USING MATRICES TO COMPUTE BASES OF KERNELS AND IMAGES

Let V and W be finite dimensional vector spaces over the field F and $T\colon V \to W$ a linear transformation. Choosing bases

$$v_1, \dots, v_n \qquad \text{and} \qquad w_1, \dots, w_m$$

of V and W, respectively, T is represented by an $m \times n$ matrix $A = (a_{ij})$ whose entries are determined by the relations

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i \qquad \text{for all } 1 \leq j \leq n.$$

Computing a basis of the kernel. Recall that a vector $x_1 v_1 + x_2 v_2 + \cdots + x_n v_n \in V$ lies in $\mathrm{Ker}(T)$ if and only if it satisfies the equation

$$T(x_1 v_1 + x_2 v_2 + \cdots + x_n v_n) = 0_W,$$

that is, if and only if

$$x_1 T(v_1) + x_2 T(v_2) + \cdots + x_n T(v_n) = 0_W. \tag{1}$$

Since each $T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$, the above equation becomes

$$\sum_{j=1}^{n} x_j \left(\sum_{i=1}^{m} a_{ij} w_i \right) = 0_W.$$

Rearranging things slightly, we obtain

$$\sum_{i=1}^{m} \left(\sum_{j=1}^{n} x_j a_{ij} \right) w_i = 0_W.$$

Because the w_i are linearly independent, the above coefficients must satisfy $\sum_{j=1}^{n} x_j a_{ij} = 0$. In other words, we obtain the following system of linear equations.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned} \tag{2}$$

Note that the augmented matrix of this system is simply $(A\ \mathbf{0})$. Let $x_{j_1}, \dots, x_{j_s}$ denote the free variables in the system (2). Our experience in solving systems of linear equations tells us that any solution to (2) is of the form

$$\mathbf{u} = x_{j_1} \mathbf{u}_1 + \cdots + x_{j_s} \mathbf{u}_s$$

where $\mathbf{u}_1, \ldots, \mathbf{u}_s$ are the vectors in F^n obtained from the reduced echelon form of $(A\ \mathbf{0})$ by collecting like terms involving $x_{j_1}, \ldots, x_{j_s}$. By the way we compute solutions to systems of linear equations, one readily verifies for each $1 \leq k \leq s$ that the j_kth entry of $\mathbf{u}$ is just x_{j_k}. For each $1 \leq k \leq s$, write

$$\mathbf{u}_k = \begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{pmatrix}.$$

Finally, for each $1 \leq k \leq s$, define the vector

$$u_k = \sum_{j=1}^{n} b_{jk} v_j$$

in V.

Proposition 1. *The vectors $u_1, \ldots, u_s$ form a basis of* $\mathrm{Ker}(T)$.

Proof. Note for each $1 \leq k \leq s$ that

$$T(u_k) = T\left(\sum_{j=1}^{n} b_{jk} v_j\right) = \sum_{j=1}^{n} b_{jk} T(v_j) = \sum_{j=1}^{n} b_{jk} \left(\sum_{i=1}^{m} a_{ij} w_i\right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} b_{jk}\right) w_i.$$

Since

$$\mathbf{u}_k = \begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{pmatrix}$$

is a solution to the system (2), all of the coefficients $\sum_{j=1}^{n} a_{ij} b_{jk}$ in the above expression are zero. This shows that $T(u_k) = 0_W$, thus $u_k \in \mathrm{Ker}(T)$.

To show that $u_1, \ldots, u_s$ span $\mathrm{Ker}(T)$, let $v \in \mathrm{Ker}(T)$ and write

$$v = c_1 v_1 + \cdots + c_n v_n \qquad \text{for some } c_1, \ldots, c_n \in F.$$

Then

$$0_W = T(v) = T\left(\sum_{j=1}^{n} c_j v_j\right) = \sum_{j=1}^{n} c_j T(v_j) = \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{m} a_{ij} w_i\right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} c_j\right) w_i.$$

