Teaching linear algebra through dialogues

Winfried Just
Department of Mathematics, Ohio University

Joint Mathematics Meetings, Baltimore, Maryland

January 17, 2019

An open problem

We all teach through dialogues. When a student comes to our office hours, we guide that student through the process of learning by trial, error, discussion, critique of prior attempts, corrections, and trying again. This method works.

Can one somehow adapt this method to large lecture courses?

One cannot put a group of *real* students in front of the class and having them discuss their attempts at selected homework problems, giving each other feedback on their trials and misconceptions, while the entire class comments on their mistakes.

But one might put *imaginary* students in front of the class, scripting their dialogues so that these characters make all the right mistakes, and have the real students join the discussion through the Top Hat electronic feedback system.

Back to reality: The course

I did developed and teach such a course in Fall Semester 2019: MATH3200, Applied Linear Algebra.

Learning objectives postulated that students will

- become familiar with the meaning and use of the main concepts of linear algebra,
- become competent in performing standard computational procedures of linear algebra, both with pencil and paper and with the computer algebra system MATLAB,
- become familiar with translations of real-world problems from selected domains into formal linear algebra problems,
- become competent in reading and writing elementary proofs in linear algebra.

But this is still our introductory course in linear algebra, intended primarily for engineering students, wth the parallell more abstract version geared towards math majors.

The format of the course

About half of the material was presented in the format of fairly traditional lectures, and about half of the material was presented in the form of dialogues between six characters who supposedly are studying together.

Both lectures and dialogues were interspersed with questions to the audience that students answer via their cell phones or laptops.

There were approximately 6–7 such questions per class period. They were automatically graded both for participation and correctness.

Students also were assigned approximately one practice module per class period, with questions that were automatically graded for correctness.

These materials essentially constitute a full-length e-textbook, and I am teaching a version of this course again this semester.

Let's look at some excerpts

Let us look at excerpts from two conversations.

The first one is on translating assumptions of theorems into mathematical symbols.

On translating assumptions of theorems

Bob: It says here in Module 8:

"Let $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ be two symmetric 2×2 matrices such that $a_{11} = a_{22}$ and $b_{11} = b_{22}$.
Prove that $\mathbf{AB} = \mathbf{BA}$."

Let's give this a try!

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

Alice: (gently) OK. How *might* the start of a proof look like?

Cindy: (sigh) Would I write:

Let $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$

Like this?

Bob: Good start, Cindy!

Choosing a convenient notation

Cindy: But I don't like all these subscripts. Could I write:

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

Alice: This notation might be fine; let's give it a try.

Now what do you want to prove?

Cindy: That $\mathbf{AB} = \mathbf{BA}$.

Alice: So what do you need to do next?

Does this work?

Cindy: (sigh) I need to calculate with symbols

$$\mathbf{AB} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} \text{ and } \mathbf{BA} = \begin{bmatrix} e & f \\ g & h \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

and verify that both calculations give the same result.

Bob: High 5, Cindy! This would be the thing to do here.

Frank: This isn't going to work out. Forget it!

Bob: Don't be so negative, Frank! It will work out fine.

Question C6.1: Who is right here? Frank or Bob?

Cindy: Please don't start an argument. Let me do the calculations and we will see who's right.

Why doesn't this work?

Cindy: I get:

$$\mathbf{AB} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + ch \end{bmatrix}$$

$$\mathbf{BA} = \begin{bmatrix} e & f \\ g & h \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ea + fc & eb + fd \\ ga + fg & hc + hd \end{bmatrix}$$

But I can't see why the two products are equal.

Frank: Of course not!

Bob: Stop it, Frank!!

Frank: It's not your fault, Cindy. As it says in Module 8, matrix multiplication is not commutative, so $\mathbf{AB} \neq \mathbf{BA}$.

Denny: Wow! The prof asked us to prove a wrong theorem. Some prof ...

Theo: Not so fast, Denny! **Sometimes** $\mathbf{AB} = \mathbf{BA}$, and the theorem says that this will be the case under the assumption that the matrices are symmetric, that $a_{11} = a_{22}$ and that $b_{11} = b_{22}$.

Cindy didn't make this assumption when she set up her notation.

Cindy: So I should have used the one with all those subscripts?

Alice: Not necessarily.

We can translate the assumptions into your notation.

Question C6.2: What does "$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is symmetric" mean?

Cindy: That $b = c$.

And $\mathbf{B} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ is symmetric when $f = g$.

Let's look at another excerpt

Now let us look at an excerpt from the conversation on linear dependence and linear independence.

From the conversation on linear (in)dependence

Bob: It says here in Question 12 of Module 22:

"Let $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2$ be any two vectors in $\mathbb{R}^3$. Is then the linear span *span*$(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2)$ always a plane in $\mathbb{R}^3$?"

Question C16.1: What is your answer?

