110.201 LINEAR ALGEBRA

Week 2 Lecture 1 Notes

1. Matrices

Denote a matrix with its entries

$$\mathbf{A}_{n\times m} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}.$$

Here, $\mathbf{A}$ has n-rows and n-columns, and these are called the *dimensions* of the two-dimensional array of numbers.

By convention, we have

(1) If $m = 1$, A is called a (column) n-vector.
(2) If $n = 1$, A is called a row m-vector.
(3) If $m = n = 1$, A is called a scalar. Typically, we do not represent scalars as 1×1-matrices. Instead, we simply treat them as (real) numbers.

Remark 1. ALL unspecified vectors are considered *column* vectors, by convention. This is similar to the notion that we always specify the horizontal axis on a graph to be the x-direction, or that numbers on the real line get larger as we move to the right. These are merely conventions to help us understand each other's more complicated constructions in math. We will see why this is useful throughout this course.

Like equations, matrices have algebraic properties:

- They can be added together if they are precisely of the same dimensions.

  $$A_{n\times m} + B_{n\times n} = C_{n\times m},$$

  where the entries of C satisfy $c_{ij} = a_{ij} + b_{ij}$.
- They can be multiplied together iff their dimensions are compatible. This means that

  $$A_{n\times m} \cdot B_{p\times r} = C_{n\times r}, \quad \text{iff} \quad m = p.$$

  Notice that the resulting matrix C has special dimensions depending on those of A and B. How does one multiply matrices?

  Definition 2. The *dot product* of a row n-vector $\mathbf{w}$ and a column n-vector $\mathbf{v}$ is defined by

  $$\mathbf{w}\cdot\mathbf{v} = \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \sum_{i=1}^{n} w_i v_i = \text{ a scalar.}$$

Date: September 11, 2013.

We use this to define a matrix product above as the matrix $\mathbf{C}_{n\times p}$, with entries

$$c_{ij} = (\text{row } i \text{ of } \mathbf{A}) \cdot (\text{column } j \text{ of } \mathbf{B}) \,.$$

Do check to see that the dimensions of $\mathbf{C}$ make sense.

Example 3. Check that this works:

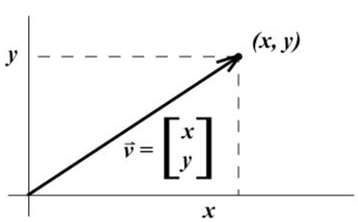
$$\begin{bmatrix} 2 & 3 \\ -1 & 0 \\ 4 & 6 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & -2 \\ 1 & 3 & -3 \end{bmatrix} = \begin{bmatrix} 2(1)+3(1) & 2(2)+3(3) & 2(-2)+3(-3) \\ -1(1)+0(1) & -1(2)+0(3) & -1(-2)+0(-3) \\ 4(1)+6(1) & 4(2)+6(3) & 4(-2)+6(-3) \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 13 & -13 \\ -1 & -2 & 2 \\ 10 & 26 & -26 \end{bmatrix}.$$

- Any matrix can be multiplied by a scalar (this is simply a multiplication of all of the entries of $\mathbf{A}$ by the scalar.) And

$$k\mathbf{A}_{n\times m} = \begin{bmatrix} ka_{11} & ka_{12} & \cdots & ka_{1m} \\ ka_{21} & ka_{22} & \cdots & ka_{2m} \\ \vdots & \vdots & \ddots & \cdots \\ ka_{n1} & ka_{n2} & \cdots & ka_{nm} \end{bmatrix}, \text{ where } k \in \mathbb{R}.$$

- We will return to this later, but the addition and multiplication of matrices satisfies
  (1) $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB} = \mathbf{AC}$,
  (2) For $k \in \mathbb{R}$, $k(\mathbf{AB}) = \mathbf{A}(k\mathbf{B})$,
  at least when the multiplications and additions make sense.



You are familiar with the notion that real n-dimensional space $\mathbb{R}^n$ is just the set of all ordered n-tuples of real numbers $(x_1, x_2, \ldots, x_n)$, where for $1 \le i \le n$, $x_i \in \mathbb{R}$. Here, we will take a slightly different view. We will consider $\mathbb{R}^n$ a *vector* space:

Definition 4. Real n-dimensional space $\mathbb{R}^n$ identifies with the set of all (column) n-vectors $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{x}$. Hence we call

$$\mathbb{R}^n = \left\{ \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \;\middle|\; x_i \in \mathbb{R},\ 1 \le i \le n \right\},$$

the (standard) n-dimensional (real) vector space.

Back to our system of equations in matrix form from last week:

$$\mathbf{Ax} = \mathbf{b}.$$

In its most general form, this matrix equation

$$\mathbf{A}_{n\times m}\mathbf{x}_{m\times 1} = \mathbf{b}_{n\times 1}$$

gives us a good new way to interpret a matrix $\mathbf{A}_{n\times m}$ beyond the idea that it is a simple bookkeeping tool for linear equations. **It is a function**, whose input is an m-vector and whose output is an n-vector. This will be quite important, and leads immediately to the definition of what kind of function to which we can associate $\mathbf{A}$:

Definition 5. The function $T : \mathbb{R}^m \to \mathbb{R}^n$, where the output n-vector $\mathbf{y} = T(\mathbf{x})$ is called a *linear transformation* if there exists an $n \times m$ matrix $\mathbf{A}$ such that $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$, for all $\mathbf{x} \in \mathbb{R}^m$.

