For your entertainment, some quotes from

www.math.vt.edu/activities/mathclub/quotes.html

- My geometry teacher was sometimes acute, and sometimes obtuse, but always, he was right.
- Old mathematicians never die; they just lose some of their functions.
- Mathematics is made of 50 percent formulas, 50 percent proofs, and 50 percent imagination.
- **Classification of mathematical problems as linear and nonlinear is like classification of the Universe as bananas and non-bananas.**
- A tragedy of mathematics is a beautiful conjecture ruined by an ugly fact.
- This is a one line proof... if we start sufficiently far to the left.
- The problems for the exam will be similar to those discussed in class. Of course, the numbers will be different. But not all of them. π will still be 3.14159...
- The highest moments in the life of a mathematician are the first few moments after one has proved the result, but before one finds the mistake.

Math 2270-6

Notes of 09/3–4/2019

1.8-1.9 Matrices and Linear Transformations

The main purpose of the next two meetings is to learn just why matrices are so important. They are because—in a sense we will discuss—

Matrices are the same as linear functions

- Remember that in Math 2210 we discussed **vector fields** which are vector valued functions of a vector argument.
- Let $\mathbf{f}$ be a function whose domain is $\mathbb{R}^n$ and whose range is (a subset of) $\mathbb{R}^m$, i.e.,

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{bmatrix}$$

  where $\mathbf{y}$ is in $\mathbb{R}^m$ and $\mathbf{x}$ is in $\mathbb{R}^n$.
- The function $\mathbf{f}$ is said to be **linear** if

$$\mathbf{f}(\mathbf{u}+\mathbf{v}) = \mathbf{f}(\mathbf{u}) + \mathbf{f}(\mathbf{v}) \quad \text{and} \quad \mathbf{f}(c\mathbf{u}) = c\mathbf{f}(\mathbf{u}),$$

  for all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ and scalars c.

- Recall that we defined the product of an $m \times n$ matrix A (with columns $\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_n$) and a vector $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ as

$$A\mathbf{x} = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \ldots \quad \mathbf{a}_n\,] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} x_i \mathbf{a}_i.$$

- We verified last week that if $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, and c is a scalar then

a. $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$ and

b. $A(c\mathbf{u}) = c(A\mathbf{u})$.

the properties a. and b. say that the function

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

is linear!

Note that property 2 of our definition implies that

$$\mathbf{f}(\mathbf{0}) = \mathbf{f}(0\mathbf{x}) = 0\mathbf{f}(x) = \mathbf{0}.$$

Note also that our use of the word **linear** is not consistent with the way we used it in the past. Then we called a (scalar) function linear if it was of the form

$$f(x) = mx + b.$$

The corresponding vector function would be

$$\mathbf{f}(\mathbf{x}) = A\mathbf{x} + \mathbf{b}. \tag{1}$$

These functions are not linear in the sense of linearity we defined above, in the case that

$$\mathbf{b} \neq \mathbf{0}.$$

When the distinction matters we would now call functions of the form (1) **affine** rather than linear.

Language

The textbook introduces new language to discuss the above familiar concepts. This can be confusing. In particular:

- The **range** of a function is the set of all function values. The **codomain** of a function is a possibly larger set that contains the range as a possibly proper subset. Another word for range is **image**. That word can also denote a single function value, for example if $f(3) = 7$ then 7 is the image of 3 (under f).
- Another word for a function from $\mathbb{R}^n$ to $\mathbb{R}^m$ is **transformation**. Naturally, a linear function from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a **linear transformation**.
- To confuse matters even more, functions are also called **mappings** or **maps**.
- The textbook also calls a function defined by

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

  a **matrix transformation**.
- What are the domain and codomain of $\mathbf{f}$ if A is an $m \times n$ matrix?

- mappings, maps, functions, transforms, and transformations are all the same thing.

The **social rule of mathematics**: If something has a lot of names it must be important!

- OK, so the formula

$$\mathbf{f}(\mathbf{x}) = A\mathbf{x} \tag{2}$$

  defines a linear function or transformation.

- It satisfies

$$\mathbf{f}(\mathbf{u}+\mathbf{v}) = \mathbf{f}(\mathbf{u}) + \mathbf{f}(v) \quad \text{and} \quad \mathbf{f}(c\mathbf{u}) = c\mathbf{f}(\mathbf{u}). \tag{3}$$

The remarkable—and, I think, not obvious—fact is that this also works the other way round: for every linear function $\mathbf{f}$ from $\mathbb{R}^m$ to $\mathbb{R}^n$ that satisfies (3) there is a matrix A such that (2) holds.

- In other words, matrices and linear functions (or transformations) are the same thing!

Construction of A

- To construct the matrix A we use the columns of the identity matrix

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

- The standard notation for the i-th column of I is $\mathbf{e}_i$ which is the vector in $\mathbb{R}^n$ all of whose entries are zero, except that the i-th entry equals 1.

- Recall that in math 2210 we used $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ in $\mathbb{R}^2$ and $\mathbb{R}^3$. We are now generalizing this idea to $\mathbb{R}^n$.

- Clearly, for any vector

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

we have

$$\mathbf{x} = I\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i.$$

- For example,

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

Suppose now that we are given a linear function (or transformation) T from $\mathbb{R}^n$ to $\mathbb{R}^m$. Let

$$\mathbf{a}_i = T(\mathbf{e}_i), \quad i = 1, 2, \ldots, n$$

Moreover, let A be the $m \times n$ matrix

$$A = [\, \mathbf{a}_1 \quad \mathbf{a}_2 \quad \ldots \quad \mathbf{a}_n \,]$$

Then

$$
\begin{aligned}
T(\mathbf{x}) &= T\left(\sum_{i=1}^{n} x_i \mathbf{e}_i\right) \\
&= \sum_{i=1}^{n} T\left(x_i \mathbf{e}_i\right) \qquad \text{(by part 1 of linearity!)} \\
&= \sum_{i=1}^{n} x_i T(\mathbf{e}_i) \qquad \text{(by part 2 of linearity!)} \\
&= \sum_{i=1}^{n} x_i \mathbf{a}_i \\
&= A\mathbf{x}.
\end{aligned}
$$

- In other words, with our choice of A,

$$T(\mathbf{x}) = A\mathbf{x}$$

- In this context, A is called the **standard matrix** of the transformation T.

- Example: Rotation in $\mathbb{R}^2$ by an angle θ:
- Exercise: Show geometrically that the rotation is linear.

- Example: Let T be the matrix that reflects the vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in the $y - z$ plane.

Interpretation of Linear Systems

- A mapping

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

  is said to be **onto $\mathbb{R}^m$** (or just **onto**) if each $\mathbf{b}$ in $\mathbb{R}^n$ is the image of at least one $\mathbf{x}$ in $\mathbb{R}^n$.

- In the terminology used by the textbook an equivalent statement is that the range is all of the codomain of T.

- A mapping

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

  is called **one-to-one** if each $\mathbf{b}$ in $\mathbf{r}^m$ is the image of at most one $\mathbf{x}$ in $\mathbb{R}^n$.

- While we are at it: a mapping that is both onto and one-to-one is called a **bijection**.

- Using this terminology and the transformation

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

  discuss the issues of existence and uniqueness of the solution of the linear system.

Caveats

- In the context of today's discussion we have assumed implicitly that functions are between finite dimensional vector spaces.
- For example, differentiation and integration are also linear, but they cannot in general be written in terms of matrices. (We will see later in the semester how to write those operators as matrices in special cases.)
- The matrix of a linear transformation depends on how we express vectors. We don't know yet how, but we don't need to use the standard vectors $\mathbf{e}_i$.