Cindy: I think it must be "yes." When we take the span of 1 vector in $\mathbb{R}^3$, we get a line, when we take the span of 2 vectors in $\mathbb{R}^3$, we get a plane, when we take the span of 3 vectors in $\mathbb{R}^3$, we get the whole space $\mathbb{R}^3$. This is such a neat pattern:

When we take the span of k vectors, we get a space of dimension k.

Denny: Nice pattern! Yeah, but ... then when we take the span of 4 vectors in $\mathbb{R}^3$ we would get a 4-dimensional subspace.

How would that look?

A counterexample

Frank: Wouldn't look like anything. There are no 4-dimensional subspaces of the 3-dimensional Euclidean space $\mathbb{R}^3$. Cindy's pattern must be wrong!

Alice: The pattern that Cindy noticed actually works much of the time, but not always.

Cindy: So does it always work for dimensions 1, 2, and 3?

Bob: I think the question of Module 22 is asking precisely whether it does work for dimension 2.

Frank: And the answer is "no."

Bob: Why?

Frank: Consider $\vec{\mathbf{v}}_1 = [1, 0, 0]$ and $\vec{\mathbf{v}}_2 = [-2, 0, 0]$.

Question C16.2: What is *span*$(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2)$ in Frank's example?

What if $\vec{\mathbf{v}}_2$ is a scalar multiple of $\vec{\mathbf{v}}_1$?

Frank: When $\vec{\mathbf{v}}_1 = [1, 0, 0]$ and $\vec{\mathbf{v}}_2 = [-2, 0, 0]$, then $span(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2)$ comprises all vectors on the x-axis, but noting more. This set is a line, not a plane.

Denny: But your example is cheating, man! Your $\vec{\mathbf{v}}_2 = -2[1, 0, 0]$, so the vectors $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2$ in your example are practically the same!

Frank: They are two different vectors though. What do you mean by "practically the same" anyway?

Alice: Good question, Frank!

Cindy: I think Denny means that when $\vec{\mathbf{v}}_2 = c\vec{\mathbf{v}}_1$ for some scalar c, then $span(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2)$ can only be a line. Is that right?

Theo: This is true, and here is why: In this case, if $\vec{\mathbf{w}}$ is in $span(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2)$, then $\vec{\mathbf{w}}$ is a linear combination of $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2$ so that for some scalars d_1, d_2:

$$\vec{\mathbf{w}} = d_1\vec{\mathbf{v}}_1 + d_2\vec{\mathbf{v}}_2 = d_1\vec{\mathbf{v}}_1 + d_2(c\vec{\mathbf{v}}_1) = (d_1 + cd_2)\vec{\mathbf{v}}_1.$$

So $span(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2) = span(\vec{\mathbf{v}}_1)$ and all of these vectors form a line, not a plane.

How about the linear span of 3 vectors?

Cindy: I see. And when we take the span $span(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3)$ of 3 vectors in $\mathbb{R}^3$, and one of them is a scalar multiple of another, then we can get at most a plane, but not the entire space $\mathbb{R}^3$, right?

Theo: Right! Essentially the same calculation shows this.

Denny: And when neither of these 3 vectors is a scalar multiple of another, we get the entire $\mathbb{R}^3$!

Theo: You are jumping to conclusions here, Denny.

Denny: How is that?

Theo: Consider $\vec{\mathbf{v}}_1 = [1, 2, 3], \vec{\mathbf{v}}_2 = [-1, 1, -2], \vec{\mathbf{v}}_3 = [3, 0, 7]$. Neither of these vectors is a scalar multiple of the other two.

Denny: So $span(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3) = \mathbb{R}^3$; that's what I meant!

Theo: But that's not the case!

Bob: But I don't see why ...

How did it go?

- Performance on tests, quizzes, and the final was a lot better than when I taught in a more traditional way.
- Students did much better with proofs, significantly better with conceptual and modeling problems, but slightly worse with routine calculations like Gaussian elimination than in a traditional course.
- Performance on tests etc. was strongly correlated with participation points awarded through the Top Hat software. Those few students who failed the course had essentially given up on participation at some point.
- Many students liked the dialogues in the earlier part of the course, on proofs and the systems of linear equations.
- But as the material became more difficult and focused on conceptual understanding, resentment started building up. On balance, students *disliked* the dialogues.

My take on this

The problem on the first slide is still open.

The data on student performance show that the method has potential, but some changes to the course are needed to make it appealing in addition to being effective.

- Too much material was presented. Some of it may be a little too sophisticated for the student's level. I need to cut out some and present the remainder at a slower pace.
- Many of the 88 units in the materials were too long, which caused time management problems. I am splitting them up into shorter units.
- The materials need to be restructured in a more linear way so that especially the dialogues on the conceptual parts are more closedly linked with preceding lectures on the relevant material.

What suggestions for improvement would you have?