Chapter 8: Linear transformations **Section 1**

Matrix transformations

What you need to know already:	**What you can learn here:**
➤ All basic properties of Euclidean vectors and matrices. ➤ What functions are.	➤ The functions that are considered in linear algebra as relating vectors. ➤ A special kind of such functions.

Are we going to transform matrices in this section?

Not quite. We have seen repeatedly how linear algebra has its own terminology and the word *transformation* is an important part of this vocabulary.

Although this word is most often accompanied by a qualifying adjective (matrix, linear, etc.) let me give you a useful definition that will provide a bridge with a more familiar concept

Definitions

A ***transformation*** is a rule that associates to each vector in $\mathbb{R}^n$ a vector in $\mathbb{R}^m$, where *n* and *m* are two integer numbers.

$\mathbb{R}^n$ is called the ***domain*** of the transformation.

$\mathbb{R}^m$ is called the ***codomain*** of the transformation.

And there is some notation associated with a transformation.

Definition

A transformation is usually denoted with a capital italic letter, followed by the domain and codomain and, if needed, a formula indicating how the transformation works:

$$T: \mathbb{R}^n \to \mathbb{R}^m,\ T(\mathbf{v}) = \mathbf{w}$$

That looks very similar to a function!

That's because it is! The only difference is in the fact that, unlike a function that may work on just some numbers, a transformation *must* work on every vector in its domain. In fact, you can start by viewing the word *transformation* as the linear algebra version of *function*. That is why I said that this concept would be familiar to you. Notice that we have already seen many transformations.

Example: Embeddings and truncations

The embedding of $\mathbb{R}^n$ into $\mathbb{R}^{n+p}$, obtained by associating to each vector **v** the vector obtained by adding *p* 0's as additional components at the end, is a transformation. We denote it by:

$$E:\mathbb{R}^n \to \mathbb{R}^{n+p},\ E\left[v_1\ v_2 \cdots v_n\right] = \left[v_1\ v_2 \cdots v_n \quad 0 \cdots 0\right]$$

Similarly, the truncation of $\mathbb{R}^{n+p}$ into $\mathbb{R}^n$ obtained by associating to each vector **v** the vector obtained by deleting the last p components, is a transformation. We denote it by:

$$P:\mathbb{R}^{n+p} \to \mathbb{R}^n,\ P\left[v_1\ v_2 \cdots v_n \cdots v_{n+p}\right] = \left[v_1\ v_2 \cdots v_n\right]$$

Wait! If you used the letter E for an embedding, why did you use P for a truncation?

Couldn't slip it past you, eh? Good! The reason is that we use the letter T for a generic transformation and I wanted to be more specific here. I used P because a truncation is a special type of projection. In fact…

Example: Projections

The projection of any vector onto a given vector u is a transformation, defined by:

$$P_{\mathbf{u}}:\mathbb{R}^n \to \mathbb{R}^n,\ P_{\mathbf{u}}(\mathbf{v}) = \frac{\mathbf{v}\bullet\mathbf{u}}{\mathbf{u}\bullet\mathbf{u}}\mathbf{u}$$

There are also some transformations that we have not seen so far and, it turns out, we shall not study in the rest of the course either!

Example: $S:\mathbb{R}^3 \to \mathbb{R}^3,\ S\left(\left[a\ \ b\ \ c\right]\right) = \left[a^2\ \ b^2\ \ c^2\right]$

Since this rule can be applied to any vector, its domain is all of $\mathbb{R}^3$ and this is a transformation.

However, the rule that associates to each 3D vector $\mathbf{v} = \left[a\ \ b\ \ c\right]$ the vector $\mathbf{v}_2 = \left[\ln a\ \ \ln b\ \ c\right]$ is NOT a transformation, because it cannot be applied to all vectors in its domain.

However, not every rule that may be viewed as a function is a transformation.

