \ldots, \phi(v_n)$ is a basis for $\operatorname{im} \phi$; $\dim \operatorname{im} \phi = n - k$.

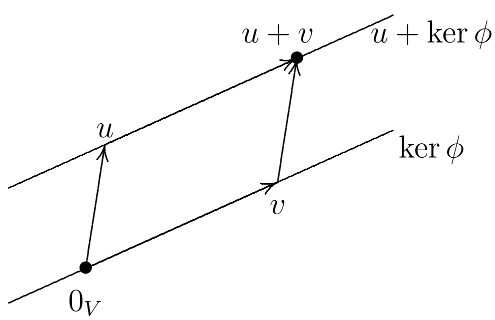
Proof. We have shown these vectors generate $\operatorname{im} \phi$. We need to show they are linearly independent. Suppose $\sum_{i=k+1}^{n} a_i \phi(v_i) = 0_W$. Then $\phi(\sum_{i=k+1}^{n} a_i v_i) = 0_W$, so $\sum_{i=k+1}^{n} a_i v_i \in \ker \phi$. Hence $\sum_{i=k+1}^{n} a_i v_i = \sum_{i=1}^{k} a_i v_i$. Since $v_1, \ldots, v_n$ are independent, all $a_i = 0$. Hence $\phi(v_{k+1}), \ldots, \phi(v_n)$ are independent.

We have shown that $\dim \ker \phi + \dim \operatorname{im} \phi = \dim V$.

If $b \notin \operatorname{im} \phi$, there is no solution $x \in V$ to $\phi(x) = b$. If $b \in \operatorname{im} \phi$ then there is some $u \in V$ with $\phi(u) = b$. Let $v \in \ker \phi$. Then $\phi(u + v) = \phi(u) + \phi(v) = b + 0_W = b$. Conversely, if $\phi(x) = b$ then $\phi(x - u) = b - b = 0_W$, so $x - u \in \ker \phi$ and $x = u + v$ for some $v \in \ker \phi$. Hence if u is a particular solution, $\phi(u) = b$, then the set of all solutions is :

$$\{x : \phi(x) = b\} = \{u + v : v \in \ker \phi\}.$$

This set of solutions is also written $u + \ker \phi$. Think of it as a k-dimensional plane parallel to $\ker \phi$ through the point u.



We return to the example on page 1. To solve this system of two equations in three unknowns, we look for a simpler system with the same solutions. We can get such an equivalent system by multiplying one equation by a nonzero number, adding a multiple of one equation to another equation, or by interchanging two equations. These operations can be done as operations on the *augmented* matrix consisting of A together with the column matrix b.

The augmented matrix in the example is

$$\begin{pmatrix} 5 & -2 & 3 & 1 \\ -4 & 3 & 2 & 2 \end{pmatrix}.$$

Add the second row to the first to get:

$$\begin{pmatrix} 1 & 1 & 5 & 3 \\ -4 & 3 & 2 & 2 \end{pmatrix}.$$

The add 4 times the first row to the second:

$$\begin{pmatrix} 1 & 1 & 5 & 3 \\ 0 & 7 & 22 & 14 \end{pmatrix}.$$

Now multiply the second row by 7^{-1}:

$$\begin{pmatrix} 1 & 1 & 5 & 3 \\ 0 & 1 & 22/7 & 2 \end{pmatrix}$$

and subtract the second row from the first:

$$\begin{pmatrix} 1 & 0 & 13/7 & 1 \\ 0 & 1 & 22/7 & 2 \end{pmatrix}.$$

The new set of equations is

$$\begin{aligned} x_1 \qquad + (13/7)x_3 &= 1 \\ x_2 + (22/7)x_3 &= 2. \end{aligned}$$

Setting the third variable, $x_3 = 0$ now determines values for the other two: $x_1 = 1$ and $x_2 = 2$, so $x = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ is a particular solution.

Replacing the last column with 0s gives the equations for the kernel. Let $x_3 = 7$ Then $x_1 = -13$ and $x_2 = -22$, so the vector $v = \begin{pmatrix} -13 \\ -22 \\ 7 \end{pmatrix}$ is in the kernel. Any choice of x_3 gives exactly one element of $\ker \phi$, so $\ker \phi$ is one-dimensional and this v is a basis for it. The general solution is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + t \begin{pmatrix} -13 \\ -22 \\ 7 \end{pmatrix} \quad \text{for any real } t.$$

The image of ϕ has dimension $\dim V - \dim \ker \phi = 2$, so $\operatorname{im} \phi = W$ and there is a solution for any $b \in W$.

Fields and vector spaces

Let K be a field with a subfield F, $F \subset K$. K is called an *extension* of F. K also has the structure of a vector space over F. In fact, the properties which must hold for K to be a vector space over F hold because K is a field with F as a subfield. Specifically:

$K, +$ is an abelian group,

for $a \in F$ and $v \in K$, $av \in K$ is defined,

the associative law $(ab)v = a(bv)$ holds,

the distributive laws $a(v_1 + v_2) = av_1 + av_2$ and $(a + b)v = av + bv$ hold, and

$1 \in F$ satisfies $1v = v$ in K.

We just forget for the moment that we can multiply any two elements of K.

Definition. The *degree* of the field K over F is the dimension of K as a vector space over F, $[K : F] = \dim_F K$.

Proposition. If $F \subset K \subset L$ are fields and $[L : F]$ is finite, then $[K : F]$ and $[L : K]$ are both finite.

Proof. $K \subset L$ are vector spaces over F, so by Proposition 1.1, since L is finitely generated over F, K has a basis over F and $\dim_F K \le \dim_F L$. For the second statement, let $z_1, \ldots, z_d$ generate L over F. Then any $w \in L$ can be written $w = \sum a_i z_i$ for $a_i \in F$. Since also $a_i \in K$, $z_1, \ldots, z_d$ generate L over K.

Theorem 1. If $F \subset K \subset L$ are fields and $[K : F]$ and $[L : K]$ are both finite, then $[L : F]$ is finite and $[L : F] = [L : K][K : F]$.

Proof. Let $v_1, \ldots, v_m$ be a basis for K over F and let $w_1, \ldots, w_n$ be a basis for L over K. Any $w \in L$ can be written as a sum $w = \sum_{j=1}^n c_j w_j$ where $c_j \in K$. Then each $c_j = \sum_{i=1}^m a_{ij} v_i$ and hence $w = \sum_{j=1}^n \sum_{i=1}^m a_{ij} v_i w_j$.

Hence $\{v_i w_j : 1 \le i \le m, 1 \le j \le n\}$ is a set of generators for L over F.

We must show that this set is linearly independent over F. Suppose that

$$\sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij} v_i w_j = 0_L.$$

Then $\sum_{j=1}^n \left(\sum_{i=1}^m a_{ij} v_i\right) w_j = 0_L$. Since $w_1, \ldots, w_n$ are linearly independent over K each coefficient $\sum_{i=1}^m a_{ij} v_i = 0_K$. Then, since $v_1, \ldots, v_m$ are linearly independent over F, each $a_{ij} = 0_F$. Therefore $\{v_i w_j : 1 \le i \le m, 1 \le j \le n\}$ is a set of vectors in L which is linearly independent over F. It follows that $\dim_F L = mn = \dim_F K \dim_K L$.