Conversation 30B: Linear transformations etc.: How are all these concepts related?

Winfried Just
Department of Mathematics, Ohio University

MATH3200: Applied Linear Algebra

Review: Matrix representations of linear transformations

In the first part of this conversation, our protagonists discussed the following result:

Theorem (Matrix representation of linear transformations)

Suppose $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation.

If both the elements of the domain $\mathbb{R}^n$ of L and the function values $L(\vec{\mathbf{x}})$ in $\mathbb{R}^m$ are treated as column vectors. Then there exists a matrix $\mathbf{A}$ of order $m \times n$ such that $L = L_{\mathbf{A}}$, that is,

$L(\vec{\mathbf{x}}) = \mathbf{A}\vec{\mathbf{x}}$ for all $\vec{\mathbf{x}}$ in $\mathbb{R}^n$.

What is this good for?

Frank: This theorem really shows that at some level linear transformations are just another way of looking at matrix multiplication. So why do we have to learn about them?

Alice: Thinking about linear transformations makes it easier to understand what all these other concepts we learned about mean, and how they are related to each other.

For example, a lot about a matrix $\mathbf{A}$ can be learned from studying the *range* of the linear transformation $L_{\mathbf{A}}$, that is, from the set of values that this function takes.

When $\mathbf{A}$ is of order $m \times n$, then we say that $L_{\mathbf{A}}$ maps $\mathbb{R}^n$ *onto* $\mathbb{R}^m$ if the range is the entire space $\mathbb{R}^m$, that is, when every vector in $\mathbb{R}^m$ is a function value of $L_{\mathbf{A}}$.

Consistency and linear transformations

Theo: Let $\mathbf{A}$ be the coefficient matrix of a system of linear equations

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

Then for a given vector $\vec{\mathbf{x}} \in \mathbb{R}^n$ we have $L_{\mathbf{A}}(\vec{\mathbf{x}}) = \mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ if, and only if, $\vec{\mathbf{x}}$ is a solution of the system. Thus $\vec{\mathbf{b}}$ would be a value of the function $L_{\mathbf{A}}$ if, and only if, the above system is consistent.

Theo: Instead of saying "$\vec{\mathbf{b}}$ is a value of the function $L_{\mathbf{A}}$," we could also say that "$\vec{\mathbf{b}}$ is in the range of $L_{\mathbf{A}}$."

Denny: Cool! So could we then say that $\vec{\mathbf{b}}$ is in the column space $CS(\mathbf{A})$ of $\mathbf{A}$ if, and only if, $\vec{\mathbf{b}}$ is in the range of $L_{\mathbf{A}}$?

Question C30.1: Did Denny get this right?

Theo: Exactly!

Consistency and the rank

Theo: In particular, when $L_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^m$ maps $\mathbb{R}^n$ *onto* $\mathbb{R}^m$ —

Cindy: Do you mean that every vector $\vec{\mathbf{b}}$ in $\mathbb{R}^m$ is a function value $L_{\mathbf{A}}(\vec{\mathbf{x}})$ for some $\vec{\mathbf{x}}$?

Theo: Exactly!

Denny: Then the system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is consistent for every possible vector $\vec{\mathbf{b}}$ in $\mathbb{R}^m$!

Cindy: Didn't we see something like this before, but about the rank of $\mathbf{A}$? Was it the same as $\mathbf{A}$ having full rank?

Question C30.2: What is the connection with $r(\mathbf{A})$ here?

Alice: You made a good connection, Cindy. But it works this way only for square matrices $m \times m$ that can have full rank. In general, when n may be different from m, the linear transformation $L_{\mathbf{A}}$ will map $\mathbb{R}^n$ onto $\mathbb{R}^m$ if, and only if, $r(\mathbf{A}) = m$, where m is the number of rows of $\mathbf{A}$.

$L_{\mathbf{A}}$ and the rank $r(\mathbf{A})$

Theo: In fact, the observations of Denny, Cindy, and Alice together essentially add up to a proof of the following theorem:

Theorem

Let $\mathbf{A}$ *be an* $m \times n$ *matrix. Then the following properties are equivalent:*

- $r(\mathbf{A}) = m$
- $L_{\mathbf{A}}$ *maps* $\mathbb{R}^n$ *onto* $\mathbb{R}^m$.
- *Every system* $\mathbf{A}\vec{x} = \vec{\mathbf{b}}$ *is consistent.*

When $m = n$*, that is, when* $\mathbf{A}$ *is a square matrix, then each of the above is also equivalent to*

- $\mathbf{A}$ *has full rank.*

Denny: What do you mean by "properties are equivalent"?

Theo: That they basically say the same thing about the matrix. Only in different terminology.

Bob: Now let's assume a system of linear equations

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

is underdetermined. Does this also have a connection with $L_{\mathbf{A}}$?

