Conversation 4: Our Second Proof

Winfried Just
Department of Mathematics, Ohio University

MATH 3200: Applied Linear Algebra

Question 5.9

Bob: It says here in Module 5:

"Prove that for any matrices **A**, **B** of the same order and scalar λ the distributivity law $\lambda\mathbf{A} + \lambda\mathbf{B} = \lambda(\mathbf{A} + \mathbf{B})$ holds.

Hint: Prove the law first for matrices of order 2×2, and then generalize to matrices of arbitrary order."

Let's give this a try!

Cindy: But I never know how to get started with these proofs!

Frank: And I don't even understand the question.
What does "scalar" mean?

Denny: They should use plain English!

Alice: Let's look up Lecture 4 and see what it says.

Bob: I think slide 12 has the info about scalars.
Let's look up that slide:

Review: Slide 12 of Lecture 4

In linear algebra, the word “scalar” simply means “number.”

For any matrix $\mathbf{A} = [a_{ij}]_{m \times n}$ and scalar λ we can define

$$\lambda \mathbf{A} = [\lambda a_{ij}]_{m \times n} = [a_{ij} \lambda]_{m \times n} = \mathbf{A} \lambda.$$

For example, when $\lambda = 3$ and $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

we get:

$$\lambda \mathbf{A} = 3\mathbf{A} = \begin{bmatrix} 3 & 6 \\ 9 & 12 \end{bmatrix}$$

Let's look for more info

Frank: OK, this clears up the "scalar" thing.

Cindy: But it doesn't help me with how to get started.

Bob: Let's follow the hint and start with matrices of order 2×2.

Cindy: But I still don't know how to get started!

Denny: Let's see what the slides say about distributivity laws.

Bob: This would be on slide 14, where we can see an example.

Review: Slide 14 of Lecture 4

- For any matrices $\mathbf{A}$, $\mathbf{B}$ of the same order and scalar λ we have $\lambda\mathbf{A} + \lambda\mathbf{B} = \lambda(\mathbf{A} + \mathbf{B})$ *(distributivity).*

For example, let $\lambda = 3$ and $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ $\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 2 & 5 \end{bmatrix}$

Then

$$\lambda\mathbf{A} + \lambda\mathbf{B} = 3\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + 3\begin{bmatrix} 1 & 0 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 9 & 12 \end{bmatrix} + \begin{bmatrix} 3 & 0 \\ 6 & 15 \end{bmatrix} = \begin{bmatrix} 6 & 6 \\ 15 & 27 \end{bmatrix}$$

Similarly,

$$\lambda(\mathbf{A} + \mathbf{B}) = 3\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 2 & 5 \end{bmatrix}\right) = 3\begin{bmatrix} 2 & 2 \\ 5 & 9 \end{bmatrix} = \begin{bmatrix} 6 & 6 \\ 15 & 27 \end{bmatrix}$$

Both calculations give the same result.

We have an example ...

Frank: So this proves that the law holds. Case closed.

Bob: That's not a proof though.

Frank: Why not?

Cindy: Didn't we already talk about this?
That a single example is not a proof?

Theo: Exactly. A proof must cover all possible examples.

Cindy: So how would the example help me with my proof?

Question C4.1: What advice would you give Cindy in response?

Theo: You need to convert the calculation for specific numbers into a calculation with symbols.

Cindy: But I don't know how!

Cindy starts her first proof

Denny: In plain English: Use letters instead of numbers.

Alice: Just give a try, Cindy. You can do it!

Cindy: You mean, like—

For example, let $\lambda = 3$ and $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $\mathbf{B} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$

Cindy: —like this? Am I on the right track?

Theo: Er, well—

Alice: Just keep going, Cindy!

