Linear Algebra, Section 1.9

First, some vocabulary: A function is a rule that associates objects in a set (the domain) to a unique object in a set (the **codomain**).

The **range** or **image** of f is:

$$\{y|y = f(x)\}$$

We don't talk about the codomain in calculus anymore for some reason... Think of the range (or image) as a subset of the codomain.

In calculus, we have the following definitions. These are necessary before we can talk about the *inverse* of a function.

- A function $y = f(x)$ is said to be *onto* (its codomain) if, for every y (in the codomain), there is an x such that $y = f(x)$.

  Note: Every function is automatically onto its image by definition (Since we only talk about the range in calculus, this is probably why the codomain is never mentioned anymore). Normally, the question is whether the function is onto its codomain. For example, $y = x^2$ is not onto the real line, but is onto its range, which is the interval $[0, \infty)$.

  If we don't want to specify that a function is onto its codomain, we will say that f maps x *into* the codomain.

- A function $y = f(x)$ is said to be $1 - 1$ if

  $$f(x_1) = f(x_2) \Rightarrow x_1 = x_2$$

  or (this is logically equivalent to the equation above):

  $$x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$$

  In words, this means that each element of the range came from a *unique* element of the domain. In calculus, you might remember this graphically as the horizontal line test- If any horizontal line passes through the graph of f in more than one place, then $f(x_1) = f(x_2)$, but x_1 is not x_2. For example, $y = x^2$ is not $1 - 1$ because $(-2)^2 = 2^2$, but $-2 \neq 2$.

In Section 1.9, the big theorem was that:

If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then there exists an $m \times n$ matrix A so that

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

Additionally, $A = [T(\mathbf{e}_1)\ \ T(\mathbf{e}_2)\ \ \ldots\ \ T(\mathbf{e}_n)]$

Another way to think of this is to say that every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be realized as matrix-vector multiplication. Now we are looking at the equation

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

as a *linear transformation*, where $\mathbf{x} \in \mathbb{R}^n$ is the domain, and $\mathbb{R}^m$ is the codomain.

As it turns out, linear transformations are also *continuous* and *differentiable*, but to show this we will need to be able to measure distances in $\mathbb{R}^n$ and $\mathbb{R}^m$.

1 Onto

When will $T(\mathbf{x}) = A\mathbf{x}$ be onto? This would imply that for every $\mathbf{b} \in \mathbb{R}^m$, there is (at least one) solution to $A\mathbf{x} = \mathbf{b}$. This is the setup for Theorem 4, page 43. We now list that theorem, together with our new terminology:
Theorem 4, Revised: Let A be an $m \times n$ matrix. The following are logically equivalent:

1. The function $A\mathbf{x} = \mathbf{b}$ is onto $\mathbb{R}^m$.
2. For each $\mathbf{b} \in \mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
3. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.
4. The columns of A span $\mathbb{R}^m$.
5. A has a pivot position in every row.

2 One to One

When will $T(\mathbf{x}) = A\mathbf{x}$ be 1-1? This means that the matrix equation $A\mathbf{x} = \mathbf{b}$ always has *at most* one solution (it might have no solution). Let's write out the counterpart to Theorem 4. To neatly summarize, you should write out a comparison chart!

The "One to One" Theorem: Let A be an $m \times n$ matrix. The following are logically equivalent:

1. The function $A\mathbf{x} = \mathbf{b}$ is $1 - 1$.
2. The equation $A\mathbf{x} = \mathbf{b}$ has at most one solution for every $\mathbf{b}$
3. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
4. The columns of A are linearly independent.
5. A has a pivot position in every column.

3 Worked Examples

1. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ by mapping the span of $(3, 2)$ and $(-1, 1)$ to the span of $(1, 1)$. Find a matrix that realizes this mapping, then determine if the mapping is 1-1 and/or onto:

   The mapping is not unique in this case because the question does not specify *how* the mapping performs this operation. However, since $A\mathbf{x}$ is a linear combination of the columns of A, we know that the columns of A must be scalar multiples of $(1, 1)$. Furthermore, we know that the span of

(3, 2) and (−1, 1), being linearly independent, is all of $\mathbb{R}^2$. We have lots of choices for A; one A might be:

$$A = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \Rightarrow \texttt{rref}(A) = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$$

This mapping is neither 1-1 nor onto, since there is only one pivot and two rows/columns.

2. Let A be $m \times n$ with $m < n$ so that A is a "wide" matrix. If it has m pivots, then the corresponding transformation $\mathbf{x} \to A\mathbf{x}$ will be onto $\mathbb{R}^m$, but will not be $1 - 1$ (because of the free variables).