Theo: Yes. Then we must have at least two different solutions, $\vec{\mathbf{x}} \neq \vec{\mathbf{y}}$, so that $L_{\mathbf{A}}(\vec{\mathbf{x}}) = L_{\mathbf{A}}(\vec{\mathbf{y}}) = \vec{\mathbf{b}}$.

So in this case $L_{\mathbf{A}}$ *cannot* be a *one-to-one* function.

Denny: Cool! So could you then say that the system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is underdetermined, if and only if, $L_{\mathbf{A}}$ is *not* one-to-one?

Question C30.3: Did Denny get this right?

Frank: Which $\vec{\mathbf{b}}$ are you talking about, Denny?

Denny: Oh, I see! $L_{\mathbf{A}}$ may not be one-to-one, but the system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ may be inconsistent. So I should have said:
"A consistent system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is underdetermined, if, and only if, $L_{\mathbf{A}}$ is not one-to-one."

Theo: This is true, but not entirely obvious.

Denny: Why not?

Theo: When $L_{\mathbf{A}}$ is not one-to-one, there is **some** vector $\vec{\mathbf{b}}'$ so that for some $\vec{\mathbf{x}} \neq \vec{\mathbf{y}}$ we have $L_{\mathbf{A}}(\vec{\mathbf{x}}) = L_{\mathbf{A}}(\vec{\mathbf{y}}) = \vec{\mathbf{b}}'$, which means that both $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ are solutions of the system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}'$.
But how do you know that $\vec{\mathbf{b}} = \vec{\mathbf{b}}'$?

Denny: Yeah, my $\vec{\mathbf{b}}$ could be different ... Now I'm confused.

Cindy: But Theo, for your vectors $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ we must then have $L_{\mathbf{A}}(\vec{\mathbf{x}} - \vec{\mathbf{y}}) = L_{\mathbf{A}}(\vec{\mathbf{x}}) - L_{\mathbf{A}}(\vec{\mathbf{y}}) = \vec{\mathbf{b}}' - \vec{\mathbf{b}}' = \vec{\mathbf{0}}$. So $\vec{\mathbf{x}} - \vec{\mathbf{y}}$ would be in the null space $N(\mathbf{A})$, and $\vec{\mathbf{x}} - \vec{\mathbf{y}} \neq \vec{\mathbf{0}}$, since $\vec{\mathbf{x}} \neq \vec{\mathbf{y}}$.

Question C30.4: What does this imply about $N(\mathbf{A})$?

$L_{\mathbf{A}}$ and the null space $N(\mathbf{A})$

Cindy: So $N(\mathbf{A})$ contains some vector $\vec{\mathbf{z}} \neq 0$. Like $\vec{\mathbf{z}} = \vec{\mathbf{x}} - \vec{\mathbf{y}}$.

Alice: Excellent observation, Cindy!

Denny: How does this help me with my $\vec{\mathbf{b}}$ if I only know about $\vec{\mathbf{b}}'$?

Cindy: As we discovered in Conversation 29, if $\vec{\mathbf{x}}$ is one solution of your system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$, —

Denny: — and I have said "consistent," so there is such a $\vec{\mathbf{x}}$!

Cindy: Then $\vec{\mathbf{x}} + \vec{\mathbf{z}}$ is also a solution for any $\vec{\mathbf{z}}$ in the null space $N(\mathbf{A})$. So, Denny, if you take $\vec{\mathbf{z}} \neq \vec{\mathbf{0}}$ in $N(\mathbf{A})$, then you get another solution, and your system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ must be underdetermined.

Denny: Sweet!

Bob: Great observations, Cindy!

Another theorem

Theo: Also recall that $dim(N(\mathbf{A})) = n - r(\mathbf{A})$. Taken together, Cindy and Denny's observations add up to the following theorem:

Theorem

Let $\mathbf{A}$ *be an* $m \times n$ *matrix. Then the null space* $N(\mathbf{A})$ *is the set of all vectors* $\vec{\mathbf{x}}$ *with* $L_{\mathbf{A}}(\vec{\mathbf{x}}) = \vec{\mathbf{0}}$*. It has dimension* $dim(N(\mathbf{A})) > 0$ *if, and only if,* $L_{\mathbf{A}}$ *is not one-to-one.*

Moreover, the following properties are equivalent:

- $r(\mathbf{A}) < n$.
- $L_{\mathbf{A}}$ *is not one-to-one.*
- *The homogeneous system* $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{0}}$ *is underdetermined.*
- *Every consistent system* $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ *is underdetermined.*

When $m = n$*, that is, when* $\mathbf{A}$ *is a square matrix, then each of the above is also equivalent to*

- $\mathbf{A}$ *does not have full rank.*

A different perspective on Theo's theorems

Frank: Let me see whether I got this straight.

I'll look at this theorem from the opposite angle.

Let me focus on the case when $\mathbf{A}$ is a square matrix.

Cindy: Yes, please do! Then $m = n$.