Because the w_i are linearly independent, we must have

$$\sum_{j=1}^{n} a_{ij} c_j = 0 \qquad \text{for all } 1 \leq i \leq m.$$

This means that the vector

$$\mathbf{u} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \in F^n$$

is a solution to the system (2). We saw that any solution of (2) is some linear combination of $\mathbf{u}_1, \ldots, \mathbf{u}_s$, hence there exist scalars $d_1, \ldots, d_s \in F$ such that

$$\mathbf{u} = d_1 \mathbf{u}_1 + \cdots + d_s \mathbf{u}_s.$$

This is equivalent to the statement

$$c_j = d_1 b_{j1} + \cdots + d_s b_{js} \qquad \text{for all } 1 \leq j \leq n.$$

We therefore have

$$v = \sum_{j=1}^{n} c_j v_j = \sum_{j=1}^{n} \left(\sum_{k=1}^{s} d_k b_{jk} \right) v_j = \sum_{k=1}^{s} d_k \left(\sum_{j=1}^{n} b_{jk} v_j \right) = \sum_{k=1}^{s} d_k u_k.$$

This exhibits v as a linear combination of $u_1, \ldots, u_s$.

It remains to show that $u_1, \ldots, u_s$ are linearly independent. Suppose that

$$c_1 u_1 + \cdots + c_s u_s = 0_W.$$

We need to show that the coefficients c_k are all zero. One computes that

$$0_W = \sum_{k=1}^{s} c_k u_k = \sum_{k=1}^{s} c_k \left(\sum_{j=1}^{n} b_{jk} v_j \right) = \sum_{j=1}^{n} \left(\sum_{k=1}^{s} c_k b_{jk} \right) v_j.$$

Since the v_j are linearly independent, we must have

$$\sum_{k=1}^{s} c_k b_{jk} = 0 \qquad \text{for all } 1 \leq j \leq n.$$

This in turn shows that

$$c_1 \mathbf{u}_1 + \cdots + c_s \mathbf{u}_s = \mathbf{0}.$$

By the remarks preceding Proposition 1, for each $1 \leq k \leq s$, the j_kth entry of the left hand side is c_k. The j_kth entry of the right hand side is zero, so $c_k = 0$. □

Computing a basis of the image. For this part we introduce a bit of notation. Write the columns of A as

$$A = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_n \end{pmatrix}.$$

Let $(B\ \mathbf{0})$ be an echelon form of $(A\ \mathbf{0})$ and write the columns of B as

$$B = \begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_n \end{pmatrix}.$$

Recall that the free and basic variables are determined by the pivot columns in $(B\ \mathbf{0})$.

Proposition 2. *The vectors* $\{T(v_j) \mid x_j \text{ is a basic variable}\}$ *form a basis of* $\operatorname{Im}(T)$.

Proof. Let $x_{j_1}, \dots, x_{j_r}$ be the *basic* variables. We will first show that $T(v_{j_1}), \dots, T(v_{j_r})$ span $\operatorname{Im}(T)$. For this it suffices to prove for each *free* variable x_{j_0} that the vector $T(v_{j_0})$ may be expressed as a linear combination of $T(v_{j_1}), \dots, T(v_{j_r})$. Since x_{j_0} is a free variable, we may choose the solution

$$\mathbf{u} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \in F^n$$

to the system (2) (and hence to the equation (1)) obtained by setting $x_{j_0} = -1$ and all other free variables equal to zero. One may check that this forces $c_{j_0} = -1$ and all other c_j with x_j free to be zero. Plugging $\mathbf{u}$ back into (1) yields the equality

$$-T(v_{j_0}) + b_{j_1}T(v_{j_1}) + \cdots + b_{j_r}T(v_{j_r}) = 0_W$$

so that

$$T(v_{j_0}) = b_{j_1}T(v_{j_1}) + \cdots + b_{j_r}T(v_{j_r}).$$

This exhibits $T(v_{j_0})$ as a linear combination of $T(v_{j_1}), \dots, T(v_{j_r})$.