Example: $S:\mathbb{R}^3 \to \mathbb{R}^3,\ S\left(\left[a\ \ b\ \ c\right]\right) = \left[\ln a\ \ \sqrt{b}\ \ c\right]$

This rule is NOT a transformation, because it cannot be applied to all vectors in $\mathbb{R}^3$. Although it may be viewed as a function, with a limited domain, it is not a transformation.

Notice that we called the other space involved in a transformation the *codomain*, NOT the *range*, as you may expect from the similar words in calculus. That is because, as in calculus, it is possible that not every vector in the codomain comes from a vector in the domain. For instance, not every vector in the codomain of an embedding comes from a vector in the domain, only those whose last *p* components are 0. So, we need another definition, which is also very familiar.

Definition

If $T:\mathbb{R}^n \to \mathbb{R}^m$ is a transformation, for any vector **v** in the domain $\mathbb{R}^n$, the vector $T(\mathbf{v})$ in the codomain $\mathbb{R}^m$ is called the ***image*** of **v** under *T*.

The subset of the codomain $\mathbb{R}^m$ consisting of ***all images*** obtained from *T* is called the ***range*** of *T*.

Example: $S:\mathbb{R}^3 \to \mathbb{R}^3, S([a\ b\ c]) = [a^2\ b^2\ c^2]$

The codomain of this transformation is $\mathbb{R}^3$ because that is where all images end up being.

However, the range consists of only those vector whose components are not negative. So, $[1\ 2\ 3]$ is in the range, while $[-1\ 2\ 3]$ is in the codomain, but not in the range.

And here is another connection with functions.

Technical fact

Two transformations with the same domain and codomain may be ***added***, as follows:

$$(T_1 + T_2)(\mathbf{v}) = T_1(\mathbf{v}) + T_2(\mathbf{v})$$

If $T_1:\mathbb{R}^n \to \mathbb{R}^m$ and $T_2:\mathbb{R}^m \to \mathbb{R}^p$ are two transformations, their ***composition*** is given by:

$$T_2 \circ T_1:\mathbb{R}^n \to \mathbb{R}^p, \quad T_2 \circ T_1(\mathbf{v}) = T_2(T_1(\mathbf{v}))$$

Been there, done that!

Glad to hear it, and for now we shall put this idea of combining functions on the backburner, but it was basic enough that it belongs here. Now for some newer concepts.

Definition

A transformation $T:\mathbb{R}^n \to \mathbb{R}^m$ is said to be:

- ***One-to-one***, or ***injective***, if for any two vectors **v** and **w** in $\mathbb{R}^n$, $T(\mathbf{v}) \neq T(\mathbf{w})$. That is, the images of different vectors are different.
- ***Onto***, or ***surjective***, if the ***range and codomain are the same***. That is, every vector in the codomain comes from a vector in the domain.

Example: ***Embeddings and truncations***

An embedding is one-to-one (injective), since adding 0's does not change the fact that the old sets of components are different. However, it is not onto (surjective), since vectors whose added components are not zero are not images of any vector in the domain.

On the other hand, a truncation is surjective (onto), since every vector in the codomain is the truncation of some higher dimensional vector with the same set of components at the beginning. However, it is not injective (one-to-one), since different vectors with the same initial components will have the same image.

One-to-one, onto, injective, surjective! Strange words!

Yes, all part of linear algebra tradition!

And of course, this being linear algebra, we shall restrict our attention only to certain types of transformations, namely those that fit in with all other concepts we have seen so far related to vectors and matrices. That is where the next type of transformations comes in.

Definition

Given an $m \times n$ matrix $\mathbf{A}$, the transformation defined by:

$$T_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^m, T_{\mathbf{A}}(\mathbf{v}) = \mathbf{A}\mathbf{v}$$

is called the ***matrix transformation*** associated with, or generated by the matrix **A**.

Notice that since the product used in the definition can always be computed (all dimensions fit), this is indeed a transformation. Just for fun, let's see a typical example.

