Condensed notes on linear algebra for IMDS

Jeffrey Giansiracusa

October 15, 2021

1 Vectors

A vector is a list of numbers (usually these will be real number, but you might encounter complex numbers or more exotic things out in the world). We will usually write vectors as columns, e.g.:

$$\mathbf{v} = \begin{pmatrix} 1 \\ 4 \\ 7 \\ \vdots \\ 9 \end{pmatrix}$$

A component of a vector is just one of the entries in the list.

We write $\mathbb{R}^n$ for the set of all vectors with n-components, where n is some number. E.g., $\mathbb{R}^2$ is the set of all vectors $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$.

We think of $\mathbb{R}^2$ as a plane (2-dimensional space) and $\mathbb{R}^3$ as 3-space, so we can draw pictures of vectors in 2 and 3 dimensions.

1.1 Vector addition

We can add vectors by adding them component by component.

Example 1.

$$\begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} + \begin{pmatrix} 0.1 \\ 0.9 \\ 2.5 \end{pmatrix} = \begin{pmatrix} 1.1 \\ 2.9 \\ 7.5 \end{pmatrix}$$

Note that you are only allowed to add vectors if they have the same length.

1.2 Scalar multiplication

When talking about vectors, we often use the word scalar as a fancy-sounding word for number. The idea is that a scalar can be used to scale (stretch) vectors (make them shorter or longer without changing their direction).

If $\lambda \in \mathbb{R}$ is a scalar and $\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$ is a vector, then $\lambda v \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix}$.

Example 2.

$$2\begin{pmatrix} 4 \\ 7 \end{pmatrix} = \begin{pmatrix} 8 \\ 14 \end{pmatrix}.$$

2 Linear combinations and span

A linear combination of a bunch of vectors $\mathbf{v}_1, \ldots \mathbf{v}_k$ is simply a vector we make from these using scalar multiplication and addition:

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n,$$

where the λ_i are scalars.

The span of a set of vectors is the set of all vectors we can make by taking linear combinations.

Example 3. The span of the vectors

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{b}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{b}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

is all of $\mathbb{R}^3$ since we can make an arbitrary vector $\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ as a linear combination

$$x\mathbf{b}_1 + y\mathbf{b}_2 + z\mathbf{b}_3.$$

The span of $\mathbf{b}_1$ and $\mathbf{b}_2$ is just the xy-plane (where $z = 0$).

2.1 Linear dependence and independence

A set of vectors $\mathbf{v}_1, \ldots, \mathbf{v_n}$ is linearly dependent if it is possible to make one of them as a linear combination of the others.

Example 4. The list of vectors

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 5 \\ 7 \end{pmatrix}$$

is a linearly dependent set of vectors because we can make v_3 as a linear combination $2\mathbf{v}_1 + \mathbf{v}_2$. In fact, in this example we can make any one of the vectors from the other two. We can plug this expression for $\mathbf{v}_3$ into any linear combination $\lambda_1\mathbf{v}_1 + \lambda\mathbf{v}_2 + \lambda\mathbf{v}_3$ to turn it into a linear combination of just $\mathbf{v}_1$ and $\mathbf{v}_2$,

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \lambda_3\mathbf{v}_3 = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \lambda_3(2\mathbf{v}_1 + \mathbf{v}_2) = (\lambda_1 + 2\lambda_3)\mathbf{v}_1 + (\lambda_2 + \lambda_3)\mathbf{v}_2.$$

Thus the span of all three is equal to the span of just the first two vectors in this example.

Example 5. For another example, consider

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}, \mathbf{v}_4 = \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}.$$

In this example, we can make $\mathbf{v}_4$ as a linear combination $\mathbf{v}_4 = \mathbf{v}_2 + \mathbf{v}_3$, and we can rearrange the equation to make either $\mathbf{v}_3$ or $\mathbf{v}_2$ out of the others. Note however that we cannot make $\mathbf{v}_1$ as a linear combination of the others.

As a special case, two vectors are linearly dependent if and only if one is scalar multiple of the other.

A set of vectors is linearly independent if we cannot make any one as a linear combination of the others. For example, if we delete $\mathbf{v}_4$ from the above set of vectors then we would have an independent set of vectors.

Here are three useful equivalent ways to think about dependence and independence:

1. A set of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ is independent if removing any one vector from the list makes the span smaller.

