Vector Spaces

A *vector space* is a set of elements V and a set of scalar elements along with two operations, addition and scalar multiplication, that satisfy the following conditions:

- $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ for all elements $\mathbf{u}$, $\mathbf{v}$ in V
- $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ for all elements $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V.
- There is a $\mathbf{0}$ element that satisfies $\mathbf{u} + \mathbf{0} = \mathbf{u}$ for all $\mathbf{u}$ in V.
- For each $\mathbf{u}$ in V there is an element $-\mathbf{u}$ that satisfies $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
- $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\, \mathbf{u} + \alpha\, \mathbf{v}$ for all scalars α and elements $\mathbf{u}$, $\mathbf{v}$ in V.
- $(\alpha + \beta)\, \mathbf{u} = \alpha\, \mathbf{u} + \beta\, \mathbf{u}$ for all scalars α and β and all elements $\mathbf{u}$ in V.
- $\alpha\, (\beta\, \mathbf{u}) = (\alpha\, \beta)\, \mathbf{u}$ for all scalars α and β and all elements $\mathbf{u}$ in V.
- $1\, \mathbf{u} = \mathbf{u}$ for all $\mathbf{u}$ in V.

Subspaces

A *subspace* U of a vector space V is a subset of V containing the $\mathbf{0}$ vector that is closed under the operations of vector addition and scalar multiplication.

Linear Operators

A function f that maps a vector space V to a vector space W is a *linear operator* if for all $\mathbf{u}$ and $\mathbf{v}$ in V and all scalars α and β we have

$$f(\alpha\, \mathbf{u} + \beta\, \mathbf{v}) = \alpha\, f(\mathbf{u}) + \beta\, f(\mathbf{v})$$

Understanding the Action of Linear Operators

A key aspect of linear algebra is understanding what effect a linear operator $f{:}V \to U$ has on vectors in V. One of the first questions to ask about a linear operator f is what its *null space* is. If $f{:}V \to U$ is a linear operator on V, the set $N(f)$ is the subset of all vectors $\mathbf{v}$ in V for which $f(\mathbf{v}) = \mathbf{0}$.

The first important fact about the set $N(f)$ is that it is a subspace of V:

- $f(\mathbf{0}) = \mathbf{0}$ for all linear operators (why?), so $\mathbf{0}$ is always in $N(f)$.
- If $\mathbf{u}$ and $\mathbf{v}$ are in $N(f)$ and α and β are any two scalars, $f(\alpha\, \mathbf{u} + \beta\, \mathbf{v}) = \alpha\, f(\mathbf{u}) + \beta\, f(\mathbf{v}) = \mathbf{0}$, so $N(f)$ is closed under addition and scalar multiples.

An important fact about null spaces is that solutions of $f(\mathbf{v}) = \mathbf{u}$ are unique if and only if $N(f) = \{\mathbf{0}\}$. This tells us that uniqueness questions concerning linear operator equations $f(\mathbf{v}) = \mathbf{u}$ can be addressed by trying to understand the null space of f.

What about the existence of solutions to linear operator equations $f(\mathbf{v}) = \mathbf{u}$? It turns out that even here null spaces have something useful to tell us. Here is a result that applies to the special case of an operator that maps the vector space $\mathbb{R}^n$ to $\mathbb{R}^n$. Such operators can be represented as matrix multiplications.

Theorem (The Fredholm Alternative) Suppose A is an n by n matrix with real entries. The mapping $f(\mathbf{x}) = A\,\mathbf{x}$ is a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$. Exactly one of the following is true:

1. The null space of $f(\mathbf{x}) = A\,\mathbf{x}$ is trivial, and for all $\mathbf{b}$ in $\mathbb{R}^n$ the equation $f(\mathbf{x}) = A\,\mathbf{x} = \mathbf{b}$ has a solution and that solution is unique.
2. The null space of $f(\mathbf{x}) = A\,\mathbf{x}$ is nontrivial, and the equation $f(\mathbf{x}) = A\,\mathbf{x} = \mathbf{b}$ has a solution if and only if for all $\mathbf{w}$ in $N(A^T)$ we have that $\mathbf{w}\cdot\mathbf{b} = 0$.

An example

Consider the matrix

$$A = \begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 1 & 4 & 2 \\ 2 & 7 & 2 & 6 \\ 1 & 4 & 3 & 4 \end{bmatrix}$$

The standard way to determine whether or not the equation $A\,\mathbf{x} = \mathbf{b}$ has a solution for some vector $\mathbf{b}$ is to form the augmented matrix

$$\begin{bmatrix} 1 & 3 & -1 & 2 & b_1 \\ 0 & 1 & 4 & 2 & b_2 \\ 2 & 7 & 2 & 6 & b_3 \\ 1 & 4 & 3 & 4 & b_4 \end{bmatrix}$$

and then do Gauss elimination on the augmented matrix. Here are the steps in that elimination

$$\begin{bmatrix} 1 & 3 & -1 & 2 & b_1 \\ 0 & 1 & 4 & 2 & b_2 \\ 0 & 1 & 4 & 2 & b_3 - 2\,b_1 \\ 0 & 1 & 4 & 2 & b_4 - b_1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & -1 & 2 & b_1 \\ 0 & 1 & 4 & 2 & b_2 \\ 0 & 0 & 0 & 0 & b_3 - 2\,b_1 - b_2 \\ 0 & 0 & 0 & 0 & b_4 - b_1 - b_2 \end{bmatrix}$$

This tells us that in order for $A\,\mathbf{x} = \mathbf{b}$ to have a solution the vector $\mathbf{b}$ has to satisfy a pair of *auxiliary conditions*: $b_4 - b_1 - b_2 = 0$ and $b_3 - 2\,b_1 - b_2 = 0$. The Fredholm alternative tells us that we can derive these same auxiliary conditions by computing the null space of A^T and then demanding that $\mathbf{b}$ be perpendicular to all the vectors in that null space.