Example: $\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 4 \end{bmatrix}$

The transformation generated by this matrix is:

$$T_{\mathbf{A}} : \mathbb{R}^3 \to \mathbb{R}^2,\ T_{\mathbf{A}} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + 2y - z \\ 3y + 4z \end{bmatrix}.$$

Notice that this transformation, by using a 2×3 matrix, changes a 3D vector into a 2D vector. Therefore, its domain is $\mathbb{R}^3$ and its codomain is $\mathbb{R}^2$, as required in the definition. Had you noticed that detail?

Is this transformation injective? Well, in order for it to be so, each 2D vector must come from a unique 3D vector. That is, the system $\mathbf{Ax} = \mathbf{c}$ must have a unique solution for any vector **c**

But the matrix **A** has rank 2 (can you see why?) and the system has a free variable, therefore it has infinitely many solutions. Therefore, it cannot be injective.

Is it surjective? Again because the system will have infinitely many solutions for any vector **c**, yes, it is surjective.

Notice how even in this basic example we used the old knowledge we have about matrices and vectors, as well as the new definition of transformation. In particular, its being injective or surjective depends on the rank of the matrix and the number of solutions available, old friends of ours by now.

But there is another important connection.

Technical fact

Two matrix transformations with the same domain and codomain may be ***added*** by ***adding the corresponding matrices***:

$$(T_{\mathbf{A}} + T_{\mathbf{B}})(\mathbf{v}) = \mathbf{Av} + \mathbf{Bv} = (\mathbf{A} + \mathbf{B})\mathbf{v} = T_{\mathbf{A}+\mathbf{B}}(\mathbf{v})$$

Two matrix transformations are ***composed by multiplying the corresponding matrices*** in the proper order:

$$(T_{\mathbf{A}} \circ T_{\mathbf{B}})(\mathbf{v}) = T_{\mathbf{A}}(T_{\mathbf{B}}(\mathbf{v})) = \mathbf{A}(\mathbf{Bv}) = (\mathbf{AB})\mathbf{v} = T_{\mathbf{AB}}(\mathbf{v})$$

So, many of the matrix algebra properties we have seen so far can be transferred to information about matrix transformations, right?

Yes! For instance:

Technical fact

The matrix transformation associated to a matrix **A** is ***invertible*** if and only if **A** is invertible.

In that case $(T_{\mathbf{A}})^{-1} = T_{\mathbf{A}^{-1}}$.

I hope you will not be surprised by the fact that I am leaving the simple proof of this fact to you! I trust it will help you understand the new concepts, as well as their connections to the old ones.

Example: $\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$

This matrix determines the transformation $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+3y \\ 2y \end{bmatrix}$.

The fact that **A** is invertible tells us that its transformation has an inverse as a function. In fact, you can check that such inverse is given by:

$$T^{-1}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x-1.5y \\ 0.5y \end{bmatrix} = \begin{bmatrix} 1 & -1.5 \\ 0 & 0.5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{A}^{-1}\begin{bmatrix} x \\ y \end{bmatrix}$$

Check what? What do you mean that the transformation is invertible?

Exactly what you would expect from calculus.

Definition

A transformation T is ***invertible*** if there is another transformation T^{-1}, called its inverse, such that for every vector **v** in the domain:

$$T\left(T^{-1}(\mathbf{v})\right) = T^{-1}\left(T(\mathbf{v})\right) = \mathbf{v}$$

Before I send you to play with these basic ideas, here is an obvious little fact that will prove very useful later:

Technical fact

For every matrix **A**, $T_{\mathbf{A}}(\mathbf{0}) = \mathbf{0}$.

And the "very simple" proof is left to me, I get it!

Summary

- A transformation is a function that associates to every vector in a Euclidean space another vector in the same space or another one.
- A matrix transformation is a transformation defined through multiplication by a constant matrix.