2. A set of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ is independent if the span gets bigger each time we add one. I.e., the span of $\{\mathbf{v}_1, \mathbf{v}_2\}$ is bigger than the span of $\mathbf{v}_1$, and the span of the first 3 is bigger than the span of the first two, and so on.

3. A set of vectors is independent if there each point in the span can be made as a linear combination in only one unique way.

Remark 1. To see why the third condition is equivalent: if there is some vector $\mathbf{u}$ that we can make as a linear combination in two different ways,

$$u = \lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n = \mu_1\mathbf{v}_1 + \cdots + \mu_n\mathbf{v}_n,$$

then the difference

$$(\lambda_1 - \mu_1)\mathbf{v}_1 + \cdots + (\lambda_n - \mu_n)\mathbf{v}_n = 0$$

gives us a nontrivial recipe for 0, and then we can use this to write one of the vectors $\mathbf{v}_i$ as a linear combination of the others as long as the coefficient $(\lambda_i - \mu_i)$ is not zero.)

Given a list of vectors that is dependent, we can always delete vectors from the list one at a time if they can be made as a linear combination of the others, until we arrive at a list of linearly independent vectors.

2.2 Determining if something is in the span

Suppose we have a vector $\mathbf{u}$ and we want to know if it is in the span of a list of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. I.e., we want to find scalars $\lambda_1, \ldots, \lambda_n$ such that

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n = \mathbf{u}.$$

When we write this out component by component, we get a system of linear equations in the variables $\lambda_1, \ldots, \lambda_n$. Either this system of equations has a solution, in which case $\mathbf{u}$ is in the span, or it does not have a solution and $\mathbf{u}$ is outside the span.

Example 6. Consider the vectors:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{u} = \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix}.$$

Is the $\mathbf{u}$ in the span of $\mathbf{v}_1$ and $\mathbf{v}_2$?

We write out the system of equations:

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix}.$$

and then try to solve for λ_1 and λ_2. Looking at the middle component, we see that $\lambda_2 = 4$. Now looking at the first component and plugging in $\lambda_2 = 4$, we have $\lambda_1 + 4 = 3$, so λ_1 must be -1. But then we find that these values don't work for the equation of the third component equation, $2\lambda_1 + 0\lambda_2 = 1$. Thus there is no solution, so $\mathbf{u}$ is not in the span.

2.3 Linear spaces: lines, planes, etc

Here is a somewhat abstract definition to start with: A <u>linear space</u> in $\mathbb{R}^n$ is something that is the span of a set of vectors, and its <u>dimension</u> is the size of the largest set of independent vectors you can find in it.

More concretely, a line in $\mathbb{R}^n$ is the span of a single (nonzero) vector.

If we take the span of two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ then there are two possibilities: either they are dependent (so $\mathbf{v}_2$ is a scalar multiple of $\mathbf{v}_1$ and the span is still a line), or they are independent and the span is something bigger.

A plane in $\mathbb{R}^n$ is what we get when we take the span of two independent vectors.

Important fact: In $\mathbb{R}^n$, you can find at most n independent vectors.

Thus you will never find a 4-dimensional space inside $\mathbb{R}^3$, etc.

2.4 Basis

A set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in $\mathbb{R}^n$ is a basis if they are linearly independent and they span all of $\mathbb{R}^n$. It turns out that this can only happen when the number of vectors is equal to the dimension n.

The most common example of a basis is the *standard basis*:

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \mathbf{b}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

(so $\mathbf{b}_i$ has a 1 in the i^{th} positions and zeros everywhere else).

This is certainly not the only example of a basis, and many tasks in geometry and data science involve finding a basis that is particularly nice for the problem you are looking at.

If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis, then any vector can be written in a unique way as a linear combination of these basis vectors:

$$\mathbf{u} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n.$$

We can think of the list of numbers λ_i as representing $\mathbf{u}$ relative to this basis.

3 Lengths, angles, and the dot product

3.1 The length of a vector

Pythagoras tells us that if we have a right triangle where the legs have length a and b and hypotenuse has length c, then

$$a^2 + b^2 = c^2.$$

We can use this to measure the lengths of vectors in the plane. Given a vector $\mathbf{v} = \begin{pmatrix} a \\ b \end{pmatrix}$, we have a right triangle by moving along the x-axis a distance a and the parallel to the y-axis a distance b. The hypotenuse is precisely our vector, and so Pythagoras says that the length or the vector, written $||\mathbf{v}||$, is $\sqrt{a^2 + b^2}$.

