MATH S-15, SUMMER 2001
GROUPS, GRAPHS, AND ALGEBRAIC STRUCTURES FOR COMPUTING
Lecture # 5, supplement

The Vector Space $\mathbb{R}^2$

Vectors in the xy-plane may be thought of as directed line segments beginning at the origin and ending at a point, whose coordinates we use to identify the vector. We will use the notation $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ to indicate a vector $\mathbf{v}$ whose endpoint has coordinates (x, y). There are many operations we can perform on vectors:

- We add vectors according to the rule: $\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} z \\ w \end{bmatrix} = \begin{bmatrix} x+z \\ y+w \end{bmatrix}$
- We multiply a vector by any real number r according to the rule:
$$r \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} rx \\ ry \end{bmatrix}$$
- The **length** or **norm** of the vector $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ is given by the formula
$$||\mathbf{v}|| = \sqrt{x^2 + y^2}.$$
- The **dot product** of the two vectors $\mathbf{u} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} z \\ w \end{bmatrix}$ is given by the formula
$$\mathbf{u} \cdot \mathbf{v} = xz + yw.$$
- The **angle** between two vectors $\mathbf{u}$ and $\mathbf{v}$ is given by the formula
$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{||\mathbf{u}||\ ||\mathbf{v}||}.$$

The Vector Space $\mathbb{R}^3$

Everything carries over in an analogous way from the consideration of $\mathbb{R}^2$ with three coordinates in place of two.

Abstract Vector Spaces

$\mathbb{R}^2$ and $\mathbb{R}^3$ are two of our most familiar examples of vector spaces, but it will pay off to outline their properties in general. A **vector space** (the elements of which are called "vectors") over a field F (the elements of which are called the "scalars") must satisfy the following for all vectors $\mathbf{u}$ and $\mathbf{v}$ and all scalars λ and μ:

1. The vectors form an abelian group under addition.
2. $\lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v}$
3. $(\lambda + \mu)\mathbf{u} = \lambda\mathbf{u} + \mu\mathbf{u}$
4. $\lambda(\mu\mathbf{u}) = (\lambda\mu)\mathbf{u}$
5. $1\mathbf{u} = \mathbf{u}$

Thus $\mathbb{R}^2$ and $\mathbb{R}^3$ are easily seen to be vector spaces, where the vectors have the form $\mathbf{u} = \begin{bmatrix} x \\ y \end{bmatrix}$ or $\mathbf{u} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, respectively, and where in each case the field of scalars is $\mathbb{R}$. The fact that $\mathbb{R}^2$ and $\mathbb{R}^3$ also have norms, dot products, and angles is not part of the definition of a vector space.

There are many other types of vector spaces, such as spaces of matrices and spaces of functions, and they may be defined over other fields as well, but $\mathbb{R}^2$ and $\mathbb{R}^3$ will remain our principle examples. Note that a trivial and yet somewhat interesting example is to consider a field to be a vector space over itself!

Linear Transformations

A function $f : V \longrightarrow W$, where V and W are vector spaces defined over the same field F, is said to be **linear** if the following rules hold for all vectors $\mathbf{u}$ and $\mathbf{v}$ in V and all scalars λ in F:

1. $f(\lambda\mathbf{u}) = \lambda f(\mathbf{u})$
2. $f(\mathbf{u}+\mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$

Note that another way to phrase the second rule is to say that f is a homomorphism from the abelian group V to the abelian group W since V and W are vector spaces.

Example: The function $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by $f(x) = 3x + 1$ is *not* a linear transformation because, for example, $f(1+2) = f(3) = 10$, while $f(1) = 4$ and $f(2) = 7$ so that $f(1) + f(2) = 11$, which violates the second rule.

Example: The only functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ that are linear transformations have the form $f(x) = ax$ for some real number a.

Example: The functions $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ which performs a reflection about the y-axis is a linear transformation. Check geometrically that this is given by the formula $f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -x \\ y \end{bmatrix}$. Then check that the two rules for linear transformations both hold.

Matrices

It can be verified that any linear transformation $A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is given by a rule $A\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} ax+by \\ cx+dy \end{bmatrix}$. In fact, a similar but much more general statement can be made about linear transformations $B : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ for any positive integers m and n.

Because of the special form that linear transformations take, we abbreviate them in a special notational form called matrices. For example, the function A in the previous paragraph is denoted

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

and is referred to as a 2×2 matrix. This notation generalizes to the case of the linear transformation B given above, which would be denoted by an $m \times n$ matrix, that is, one with n rows and m columns.

Operations with Matrices

The first operation that one performs with a matrix is the one from which it is derived, that is, multiplying a vector by a matrix. If $A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is a linear transformation given by the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} x \\ y \end{bmatrix}$, then

$$A\mathbf{u} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}.$$

Note that the result is a vector in $\mathbb{R}^2$, which is consistent with the range of the function A.

There are several other important operations we perform with matrices. In the following, we will illustrate only with 2×2 matrices, but you are strongly encouraged to work a few examples with matrices of different sizes.

- Adding two matrices (adding two linear transformations):

  Note that if $A : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ and $B : \mathbb{R}^m \longrightarrow \mathbb{R}^n$, then $A + B$ is also a function of the same form, that is, $A + B : \mathbb{R}^m \longrightarrow \mathbb{R}^n$.

  If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$, then

  $$A + B = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}.$$

- Multiplying two matrices (*composing* two linear transformations):

  Note that this is much trickier than addition. If $A : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ and $B : \mathbb{R}^q \longrightarrow \mathbb{R}^m$, then $A \circ B$, most often written simply as AB, acts by the composition of functions and we get a function of the form $AB : \mathbb{R}^q \longrightarrow \mathbb{R}^n$.

  Again, composition works from right-to-left, so if $\mathbf{u}$ is a vector in $\mathbb{R}^q$, then $B(\mathbf{u})$ is a vector in $\mathbb{R}^m$ and $(AB)(\mathbf{u})) = A(B(\mathbf{u}))$ is a vector in $\mathbb{R}^n$.

  In the 2×2 case, this rule works as follows:

  If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$, then $AB = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$.

  If we allow ourselves to think of the rows and columns of our matrices as vectors, then another way of thinking of this is to say that the i-th row, j-th column entry of AB is the dot product of the i-th row of A with the j-th column of B.

The following operations apply to square $n \times n$ matrices only, For the $n = 2$ case, we use $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, and for $n = 3$, we use $B = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$.

- The **determinant** of a matrix M is denoted $\det(M)$ and is given by:

$$\det(A) = ad - bc$$
$$\det(B) = aei - afh + bfg - bdi + cdh - ceg$$

- The **trace** of a matrix M is denoted $\operatorname{tr}(M)$ and is given by:

$$\operatorname{tr}(A) = a + d$$
$$\operatorname{tr}(B) = a + e + i$$

- The **transpose** of a matrix M is denoted M^t and is given by:

$$A^t = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$
$$B^t = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}$$

Note that the definitions of trace and transpose generalize very naturally to the cases when $n > 3$. The trace of any square matrix is the sum of its diagonal entries. The transpose of any square matrix may be obtained by reflecting it about its diagonal.

On the other hand, the notion of determinant is much more involved and will require greater discussion.