Cindy completes her first proof

For example, let $\lambda = 3$ and $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $\mathbf{B} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ Then

$$\lambda\mathbf{A} + \lambda\mathbf{B} = 3\begin{bmatrix} a & b \\ c & d \end{bmatrix} + 3\begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} 3a & 3b \\ 3c & 3d \end{bmatrix} + \begin{bmatrix} 3e & 3f \\ 3g & 3h \end{bmatrix}$$

$$\lambda\mathbf{A} + \lambda\mathbf{B} = \begin{bmatrix} 3a+3e & 3b+3f \\ 3c+3g & 3d+3h \end{bmatrix}$$

$$\lambda(\mathbf{A} + \mathbf{B}) = 3\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix}\right) = 3\begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}$$

$$\lambda(\mathbf{A} + \mathbf{B}) = \begin{bmatrix} 3(a+e) & 3(b+f) \\ 3(c+g) & 3(d+h) \end{bmatrix} = \begin{bmatrix} 3a+3e & 3b+3f \\ 3c+3g & 3d+3h \end{bmatrix}$$

Both calculations give the same result!

Question C4.2: What has Cindy shown here?

Cindy has proved a result

Cindy: Whew! This wasn't too terribly hard.

Bob: Congratulations Cindy! This is almost perfect!

Cindy: Thank you Bob for saying it so nicely!
But what is wrong with my proof?

Alice: You wanted to prove that for all scalars λ and all 2×2 matrices $\mathbf{A}, \mathbf{B}$ the equality $\lambda\mathbf{A} + \lambda\mathbf{B} = \lambda(\mathbf{A} + \mathbf{B})$ holds. You actually have proved a result, but it is a slightly different one.

Cindy: What do you mean?

Alice: Take another look at your proof and tell us what it really shows.

Cindy checks her first proof

For example, let $\lambda = 3$ and $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $\mathbf{B} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ Then

$$\lambda\mathbf{A} + \lambda\mathbf{B} = 3\begin{bmatrix} a & b \\ c & d \end{bmatrix} + 3\begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} 3a & 3b \\ 3c & 3d \end{bmatrix} + \begin{bmatrix} 3e & 3f \\ 3g & 3h \end{bmatrix}$$

$$\lambda\mathbf{A} + \lambda\mathbf{B} = \begin{bmatrix} 3a + 3e & 3b + 3f \\ 3c + 3g & 3d + 3h \end{bmatrix}$$

$$\lambda(\mathbf{A} + \mathbf{B}) = 3\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix}\right) = 3\begin{bmatrix} a + e & b + f \\ c + g & d + h \end{bmatrix}$$

$$\lambda(\mathbf{A} + \mathbf{B}) = \begin{bmatrix} 3(a + e) & 3(b + f) \\ 3(c + g) & 3(d + h) \end{bmatrix} = \begin{bmatrix} 3a + 3e & 3b + 3f \\ 3c + 3g & 3d + 3h \end{bmatrix}$$

Both calculations give the same result.

Cindy: Oh I see! I have only proved that $3\mathbf{A} + 3\mathbf{B} = 3(\mathbf{A} + \mathbf{B})$.
How do I get rid of the 3 so that the proof works for all λ?

Theo fixes the proof

Theo: You need to start by writing:

Let λ be any scalar and let $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $\mathbf{B} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ Then

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

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

$$\lambda(\mathbf{A} + \mathbf{B}) = \lambda \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} \right) = \lambda \begin{bmatrix} a + e & b + f \\ c + g & d + h \end{bmatrix}$$

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

Both calculations give the same result. This proves that $\lambda\mathbf{A} + \lambda\mathbf{B} = \lambda(\mathbf{A} + \mathbf{B})$ for any scalar λ and any 2×2 matrices $\mathbf{A}, \mathbf{B}$.

How about matrices of arbitrary order?

Bob: That's exactly like Cindy's proof, but with writing "λ" instead of "3". But what if $\mathbf{A}, \mathbf{B}$ could be of any order?

Alice: We might start with: Let m, n be arbitrary positive integers, let λ be any scalar, and let $\mathbf{A}, \mathbf{B}$ be arbitrary matrices of the form

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

Denny: And then we let Cindy do the calculations of $\lambda\mathbf{A} + \lambda\mathbf{B}$ and $\lambda(\mathbf{A} + \mathbf{B})$ for us.

Bob: How about you doing them for us, Denny?

Denny: (intensely staring at his phone)

Alice: Cindy, do you think that you could do these calculations with the symbols that we have introduced here?