3. Let A be $m \times n$ with $m > n$ so that A is a "tall" matrix. If it has the maximum number of pivots possible, n, then the transformation $\mathbf{x} \to A\mathbf{x}$ will be 1-1 (a pivot in every column, no free variables, so that $A\mathbf{x} = \mathbf{0}$ has only the trivial solution), but A will not be onto since there is not a pivot in every row.

   NOTE: The implication of numbers 2 and 3 is that, for $\mathbf{x} \to A\mathbf{x}$ to be 1-1 and onto, the matrix must be square, and the RREF of A must be the identity matrix.

4. Construct a mapping $\mathbf{x} \to A\mathbf{x}$ so that $(-1, 1)$ is mapped to $(3, 1)$ and $(1, 1)$ is mapped to $(2, 1)$. State whether the mapping is $1 - 1$ and/or onto:

   We'll have four unknowns for the entries of A. We cannot use our "big theorem" since we don't know how the mapping transforms $\mathbf{e}_1$ and $\mathbf{e}_2$, so we set up some equations:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

   Which translates to:

$$\begin{aligned} -a + b &= 3 \\ a + b &= 2 \end{aligned} \qquad \begin{aligned} -c + d &= 1 \\ c + d &= 1 \end{aligned}$$

   Which translates to the following two augmented matrices:

$$\left[\begin{array}{cc|c} -1 & 1 & 3 \\ 1 & 1 & 2 \end{array}\right] \qquad \left[\begin{array}{cc|c} -1 & 1 & 1 \\ 1 & 1 & 1 \end{array}\right]$$

   Whose RREF are:

$$\left[\begin{array}{cc|c} 1 & 0 & -1/2 \\ 0 & 1 & 5/2 \end{array}\right] \qquad \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 1 \end{array}\right]$$

   So the matrix A is $\begin{bmatrix} -1/2 & 5/2 \\ 0 & 1 \end{bmatrix}$, and the mapping $\mathbf{x} \to A\mathbf{x}$ will be both $1 - 1$ and onto.

   Note: A linear mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$ is uniquely defined by its action on two points.

4 Extra Practice with Functions

We've introduced linear functions from $\mathbb{R}^n$ to $\mathbb{R}^m$- how do we write nonlinear functions F from $\mathbb{R}^n$ to $\mathbb{R}^m$?

First note that there are n inputs and m outputs to F. We will write an input as $(x_1, \ldots, x_n)$, and the outputs as $(F_1, \ldots, F_m)$, where each F_i maps $\mathbb{R}^n$ to $\mathbb{R}$. Here's an example before we make the general statement. Here, F maps $\mathbb{R}^2$ to $\mathbb{R}^2$:

$$F(\mathbf{x}) = \begin{bmatrix} x_1^2 + x_2^2 \\ x_2 - x_1 \end{bmatrix} = \begin{bmatrix} F_1(x_1, x_2) \\ F_2(x_1, x_2) \end{bmatrix}$$

We write a generic function from $\mathbb{R}^n$ into $\mathbb{R}^m$ by its coordinate functions,

$$\mathbf{y} = F(\mathbf{x}) = \begin{bmatrix} f_1(x_1, x_2, \ldots, x_n) \\ f_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ f_m(x_1, x_2, \ldots, x_n) \end{bmatrix}$$

where each $f_i : \mathbb{R}^n \to \mathbb{R}$.

Here are some examples:

- $F : \mathbb{R}^2 \to \mathbb{R}^3$ by:

$$F(x, y) = \begin{bmatrix} x^2 + \sin(x) \\ x(y - 2) \\ y^2 - 3xy \end{bmatrix}$$

- $F : \mathbb{R}^3 \to \mathbb{R}^2$ by:

$$F(x, y, z) = \begin{bmatrix} x^2 - y^2 \\ 3xyz - 5 \end{bmatrix}$$

- Solving $\mathbf{y} = F(\mathbf{x})$ cannot generally be done in closed form. Normally, we would need to approximate a solution, then find a method to give us better and better approximations. Sometimes we get lucky, and the problem is easy enough to solve directly- For example, solve for x, y so that:

$$\begin{bmatrix} x^2 - y^2 \\ xy \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

  so that $x^2 = y^2$ and $xy = 1$. Therefore, $(1, 1)$ and $(-1, -1)$ are solutions.

Why study *linear* functions? While general nonlinear functions may be the norm, many times we can replace the difficult, nonlinear function by its local, linear approximation (remember doing that with the tangent line?)... More on that later in class.