To show that $T(v_{j_1}), \dots, T(v_{j_r})$ are linearly independent, consider the equation

$$x_{j_1}T(v_{j_1}) + \cdots + x_{j_r}T(v_{j_r}) = 0_W. \tag{3}$$

As above, this gives rise to the following system of linear equations.

$$\begin{aligned} a_{1j_1}x_{j_1} + \ a_{1j_2}x_{j_2} + \ \cdots + \ a_{1j_r}x_{j_r} &= 0 \\ a_{2j_1}x_{j_1} + \ a_{2j_2}x_{j_2} + \ \cdots + \ a_{2j_r}x_{j_r} &= 0 \\ &\vdots \\ a_{mj_1}x_{j_1} + a_{mj_2}x_{j_2} + \ \cdots + a_{mj_r}x_{j_r} &= 0 \end{aligned} \tag{4}$$

The augmented matrix of this system is

$$\begin{pmatrix} \mathbf{a}_{j_1} & \mathbf{a}_{j_2} & \dots & \mathbf{a}_{j_r} & \mathbf{0} \end{pmatrix}.$$

The same elementary row operations as were used to obtain $(B\ \mathbf{0})$ from $(A\ \mathbf{0})$ produce the echelon form matrix

$$\begin{pmatrix} \mathbf{b}_{j_1} & \mathbf{b}_{j_2} & \dots & \mathbf{b}_{j_r} & \mathbf{0} \end{pmatrix}.$$

There is a pivot in every column of this matrix except for the last one, thus every variable in (4) is basic. This means that the system (4) has only one solution, namely the zero solution. It follows that the equation (3) has only the zero solution, hence

$$T(v_{j_1}), \dots, T(v_{j_r})$$

are linearly independent. □

Remark. The basis elements we found for $\operatorname{Ker}(T)$ were in one to one correspondence with the free variables in (2), and the basis elements for $\operatorname{Im}(T)$ were in one to one correspondence with the basic variables. There are n total variables, so we have

$$\dim \operatorname{Im}(T) + \dim \operatorname{Ker}(T) = n.$$

In this way, we obtain an alternative proof of the rank-nullity theorem.

A moderate example. Let $T\colon \mathcal{P}_4(\mathbb{Q}) \to \mathcal{P}_2(\mathbb{Q})$ be the linear transformation represented by the matrix

$$A = \begin{pmatrix} 0 & 2 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & -2 & 2 & -4 & 1 \end{pmatrix}$$

with respect to the bases

$$\underbrace{x^4 - x^3}_{v_1}, \underbrace{x^4 + x^2}_{v_2}, \underbrace{x^2 - x}_{v_3}, \underbrace{x^2 + x}_{v_4}, \underbrace{x^2 + 1}_{v_5} \qquad \text{and} \qquad \underbrace{x^2}_{w_1}, \underbrace{x + 1}_{w_2}, \underbrace{x - 1}_{w_3}$$

of $\mathcal{P}_4(\mathbb{Q})$ and $\mathcal{P}_2(\mathbb{Q})$, respectively. (One may check that T is given by

$$\begin{aligned} &T(ax^4 + bx^3 + cx^2 + dx + e) = \\ &= (2a + 2b + d)x^2 + \left(\tfrac{1}{2}a + \tfrac{1}{2}b - \tfrac{3}{2}c - \tfrac{7}{2}d + \tfrac{5}{2}e\right) x + \left(\tfrac{5}{2}a + \tfrac{5}{2}b + \tfrac{1}{2}c + \tfrac{5}{2}d - \tfrac{3}{2}e\right), \end{aligned}$$

but that isn't important for this method.) In order to compute a basis of $\operatorname{Ker}(T)$, the above procedure tells us to reduce the augmented matrix

$$\begin{pmatrix} A & \mathbf{0} \end{pmatrix} = \begin{pmatrix} 0 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & -2 & 2 & -4 & 1 & 0 \end{pmatrix}$$

to its reduced echelon form

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

This is the augmented matrix for the system of linear equations

$$\begin{array}{rcrcrcrcl} x_2 & + & & - & x_4 & & & = & 0 \\ & + & x_3 & - & 3x_4 & & & = & 0 \\ & & & & & & x_5 & = & 0 \end{array}$$

Any solution to this will be of the form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_4 \\ 3x_4 \\ x_4 \\ 0 \end{pmatrix} = x_1 \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\mathbf{u}_1} + x_4 \underbrace{\begin{pmatrix} 0 \\ 1 \\ 3 \\ 1 \\ 0 \end{pmatrix}}_{\mathbf{u}_2}.$$

By the above results, we know that

$$1(\underbrace{x^4 - x^3}_{v_1}) + 0(\underbrace{x^4 + x^2}_{v_2}) + 0(\underbrace{x^2 - x}_{v_3}) + 0(\underbrace{x^2 + x}_{v_4}) + 0(\underbrace{x^2 + 1}_{v_5}) = x^4 - x^3$$

and

$$0(x^4 - x^3) + 1(x^4 + x^2) + 3(x^2 - x) + 1(x^2 + x) + 0(x^2 + 1) = x^4 + 5x^2 - 2x$$

form a basis of $\mathrm{Ker}(T)$.