It follows that if we want to measure the length of a vector $= \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ in $\mathbb{R}^n$, then we should use the formula

$$||\mathbf{v}|| = \sqrt{v_1^2 + \cdots + v_n^2}.$$

3.2 Angles

Suppose we have two vectors $\mathbf{u}$ and $\mathbf{v}$. If they are linearly dependent then we say that the angle between them is either 0 or 180 degrees (depending on whether they point in the same or opposite directions). If they are independent then they span a plane and in that plane we can measure the angle between them. (Note that we can't really tell the difference between an angle θ and $-\theta$ or $360 - \theta$ since we haven't decided which side of the plane we should look at.)

3.3 The dot product

The dot product of two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ is the number we obtain by multiplying corresponding components of the vectors and summing these up. We write $\mathbf{u} \cdot \mathbf{v}$.

Example 7.

$$\begin{pmatrix} 5 \\ 2 \\ 9 \end{pmatrix} \cdot \begin{pmatrix} 7 \\ -1 \\ 6 \end{pmatrix} = (5)(7) + (2)(-1) + (9)(6) = 35 - 2 + 54 = 87.$$

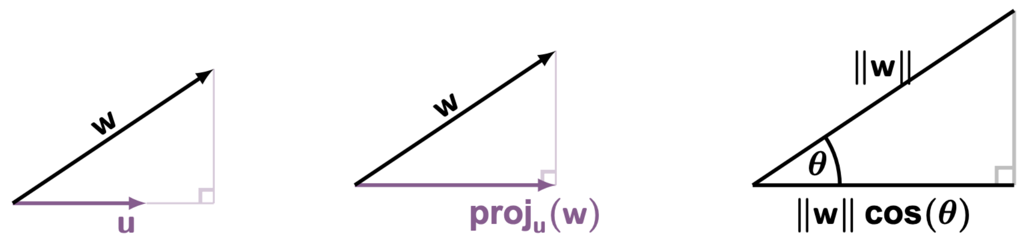
The dot product seems a bit weird, but it has a lovely geometric interpretation in terms of lengths and angles:

Fact: $\mathbf{u} \cdot \mathbf{v} = ||\mathbf{u}|| \, ||\mathbf{v}|| \cos\theta$, where θ is the angle between the two vectors.

Since the cosine of an angle is 0 if and only if the angle is a right angle (90 or -90 or 270), the dot product of two vectors is 0 if and only if they make a right angle.

3.4 Projection

Let $\mathbf{u}$ be a nonzero vector. There is a useful way to write any vector $\mathbf{w}$ as the sum of two pieces: one that points along the direction of $\mathbf{u}$, and one that is perpendicular to $\mathbf{u}$.



The piece of $\mathbf{w}$ that points along the direction of $\mathbf{u}$ is called the *projection* of $\mathbf{w}$ onto $\mathbf{u}$, and we write $\text{proj}_{\mathbf{u}}(\mathbf{w})$. This is a vector proportional to $\mathbf{u}$, and the length is $||\mathbf{w}|| \cos\theta = \frac{1}{||\mathbf{u}||}(\mathbf{w} \cdot \mathbf{u})$. We thus make $\mathbf{u}$ into the length we want by diving by the length of $\mathbf{u}$ and then multiplying by $\frac{1}{||\mathbf{u}||}(\mathbf{w} \cdot \mathbf{u})$. Thus

$$\text{proj}_{\mathbf{u}}(\mathbf{w}) = \left(\frac{\mathbf{u} \cdot \mathbf{w}}{\mathbf{u} \cdot \mathbf{u}}\right) \mathbf{u}.$$

The piece of $\mathbf{w}$ that is perpendicular to $\mathbf{u}$ is then what is left over: $\mathbf{w} - \text{proj}_{\mathbf{u}}(\mathbf{w})$.

4 Matrices

A matrix is a rectangular array of numbers (or variables). We say that a matrix A has shape $n \times k$ if it has n rows and k columns. For example

$$\begin{pmatrix} 2 & 4 & 5 \\ 1 & 0 & 9 \end{pmatrix}$$

has shape 2×3. Note that we can think of the vectors we've been working with as $n \times 1$ matrices. Given a matrix A, we will often write A_{ij} for the component in row i and column j. In the matrix above, we have $A_{12} = 4$ and $A_{21} = 1$.