Cindy gets started

Cindy: You mean, like—

$\lambda\mathbf{A} + \lambda\mathbf{B}$

$$= \lambda \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} + \lambda \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{bmatrix} + \begin{bmatrix} \lambda b_{11} & \lambda b_{12} & \dots & \lambda b_{1n} \\ \lambda b_{21} & \lambda b_{22} & \dots & \lambda b_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda b_{m1} & \lambda b_{m2} & \dots & \lambda b_{mn} \end{bmatrix}$$

Cindy: —and so on?

Alice: Exactly!

Cindy could do it, but—

Cindy: Yes, I could do this.

Frank: I can't believe you are saying that, Cindy!

Cindy: Now that you mentioned it, I cannot quite believe it either that I said I could do a proof.

Bob: Of course you could, Cindy.
You already constructed a proof earlier today.

Cindy: Maybe I could, but writing all this out would be a lot of work, and with all these "..." I would certainly get mixed up somewhere in the middle.

Denny: I think there must be a simpler way of writing up this argument.

Theo: Yes there is. Look at slide 12 of Lecture 4 again.

Review: Multiplication of a matrix by a scalar

In linear algebra, the word "scalar" simply means "number."

For any matrix $\mathbf{A} = [a_{ij}]_{m \times n}$ and scalar λ we can define

$$\lambda \mathbf{A} = [\lambda a_{ij}]_{m \times n} = [a_{ij} \lambda]_{m \times n} = \mathbf{A} \lambda.$$

For example, when $\lambda = 3$ and $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

we get:

$$\lambda \mathbf{A} = 3\mathbf{A} = \begin{bmatrix} 3 & 6 \\ 9 & 12 \end{bmatrix}$$

Bob starts a proof—and Cindy continues it

Bob: We might start with: Let m, n be arbitrary positive integers, let λ be any scalar, and let $\mathbf{A}, \mathbf{B}$ be arbitrary matrices of the form

$\mathbf{A} = [a_{ij}]_{m \times n}$ and $\mathbf{B} = [b_{ij}]_{m \times n}$.

Cindy: What do we do next?

Bob: Then we calculate:

$$\lambda\mathbf{A} + \lambda\mathbf{B} = \lambda\,[a_{ij}]_{m\times n} + \lambda\,[b_{ij}]_{m\times n} = [\lambda a_{ij}]_{m\times n} + [\lambda b_{ij}]_{m\times n}$$
$$\lambda\mathbf{A} + \lambda\mathbf{B} = [\lambda a_{ij} + \lambda b_{ij}]_{m\times n}.$$

Cindy: Oh,I see! Next we calculate:

$$\lambda(\mathbf{A} + \mathbf{B}) = \lambda\left([a_{ij}]_{m\times n} + [b_{ij}]_{m\times n}\right) = \lambda\,[(a_{ij} + b_{ij})]_{m\times n}$$
$$\lambda(\mathbf{A} + \mathbf{B}) = [\lambda\,(a_{ij} + b_{ij})]_{m\times n} = [\lambda a_{ij} + \lambda b_{ij}]_{m\times n}.$$

Bob: The two calculations give the same result.

Question C4.3: What have Bob and Cindy shown here?

They have proved the Distributivity Law $\lambda\mathbf{A} + \lambda\mathbf{B} = \lambda(\mathbf{A} + \mathbf{B})$ for any scalar λ and any matrices $\mathbf{A}, \mathbf{B}$ of the same order.

Take-home message

- Writing proofs is a lot easier than most people believe.
- First make sure you fully understand the meaning of the theorem. You may want to look up the relevant notions in the lectures.
- Numerical examples can guide us in finding proofs. Sometimes it suffices to redo the same calculations as in a numerical example with symbols.
- When translating numerical calculations into symbolic ones, make sure you use symbols for every quantity the theorem talks about.
    - In our first example, these quantities would be $\mathbf{A}, \mathbf{B}$ and λ.
    - In our second example, they would be $\mathbf{A}, \mathbf{B}, \lambda, m$, and n.
- Your first attempt may only be partially successful. This is normal. Take a close look at your first version of a proof and fix it if need be.
- Using a convenient notation may simplify your work.