Common errors to avoid

- Watch out for notation and terminology: some of it is new and mysterious, some uses old friends with new meanings. Use them all correctly and in their proper context.

Learning questions for Section LA 8-1

Review questions:

1. Explain what a transformation is.
2. Explain what a matrix transformation is.
3. Describe how addition and composition of matrix transformations corresponds to matrix operations.
4. Explain the difference between the range and the codomain of a transformation.

Memory questions:

1. When is a matrix transformation *injective*?
2. When is a matrix transformation *surjective*?
3. What is the domain of a transformation?
4. What is the codomain of a transformation?
5. What is the range of a transformation?
6. What is the image of the **0** vector under a matrix transformation?

Computation questions:

1. Explain why the transformation $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y+3 \\ y+z-2 \end{bmatrix}$ is not a matrix transformation.

2. Explain why the transformation $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ xy \end{bmatrix}$ is not a matrix transformation.

For each of the matrices provided in questions 3-6, construct the formula that defines the image of a generic vector, find the domain and codomain and determine if it is injective and/or if it is surjective.

3. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

4. $\begin{bmatrix} 2 & -3 & 0 \\ 0 & 2 & 3 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

6. $\begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix}$

Show that each of the transformations provided in question 7-10 is a matrix transformation by constructing its matrix. Then determine its domain, codomain and range and decide whether it is injective and/or surjective.

7. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x\cos 1 + y\sin 1 \\ y\cos 1 + z\sin 1 \end{bmatrix}$

8. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x-3y \\ 4y-2x \end{bmatrix}$

9. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ y-z \\ 2x-y+z \end{bmatrix}$

10. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+2y \\ x-2y \\ x+y \end{bmatrix}$

Determine which of the matrices provided in questions 11-14 are associated with an invertible transformation.

11. $\begin{bmatrix} 2 & 1 \\ 1 & -3 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

13. $\begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix}$

14. $\begin{bmatrix} 2 & -1 & 0 \\ -4 & 3 & 0 \\ 1 & 1 & 1 \end{bmatrix}$

15. Construct the matrix of the transformation that projects a vector in R^2 perpendicularly onto the line $y = x$.

16. Find the matrix of the transformation that switches the vectors $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ 5 \end{bmatrix}$ and determine whether it is one-to-one.

17. Consider the transformations $T_{\mathbf{A}}$ and $T_{\mathbf{B}}$ determined by the matrices $\mathbf{A} = \begin{bmatrix} 2 & 4 & 6 \\ -3 & 0 & 1 \\ 0 & 5 & 8 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} -4 & 2 & 0 \\ 3 & 7 & 1 \\ 8 & 8 & 2 \end{bmatrix}$. What are the images of the standard unit vectors under the composition $T_{\mathbf{A}} \circ T_{\mathbf{B}}$?

18. Show that the transformation that associates to any vector in $\mathbb{R}^3$ its dot product with the vector $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ is a matrix transformation and identify its domain, codomain and range.

Theory questions:

1. What type of calculus function are one-to-one, like injective transformations?
2. Identify one property of all transformations that is NOT shared by all calculus functions.
3. Which transformation is determined by the **0** matrix?
4. If T is a matrix transformation, under what conditions on the determinant of its matrix **A** is T invertible?

Proof questions:

1. Prove that for every matrix **A**, $T_{\mathbf{A}}(\mathbf{0}) = \mathbf{0}$.
2. Prove that the matrix transformation associated to a matrix **A** is invertible if and only if **A** is invertible, and that in such case $(T_{\mathbf{A}})^{-1} = T_{\mathbf{A}^{-1}}$.

Templated questions:

1. Construct a small matrix (not too small!) and then obtain the formula for its transformation and check if it is injective and/or surjective.
2. Construct a formula for a transformation, determine its domain, codomain and range and check if it is a matrix transformation.

What questions do you have for your instructor?