Frank: If $\mathbf{A}$ does have full rank, then this theorem implies that $L_{\mathbf{A}}$ is one-to-one, and the previous one says that then $L_{\mathbf{A}}$ is onto. Then for every $\vec{\mathbf{b}}$ in $\mathbb{R}^n$ there exists exactly one vector $\vec{\mathbf{x}}$ that is the solution of the system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$. Let me call this vector $\vec{\mathbf{x}} = L_{\mathbf{A}}^{-1}(\vec{\mathbf{b}})$.

Alice: What you have defined here is the *inverse* $L_{\mathbf{A}}^{-1}$ of the linear transformation $L_{\mathbf{A}}$. This is also a linear transformation from the space of column vectors $\mathbb{R}^n$ to the space of column vectors $\mathbb{R}^n$. It exists if, and only if, the transformation $L_{\mathbf{A}}$ is one-to-one and onto.

Frank: So we must then have $L_{\mathbf{A}}^{-1} = L_{\mathbf{B}}$ for some $n \times n$ matrix $\mathbf{B}$.

Question C30.5: What is this matrix $\mathbf{B}$?

A different perspective on Theo's theorems, completed

Cindy: Is it $\mathbf{B} = \mathbf{A}^{-1}$?

Alice: Yes, Cindy. That's the one.

We can understand why when we think about multiplication with $\mathbf{A}^{-1}$ **undoing** the effect of multiplication with $\mathbf{A}$, and of Frank's inverse function $L_{\mathbf{A}}^{-1}$ as **undoing** the effect of the transformation $L_{\mathbf{A}}$, as taking us back to our input vector.

Thus we must have $L_{\mathbf{A}}^{-1} = L_{\mathbf{A}^{-1}}$.

Cindy: Wait! Does it follow from what Alice and Frank told us that an $n \times n$ matrix $\mathbf{A}$ is invertible if, and only if, $r(\mathbf{A}) = n$?

Theo: Exactly. Our discussion here outlines a formal proof of this result.

Take-home message

While in general there are various kinds of linear transformation, for each linear transformation L from the space of column vectors $\mathbb{R}^n$ into the space of column vectors $\mathbb{R}^m$ there exists an $m \times n$ matrix $\mathbf{A}$ such that $L(\vec{\mathbf{x}}) = L_{\mathbf{A}}(\vec{\mathbf{x}}) = \mathbf{A}\vec{\mathbf{x}}$.

The columns of the matrix $\mathbf{A}$ for which $L = L_{\mathbf{A}}$ are the function values $L(\vec{\mathbf{e}}_i)$ of the standard basis vectors.

By considering the linear transformation $L_{\mathbf{A}}$ we can see important connections between the various concepts that were covered in Chapter 3.

The two theorems on the next two slides show how formally different properties relate to each other in the case of square matrices.

Take-home message: Summary of some results from Chapter 3 for the special case of a square matrix

Theorem

Let $\mathbf{A}$ be an $n \times n$ matrix. Then the following properties are equivalent:

- $r(\mathbf{A}) = n$, *that is,* $\mathbf{A}$ *has full rank.*
- *The column vectors of* $\mathbf{A}$ *form a linearly independent set.*
- *The row vectors of* $\mathbf{A}$ *form a linearly independent set.*
- *Every* $\vec{\mathbf{b}}$ *in* $\mathbb{R}^n$ *is a linear combination of the columns of* $\mathbf{A}$.
- $L_{\mathbf{A}}$ *maps* $\mathbb{R}^n$ *onto* $\mathbb{R}^n$.
- *Every system* $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ *is consistent.*
- *The linear transformation* $L_{\mathbf{A}}$ *is a one-to-one map.*
- *The system* $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{0}}$ *has exactly one solution.*
- *Every system* $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ *has exactly one solution.*
- $\mathbf{A}$ *is invertible, that is,* $\mathbf{A}^{-1}$ *exists.*

Take-home message: An alternative version of this theorem

Theorem

Let $\mathbf{A}$ *be an* $n \times n$ *matrix. Then the following properties are equivalent:*

- $r(\mathbf{A}) < n$, *that is,* $\mathbf{A}$ *does not have full rank.*
- *The column vectors of* $\mathbf{A}$ *form a linearly dependent set.*
- *The row vectors of* $\mathbf{A}$ *form a linearly dependent set.*
- *Some* $\vec{\mathbf{b}}$ *in* $\mathbb{R}^n$ *is not a linear combination of the columns of* $\mathbf{A}$.
- $L_{\mathbf{A}}$ *does not map* $\mathbb{R}^n$ *onto* $\mathbb{R}^n$.
- *Some system* $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ *is inconsistent.*
- *The linear transformation* $L_{\mathbf{A}}$ *is not a one-to-one map.*
- *The system* $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{0}}$ *is underdetermined.*
- *Every consistent system* $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ *is underdetermined.*
- $\mathbf{A}$ *is not invertible, that is,* $\mathbf{A}^{-1}$ *does not exist.*