To compute the null space of A^T we do Gauss elimination on the augmented matrix

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 0 \\ 3 & 1 & 7 & 4 & 0 \\ -1 & 4 & 2 & 3 & 0 \\ 2 & 2 & 6 & 4 & 0 \end{bmatrix}$$

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

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

We read off from this that vectors in the null space of A^T are combinations of the vectors

$$\begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

The Fredholm alternative tells us that for $A\ \mathbf{x} = \mathbf{b}$ to have a solution we must have $\mathbf{b}$ perpendicular to all vectors in the null space of A^T. This requires that

$$\begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = b_3 - 2\, b_1 - b_2 = 0$$

$$\begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = b_4 - b_1 - b_2 = 0$$

These are just the auxiliary conditions we derived earlier.

Linear Independence, Span, and Basis

A set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ is *independent* if the only solution to the equation

$$c_1\, \mathbf{v}_1 + c_2\, \mathbf{v}_2 + \cdots + c_k\, \mathbf{v}_k = \mathbf{0}$$

is the trivial solution $c_1 = c_2 = \cdots = c_k = 0$.

A set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ *spans* a vector space (or subspace) if any vector $\mathbf{u}$ in that space can be written as a combination of the vectors $\mathbf{v}_j$:

$$c_1\,\mathbf{v}_1 + c_2\,\mathbf{v}_2 + \cdots + c_k\,\mathbf{v}_k = \mathbf{u}$$

A set of vectors that are both linearly independent and span a particular vector space is said to be a *basis* for that subspace. The number of vectors in a basis for a vector space determines that vector space's *dimension*.

Note that bases are not unique. Often more than one basis is possible for a vector space, with some bases being more "useful" than others.

Representations of Linear Operators

We have seen that the linear operator which is easiest to work with is the linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$ given by

$$f(\mathbf{x}) = A\,\mathbf{x}$$

where A is an m by n matrix. For example, if we want to solve the operator equation

$$f(\mathbf{x}) = \mathbf{b}$$

we simply have to use Gauss elimination on the matrix equation

$$A\,\mathbf{x} = \mathbf{b}$$

Given some other linear operator f that maps vectors from an n dimensional vector space V to an m dimensional vector space U, there is a procedure for constructing a special matrix, called a representation, that allows us to convert the operator equation $f(\mathbf{v}) = \mathbf{u}$ into an equivalent matrix equation.

Here is how that process works.

1. Find a basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ for the vector space V and a basis $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ for the vector space U.
2. Given some vector $\mathbf{v}$ in V, express $\mathbf{v}$ as a combination of basis vectors:

$$c_1\,\mathbf{v}_1 + c_2\,\mathbf{v}_2 + \cdots + c_n\,\mathbf{v}_n = \mathbf{v}$$

3. The vector $\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$ is called the *representation* of the vector $\mathbf{v}$ with respect to the basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ for the vector space V.
4. Likewise, we can express $f(\mathbf{v}) = \mathbf{u}$ as a combination of basis vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ for the vector space U.

$$d_1\,\mathbf{u}_1 + d_2\,\mathbf{u}_2 + \cdots + d_m\,\mathbf{u}_m = \mathbf{u} = f(\mathbf{v})$$

5. The vector $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}$ is called the representation of the vector $\mathbf{u}$ with respect to the basis $\mathbf{u}_1, \mathbf{u}_2, \ldots,$

$\mathbf{u}_m$ for the vector space U.≠

6. The m by n matrix A with the property that $A\,\mathbf{c} = \mathbf{d}$ is called the representation matrix for the linear operator f with respect to the given bases for V and U.

Constructing representations

The process outlined above gives us hope that any linear operator mapping vectors from one finite dimensional vector space to another can be converted to a matrix multiplication. There are unfortunately two things that the outline does not tell us how to do. The first of these is how to find the coordinates of a vector's representation. The second is how to actually determine the entries of the representation matrix A.

Assuming for the moment that we can find a way to easily solve the first problem, here is a clever method to solve the second problem.

1. The basis vectors $\mathbf{v}_k$ have particularly simple coordinate representations:

$$\mathbf{v}_k = 0\,\mathbf{v}_1 + 0\,\mathbf{v}_2 + \cdots + 1\,\mathbf{v}_k + \cdots + 0\,\mathbf{v}_n$$

2. Let $\mathbf{u}_k = f(\mathbf{v}_k)$ and let $\mathbf{d}_k$ be its coordinate representation with respect to the basis $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ for the vector space U.

$$d_1\,\mathbf{u}_1 + d_2\,\mathbf{u}_2 + \cdots + d_m\,\mathbf{u}_m = \mathbf{d}_k = f(\mathbf{v}_k)$$

3. We seek the matrix A such that

$$A \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{bmatrix}$$

4. The properties of matrix multiplication tell us that the product $A \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$ returns the k^{th} column of the matrix A. Thus the vector $\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{bmatrix}$ is the k^{th} column of the matrix A.

5. By allowing k to vary from 1 to n we will be able to construct all n columns of the m by n matrix A.

In our next lecture we will see how to solve the first problem, thus completing this representation algorithm.