There are 4 operations we often do with matrices:

- We can add two matrices if they have exactly the same shape. As with vectors, we simply add them component by component. We can express this as: $(A + B)_{ij} = A_{ij} + B_{ij}$.
- We can multiply a matrix A by a scalar λ in essentially same way as with vectors. Just multiply each component of A by λ. In terms of symbols, this is $(\lambda A)_{ij} = \lambda(A_{ij})$.

- You may encounter the <u>transpose</u> of a matrix, denoted A^T. This is simply the matrix whose rows are the columns of A. I.e., $(A^T)_{ij} = A_{ji}$.
- Matrix multiplication: this one is a little more complicated, so we'll devote the next section to it. To multiply AB we will require that the number of columns of A is equal to the number of rows of B.

4.1 Multiplying matrices

Suppose A has shape $n \times k$ and B has shape $k \times m$. Then we define the product AB by the rule: $(AB)_{ij}$ is the dot product of the i^{th} row of A with the j^{th} column of B.

Example 8. Consider the matrices

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \\ 10 & 6 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 4 \\ 4 & 1 & 5 \\ 3 & 1 & 1 \end{pmatrix}.$$

We have

$$AB = \begin{pmatrix} 18 & 7 & 17 \\ 21 & 9 & 29 \\ 37 & 27 & 71 \end{pmatrix}, \quad BA = \begin{pmatrix} 45 & 34 & 9 \\ 56 & 42 & 18 \\ 15 & 16 & 11 \end{pmatrix}.$$

Example 9. Consider the matrices

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 0 \\ 4 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Which multiplication is allowed: AB, BA, neither or both? Since A is 2×3 matrix and B is a 3×3 matrix, we are only allowed to compute AB. We find that

$$AB = \begin{pmatrix} 12 & 5 & 2 \\ 15 & 4 & 3 \end{pmatrix}.$$

4.2 Matrix times vector

A very important special case of matrix multiplication is when we multiply an $n \times m$ matrix A by a vector $\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$ of length m (which is the same as a matrix of shape $m \times 1$).

Important fact: $A\mathbf{v}$ is a linear combination of the columns of A. Take v_j times the j^{th} column of A and sum these up.

Explanation: If we let A_{*j} denote the j^{th} column of A, then the we can use the components of $\mathbf{v}$ to make a linear combination of the columns of A by

$$v_1 A_{*1} + \cdots + v_m A_{*m}$$

and the i^{th} component of this vector is $\sum_{j=1}^{m} v_j A_{ij}$, which is exactly the i^{th} component of $A\mathbf{v}$.

Example 10.

$$\begin{pmatrix} 2 & 3 & 5 \\ 1 & 9 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 2 \\ 1 \end{pmatrix} + y \begin{pmatrix} 3 \\ 9 \end{pmatrix} + z \begin{pmatrix} 5 \\ -2 \end{pmatrix}$$

5 Solutions to systems of linear equations

Now we will look at the connection between matrices, linear (in)dependence, and finding solutions to systems of linear equations. Suppose we have a list of unknowns $x_1, \ldots, x_n$. A linear equation is an equation that says a certain linear combination of the unknowns is equal to a certain number. E.g., $3x_1 + 2x_2 + 9x_3 = 1$. A system of linear equations is a list of linear equations:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

We can collect the unknowns x_i together into a vector $\mathbf{x}$. We can also collect the right hand side numbers b_i into a vector $\mathbf{b}$, and we collect the coefficients a_{ij} into a matrix A:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

The we can write the system of linear equations in the very compact form

$$A\mathbf{x} = \mathbf{b}$$

Example 11. Consider

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 9x_1 - x_2 &= 10 \end{aligned}$$

In terms of matrices, we can rewrite this as

$$\begin{pmatrix} 2 & 3 \\ 9 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 10 \end{pmatrix}$$

Recalling from Section 4.2 that a matrix times a vector gives us a linear combination of the columns of the matrix, trying to find a solution to an equation of the form

$$A\mathbf{x} = \mathbf{b}$$

is the same as trying to find a way to make $\mathbf{b}$ as a linear combination of the columns of A. There are 3 possibilities:

1. $\mathbf{b}$ is not in the span of the columns, so we can't make $\mathbf{b}$ as a linear combination of the columns and hence there is no solution.
2. $\mathbf{b}$ is in the span of the columns, and the columns are linearly independent, so there is a unique recipe to make $\mathbf{b}$ as a linear combination, and hence there is exactly one solution.
3. $\mathbf{b}$ is in the span of the columns, and the columns are linearly dependent. In this case, there are infinitely many solutions because there are infinitely many ways to make $\mathbf{b}$ as a linear combination of the columns.

