EECS 16A Designing Information Devices and Systems I
Spring 2018 Lecture Notes Note 4

Mathematical Thinking & Derivation

A mathematical proof provides a means for guaranteeing that a statement is true. So what is a proof? A proof is a finite sequence of steps, called logical deductions, which establishes the truth of a desired statement. In particular, the power of a proof lies in the fact that using finite means, we can guarantee the truth of a statement with infinitely many cases.

More specifically, a proof is typically structured as follows. Recall that there are certain statements, called axioms or postulates, that we accept without proof (we have to start somewhere). Starting from these axioms, a proof consists of a sequence of logical deductions: Simple steps that apply the rules of logic. This results in a sequence of statements where each successive statement is necessarily true if the previous statements were true. This property is enforced by the rules of logic: Each statement follows from the previous statements. These rules of logic are a formal distillation of laws that were thought to underlie human thinking.

In this note, we are going to guide you through the process of developing proofs with a few examples. In particular, we aim to demonstrate the thought process of turning the problem statement into mathematical form and deriving successive mathematically rigorous statements that leads to the desired result.

When we encounter a proof problem, we generally try to understand the problem by asking the following questions:

- "What are the things we can assume based on the problem statement?"
- "What is it that we would like to show?"

The answer to the first question gives us the condition that we are working under and the answer to the second question gives us a clear picture of our goal. Then, we ask the question

- "How can we utilize what we know to get to what we would like to show under the specified condition?"

This is usually where the bulk of the work lies – this is the step at which we try to "fill in the gaps," developing a sequence of logical mathematical arguments that take us from what we know to the thing we would like to show. Sometimes we need to rewrite what we know in an alternative way; sometimes we try to establish connections between what we know and what we would like to show by introducing a new concept; sometimes we try to work backwards to give us some hints on the intermediate steps. Often times, we need to explore many different ways before we find a proof that works – but this is also where the fun lies!

Let's illustrate the idea above on the following two problems.

1. Let $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ be a set of linearly dependent vectors in $\mathbb{R}^n$. Take any matrix $A \in \mathbb{R}^{m \times n}$. Prove that the set of vectors $\{A\vec{v}_1, A\vec{v}_2, \ldots, A\vec{v}_n\}$ is linearly dependent.

Proof: (1) What do we know? Based on the problem statement, we know that $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is a set of linearly dependent vectors. How do we translate this into mathematical form? Recall one of the two definitions of linear dependence we introduced in the last lecture – the set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is linearly dependent if there exist an index i and scalars α_j's such that

$$\vec{v}_i = \sum_{j \neq i} \alpha_j \vec{v}_j. \tag{1}$$

(2) What would we like to show? We would like to show that the set of vectors $\{A\vec{v}_1, A\vec{v}_2, \ldots, A\vec{v}_n\}$ is linearly dependent. Again using the definition of linear dependence, we can translate it into a mathematical statement – we would like to show that there exist index k and scalars β_l's such that

$$A\vec{v}_k = \sum_{l \neq k} \beta_l \left(A\vec{v}_l\right). \tag{2}$$

(3) Now, how do we use what we know mathematically from (1) to prove the mathematical statement in (2)? We somehow would like to get vectors of the form $A\vec{v}$. How could we do that? Let's multiply both sides of equation (1) by the matrix A and see what happens:

$$A\vec{v}_i = A\left(\sum_{j \neq i} \alpha_j \vec{v}_j\right). \tag{3}$$

By distributivity of matrix-vector multiplication, we know that

$$A\left(\sum_{j \neq i} \alpha_j \vec{v}_j\right) = \sum_{j \neq i} A\left(\alpha_j \vec{v}_j\right) = \sum_{j \neq i} \alpha_j \left(A\vec{v}_j\right). \tag{4}$$

Now, we have that

$$A\vec{v}_i = \sum_{j \neq i} \alpha_j \left(A\vec{v}_j\right), \tag{5}$$

which is in exactly the mathematical form we would like to show according to step (2). Hence, we have completed our proof!

2. $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_1 + \vec{v}_2$ are all solutions to the system of linear equation $A\vec{x} = \vec{b}$. Prove that $\vec{b}$ must be a zero vector.

Proof: What does it mean for $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_1 + \vec{v}_2$ to be the solutions to $A\vec{x} = \vec{b}$? It means these vectors must satisfy the following equations:

$$A\vec{v}_1 = \vec{b} \tag{6}$$

$$A\vec{v}_2 = \vec{b} \tag{7}$$

$$A\left(\vec{v}_1 + \vec{v}_2\right) = \vec{b} \tag{8}$$

Notice that using distributivity of matrix-vector multiplication, equation (8) can be rewritten as

$$A\vec{v}_1 + A\vec{v}_2 = \vec{b}. \tag{9}$$

Now from equation (6) and (7), we can substitute $A\vec{v}_1$ and $A\vec{v}_2$ with the vector $\vec{b}$, which leads us to

$$\vec{b} + \vec{b} = \vec{b}. \tag{10}$$

Subtracting $\vec{b}$ from both sides of the equation above, we have

$$\vec{b} = \vec{0}. \tag{11}$$

Hence $\vec{b}$ is the zero vector, as desired.