To compute a basis of $\mathrm{Im}(T)$, we now turn to the basic variables x_2, x_3, x_5. According to the above results,

$$T(\underbrace{x^4 + x^2}_{v_2}) = 2(\underbrace{x^2}_{w_1}) + 1(\underbrace{x + 1}_{w_2}) - 2(\underbrace{x - 1}_{w_3}) = 2x^2 - x + 3,$$

$$T(x^2 - x) = -1(x^2) + 0(x + 1) + 2(x - 1) = -x^2 + 2x - 2,$$

and

$$T(x^2 + 1) = 0(x^2) + 0(x + 1) + 1(x - 1) = x - 1$$

form a basis of $\mathrm{Im}(T)$.

Another example. We now consider the very special case in which $V = F^n$, $W = F^m$ and $T \colon F^n \to F^m$ is given via matrix multiplication by an $m \times n$ matrix A, i.e.,

$$T(\mathbf{v}) = A\mathbf{v}.$$

Recall that A represents T with respect to the standard bases

$$\mathbf{e}_1, \dots, \mathbf{e}_n \qquad \text{and} \qquad \mathbf{f}_1, \dots, \mathbf{f}_m$$

of F^n and F^m, respectively. We also have

$$A = \begin{pmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \dots & T(\mathbf{e}_n) \end{pmatrix}.$$

In this situation, the kernel of T is called the *null space* of A, and the image of T is called the *column space* of A. These are denoted by $\mathrm{Nul}\, A$ and $\mathrm{Col}\, A$, respectively.

Consider the matrix

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

In order to compute a basis of $\operatorname{Nul} A$, the above procedure tells us to reduce the augmented matrix $(A\ \mathbf{0})$ to its reduced echelon form

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

This is the augmented matrix for the system of linear equations

$$\begin{array}{ccccccccc} x_1 & & + & x_3 & & - & x_5 & = & 0 \\ & x_2 & + & 2x_3 & & + & x_5 & = & 0 \\ & & & & x_4 & - & x_5 & = & 0 \end{array}$$

Any solution to this will be of the form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -x_3 + x_5 \\ -2x_3 - x_5 \\ x_3 \\ x_5 \\ x_5 \end{pmatrix} = x_3 \underbrace{\begin{pmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}}_{\mathbf{u}_1} + x_5 \underbrace{\begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix}}_{\mathbf{u}_2}.$$

By the above results, we know that

$$-1\mathbf{e}_1 - 2\mathbf{e}_2 + 1\mathbf{e}_3 + 0\mathbf{e}_4 + 0\mathbf{e}_5 \qquad \text{and} \qquad 1\mathbf{e}_1 - 1\mathbf{e}_2 + 0\mathbf{e}_3 + 1\mathbf{e}_4 + 1\mathbf{e}_5$$

form a basis of $\operatorname{Nul} A$. But these are just the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$. In other words,

$$\mathbf{u}_1 = \begin{pmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

form a basis of $\operatorname{Nul} A$.

To compute a basis of $\operatorname{Col} A$, we now turn to the basic variables x_1, x_2, x_4. According to the above results, $T(\mathbf{e}_1), T(\mathbf{e}_2), T(\mathbf{e}_4)$ form a basis of $\operatorname{Col} A$. But these are just

$$\mathbf{a}_1 = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{a}_4 = \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix},$$

the columns of A that correspond to the free variables x_1, x_2, x_4.