A couple of notes:

- Writing this as $\mathbf{A}\mathbf{x} = \mathbf{y}$, we again recover the matrix form of a system of m linear equations in n unknowns. The function values $\mathbf{y}$ are easily computed for any input $\mathbf{X}$ via standard matrix multiplication. On the other hand, given a value for $\mathbf{y}$, and find a input $\mathbf{x}$ which is mapped to the chosen $\mathbf{y}$, means that one would need to "solve the system, like in Chapter 1. In this system, the matrix $\mathbf{A}$, which defines the linear transformation, is just the coefficient matrix.
- we can also write the function in the following ways: Either as $\mathbf{x} \overset{T}{\mapsto} \mathbf{y}$, or $\mathbf{x} \overset{\mathbf{A}}{\mapsto} \mathbf{y}$, and it is said that "$\mathbf{x}$ is mapped to $\mathbf{y}$ by T".
- The expression "there is", or "there exists" is written in mathematical shorthand "$\exists$". We will use this notation in the future at times.
- When the dimensions of $\mathbf{A}$ are equal (i.e., $m - n$), we call the matrix *square*. Square matrices can have some special properties. For example, when $\mathbf{A}$ is square, the linear transformation $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$ may have an inverse as a function.
- When $m = 1$, the linear transformation

$$T : \mathbb{R} \to \mathbb{R}^n, \quad T(x) = \begin{bmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_n(x) \end{bmatrix}$$

  specifies a parameterized curve (really a line) in $\mathbb{R}^n$.
- For a positive integer n, the *identity matrix* $\mathbf{I}_n$ is a special square matrix defined as

$$\mathbf{A}_{n\times n} = \mathbf{I}_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

  Here $\mathbf{I}_n\mathbf{x} = \mathbf{x}$, $\forall \mathbf{x} \in \mathbb{R}^n$. (Notice that we cannot write in this instance, $\mathbf{x}\mathbf{I}_n$. WHy not?)
- The expression "for all", or "for every" is written in mathematical shorthand "$\forall$". We will also use this notation in the future at times.

So, if a function $T : \mathbb{R}^m \to \mathbb{R}^n$ is specified by a matrix $\mathbf{A}_{n\times n}$, where $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$, then it is a linear transformation. So what happens if you do not have the matrix $\mathbf{A}$, or if you are not sure you actually have a linnear transformation. It turns out, there is a "test" for linearity: A function T will be linear iff it satisfies

(1) $T(\mathbf{v}+\mathbf{w}) = T(\mathbf{v}) + T(\mathbf{w}), \quad \forall \mathbf{v}, \mathbf{w} \in \mathbb{R}^m$, and
(2) $T(k\mathbf{v}) = kT(\mathbf{v}), \quad \forall \mathbf{v} \in \mathbb{R}^m, \quad \forall k \in \mathbb{R}$.

Remark 6. The plus signs in the first condition are quite different form each other. he one on the left is addition in $\mathbb{R}^m$, while the one on the right is addition in $\mathbb{R}^n$. This is kind of an important distinction.

Remark 7. We can combine these to say that a function T is *linear* if the function evaluated on a linear combination of input values, is a linear combination of function values. This means that

$$T(k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \ldots + k_n\mathbf{v}_n) = k_1T(\mathbf{v}_1) + k_2T(\mathbf{v}_2) + \ldots + k_nT(\mathbf{v}_n), \quad \forall k_i \in \mathbb{R},\ \mathbf{v}_i \in \mathbb{R}^m.$$

Example 8. Let $T : \mathbb{R} \to \mathbb{R}$ be defined by $T(x) = 3x$. Since $T(kx) = 3(kx) = k(3x) = kT(x)$, we know that T is linear.

Example 9. Let $T : \mathbb{R} \to \mathbb{R}$ be defined by $T(x) = x^2$. Since $T(kx) = (kx)^2 = k^2x^2 = k^2T(x)$, we know that, as long as $k \neq 1$, T is not linear.

Example 10. Is $T : \mathbb{R} \to \mathbb{R}$, where $T(x) = 3x + 2$ a linear function? Be very careful here!!

When T is "linear" (that is satisfies both condition above), then a matrix $\mathbf{A}$ will exist so that $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$.

Here are two useful constructions:

(I) Just like functions, one-to-one linear transformations have inverses from the range back to the domain. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be linear and one-to-one (don't worry about this now, but it will also be onto in this case). Then the inverse function T^{-1} will exist and $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ will be linear. Hence it will have a matrix $\mathbf{B}$, where $T^{-1}(\mathbf{x}) = \mathbf{B}\mathbf{x}$. Here, $\mathbf{B}$ is called the *inverse* matrix of $\mathbf{A}$, in the sense that

$$\left(T \circ T^{-1}\right)(\mathbf{x}) = \mathbf{A}(\mathbf{B}\mathbf{x}) = \mathbf{A}\mathbf{B}\mathbf{x} = \mathbf{I}_n\mathbf{x} = \mathbf{x} = \mathbf{B}\mathbf{A}\mathbf{x} = \mathbf{B}(\mathbf{A}\mathbf{x}) = \left(T^{-1} \circ T\right)(\mathbf{x}). \tag{1}$$

Example 11. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T(\mathbf{x}) = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\mathbf{x}$. Here, as a function T has an inverse, and can be calculated. Here, we give it to you: $T^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$ is $T^{-1}(\mathbf{x}) = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}\mathbf{x}$.

- Check the parts of Equation 1 above, for $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, and $\mathbf{B} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$.

- On vectors, choose for example $\mathbf{x} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$. Then

$$T\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 5 \end{bmatrix}, \quad \text{and} \quad T\begin{bmatrix} 8 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 8 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$