6 Linear transformations

Definition 1. A function (we often say 'map' or 'mapping') $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a linear transformation if

- For any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ we have that

$$T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) \quad \textit{addition condition}.$$

- For any $\mathbf{u} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ we have that

$$T(\lambda\mathbf{u}) = \lambda T(\boldsymbol{u}) \quad \textit{scalar multiplication condition}.$$

A few notations we should remember when considering linear transformations:

- The *domain* of T is the set of allowed inputs for T. In the case above, this it $\mathbb{R}^n$.
- The *codomain* of T is the set in which the outputs of T live. In the case above, this is $\mathbb{R}^m$. It might be the case that there are vectors in the codomain that are not hit by T of anything.
- The *range* of T is the set of outputs that can actually be produced by T:

$$\text{Range}(T) = \{\mathbf{w} \in \mathbb{R}^m \mid \text{there exists} \mathbf{v} \in \mathbb{R}^n \text{ such that } T(\mathbf{v}) = \mathbf{w}\}.$$

  The range of a transformation is contained in its codomain.

Example 12. Find the domain, codomain of the following maps, and determine weather or not they are linear. In the case that they are, find Range (T)

a) $T\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}y\\x\end{pmatrix}$.

b) $T\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}y\\1\end{pmatrix}$.

c) $T(x)=\begin{pmatrix}x\\x^2\\2x\end{pmatrix}$.

Solution. a) T acts on vectors in $\mathbb{R}^2$ and returns vectors in $\mathbb{R}^2$. Thus, $\mathrm{Dom}(T)=\mathbb{R}^2$ and $\mathrm{Codom}(T)=\mathbb{R}^2$. To show linearity we check the addition and scalar multiplication conditions:
<u>addition condition</u> we have that

$$T\left(\begin{pmatrix}x_1\\y_1\end{pmatrix}+\begin{pmatrix}x_2\\y_2\end{pmatrix}\right)=T\begin{pmatrix}x_1+x_2\\y_1+y_2\end{pmatrix}=\begin{pmatrix}y_1+y_2\\x_1+x_2\end{pmatrix}$$

$$\underset{\text{separate the vectors}}{=}\begin{pmatrix}y_1\\x_1\end{pmatrix}+\begin{pmatrix}y_2\\x_2\end{pmatrix}=T\begin{pmatrix}x_1\\y_1\end{pmatrix}+T\begin{pmatrix}x_2\\y_2\end{pmatrix}.$$

<u>multiplication condition</u> we see that

$$T\left(\lambda\begin{pmatrix}x\\y\end{pmatrix}\right)=T\begin{pmatrix}\lambda x\\\lambda y\end{pmatrix}=\begin{pmatrix}\lambda y\\\lambda x\end{pmatrix}=\lambda\begin{pmatrix}y\\x\end{pmatrix}=\lambda T\begin{pmatrix}x\\y\end{pmatrix}.$$

Both conditions are valid, and as such T is a linear transformation. Let us find the range:
<u>Range (T)</u>: We'd like to find for which $u,v\in\mathbb{R}$ we can find $x,y\in\mathbb{R}$ such that

$$T\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}u\\v\end{pmatrix}.$$

As

$$T\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}y\\x\end{pmatrix}$$

we see that $u=y$ and $v=x$. This means that we can get *any vector*. Indeed

$$T\begin{pmatrix}v\\u\end{pmatrix}=\begin{pmatrix}u\\v\end{pmatrix},$$

showing that $\mathrm{Range}(T)=\mathbb{R}^2$.
There is a simpler way to find Range (T), which we will discuss shortly.

b) T acts on vectors in $\mathbb{R}^3$ and returns vectors in $\mathbb{R}^2$. Thus, $\mathrm{Dom}(T) = \mathbb{R}^3$ and $\mathrm{Codom}(T) = \mathbb{R}^2$. To show linearity we check the addition and scalar multiplication conditions:
addition condition we have that

$$T\left(\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}\right) = T\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix} = \begin{pmatrix} y_1 + y_2 \\ 1 \end{pmatrix}.$$

On the other hand

$$T\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + T\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ 1 \end{pmatrix} + \begin{pmatrix} y_2 \\ 1 \end{pmatrix} = \begin{pmatrix} y_1 + y_2 \\ 2 \end{pmatrix}.$$