Application: Finding a basis for the span of vectors. Let

$$\mathcal{S} = \{v_1, \dots, v_n\}$$

be a finite collection of vectors in a vector space V. We learned earlier in the course how one is able to remove vectors from $\mathcal{S}$, one by one, until they obtain a basis $\mathcal{B} \subseteq \mathcal{S}$ of $\operatorname{Span} \mathcal{S}$. The procedure given at that time was very clunky, however, requiring one to solve systems of equations over and over again. The above theory provides a shortcut.

Define $T\colon F^n \to V$ to be the linear transformation given by

$$T(\mathbf{e}_j) = v_j \qquad \text{for all } 1 \leq j \leq n.$$

The idea in doing this is that we will have $\operatorname{Im}(T) = \operatorname{Span} \mathcal{S}$ by construction. If V is finite dimensional, say of dimension m, then choosing a basis $w_1, \dots, w_m$ of V, we may represent T by an $m \times n$ matrix A with respect to the bases

$$\mathbf{e}_1, \dots, \mathbf{e}_n \qquad \text{and} \qquad w_1, \dots, w_m$$

of F^n and V, respectively. The above theory tells us that the collection

$$\mathcal{B} = \{T(\mathbf{e}_j) \mid x_j \text{ is a basic variable in } A\mathbf{x} = \mathbf{0}\}$$

will be a basis of $\operatorname{Im}(T)$. In practice, choosing the basis of V and finding the matrix A will usually be straightforward.

Example. Consider the collection of vectors

$$\mathcal{S} = \left\{x^3 - x,\ -2x^3 + 2x,\ -x^3 + 2x^2 + x + 1,\ 2x^2 + 1\right\}$$

in $\mathcal{P}_3(\mathbb{Q})$. To find a basis $\mathcal{B} \subseteq \mathcal{S}$ of $\operatorname{Span} \mathcal{S}$, we use the linear transformation

$$T\colon \mathbb{Q}^4 \longrightarrow \mathcal{P}_3(\mathbb{Q})$$

defined by the assignments

$$\mathbf{e}_1 \longmapsto x^3 - x, \quad \mathbf{e}_2 \longmapsto -2x^3 + 2x, \quad \mathbf{e}_3 \longmapsto -x^3 + 2x^2 + x + 1, \quad \mathbf{e}_4 \longmapsto 2x^2 + 1$$

so that $\operatorname{Im}(T) = \operatorname{Span} \mathcal{S}$. The matrix A that represents T with respect to the bases

$$\mathbf{e}_1,\ \mathbf{e}_2,\ \mathbf{e}_3,\ \mathbf{e}_4 \qquad \text{and} \qquad x^3,\ x^2,\ x,\ 1$$

is

$$A = \begin{pmatrix} 1 & -2 & -1 & 0 \\ 0 & 0 & 2 & 2 \\ -1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

The matrix

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

is an echelon form of $(A\ \mathbf{0})$. It is apparent from this that the basic variables in the system of linear equations $A\mathbf{x} = \mathbf{0}$ are x_1 and x_3. It follows that

$$T(\mathbf{e}_1) = x^3 - x \qquad \text{and} \qquad T(\mathbf{e}_3) = -x^3 + 2x^2 + x + 1$$

form a basis of $\operatorname{Im}(T) = \operatorname{Span}\mathcal{S}$.

A special case. If one is given a collection of column vectors

$$\mathcal{S} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$$

in F^m, the above procedure tells us use the linear transformation $T \colon F^n \to F^m$ defined by

$$T(\mathbf{e}_j) = \mathbf{v}_j \qquad \text{for all } 1 \le j \le n.$$

But in this case, the matrix A that represents T with respect to the standard bases

$$\mathbf{e}_1, \ldots, \mathbf{e}_n \qquad \text{and} \qquad \mathbf{f}_1, \ldots, \mathbf{f}_m$$

is none other than

$$A = \begin{pmatrix} \mathbf{v}_1 & \ldots & \mathbf{v}_n \end{pmatrix}.$$

The procedure in this special case is therefore more streamlined.

- Form the matrix $A = (\mathbf{v}_1\ \ldots\ \mathbf{v}_n)$.
- Row reduce the augmented matrix $(A\ \mathbf{0})$ to an echelon form $(B\ \mathbf{0})$ (or equivalently, row reduce A to B).
- The columns of A that correspond to pivot columns of B form a basis of $\operatorname{Span}\mathcal{S}$.