The addition condition is not satisfied, so the map is not a linear transformation. Regardless, let is check the second condition.
multiplication condition we see that

$$T\left(\lambda \begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = T\begin{pmatrix} \lambda x \\ \lambda y \\ \lambda z \end{pmatrix} = \begin{pmatrix} \lambda y \\ 1 \end{pmatrix}.$$

On the other hand

$$\lambda T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \lambda \begin{pmatrix} y \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda y \\ \lambda \end{pmatrix}.$$

In order for the scalar multiplication condition to be satisfied we must have that $\lambda = 1$. This means that the scalar multiplication is not valid for all $\lambda \in \mathbb{R}$ - showing us again that the map can't be a linear transformation.

c) T acts on numbers in $\mathbb{R}$ and returns vectors in $\mathbb{R}^3$. Thus, $\mathrm{Dom}(T) = \mathbb{R}$ and $\mathrm{Codom}(T) = \mathbb{R}^3$. To show linearity we check the addition and scalar multiplication conditions:
addition condition we have that

$$T(x+y) = \begin{pmatrix} x+y \\ (x+y)^2 \\ 2(x+y) \end{pmatrix} = \begin{pmatrix} x+y \\ x^2 + 2xy + y^2 \\ 2(x+y) \end{pmatrix}.$$

On the other hand

$$T(x) + T(y) = \begin{pmatrix} x \\ x^2 \\ 2x \end{pmatrix} + \begin{pmatrix} y \\ y^2 \\ 2y \end{pmatrix} = \begin{pmatrix} x+y \\ x^2 + y^2 \\ 2(x+y) \end{pmatrix}.$$

We see that the addition condition is satisfied if and only if $2xy = 0$, i.e. $x = 0$ or $y = 0$. This means that it is not satisfied for all x and y in $\mathbb{R}$, showing that the map is not a linear transformation. Like before, let us check the scalar multiplication as well.

multiplication condition we see that

$$T(\lambda x) = \begin{pmatrix} \lambda x \\ (\lambda x)^2 \\ 2\lambda x \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda^2 x^2 \\ 2\lambda x \end{pmatrix}.$$

On the other hand

$$\lambda T(x) = \lambda \begin{pmatrix} x \\ x^2 \\ 2x \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda x^2 \\ 2\lambda x \end{pmatrix}.$$

In order for the scalar multiplication condition to be satisfied we must have that $\lambda x^2 = \lambda^2 x^2$, i.e. $x = 0$ or $\lambda = 0$ or $\lambda = 1$. This means that it is not satisfied for all $x \in \mathbb{R}$ and $\lambda \in \mathbb{R}$, showing (yet again) that the map is not a linear transformation. $\square$

While it may seem that linear transformation are detached from what we've learned so far, such maps are, in fact, intimately connected to matrices. This is expressed in the following theorem:

Theorem 6.1. *Any $n \times m$ matrix A determines a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$ by the simple formula $T(\mathbf{v}) = A\mathbf{v}$, and every linear transformation comes from a matrix in this way.*

Proof. You should check that the function defined by a matrix is a linear transformation. To show that every linear transformation comes from a matrix, we will start from a linear transformation T and build a matrix $[T]$ from it, and then show that $T(\mathbf{v})$ is equal to the product of the matrix $[T]$ times the vector $\mathbf{v}$.

To build $[T]$, we will need the list of of standard basis vectors

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \mathbf{b}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

(so $\mathbf{b}_i$ has a 1 in the i^{th} positions and zeros everywhere else). We apply T to these to get a new list of vectors $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)$. Now we make a matrix by putting $T(\mathbf{b}_i)$ as the i^{th} column.

Then we just need to check that the matrix multiplication $[T]\mathbf{v}$ gives the same result as $T(\mathbf{v})$ for any vector $\begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$. The matrix multiplication $[T]\mathbf{v}$ is the linear combination of the columns of $[T]$ with coefficients v_i, so this is

$$\sum_{i=1}^{m} v_i T(\mathbf{b}_i).$$

On the other hand, we can write $\mathbf{v} = \sum_{i=1}^{m} v_i \mathbf{b}_i$, and then we use the fact that T is a linear transformation to get

$$T(\mathbf{v}) = T\left(\sum_{i=1}^{m} v_i T(\mathbf{b}_i)\right) = \sum_{i=1}^{m} v_i T(\mathbf{b}_i). \qquad \square$$

Example 13. Consider the following map $T : \mathbb{R}^3 \to \mathbb{R}^2$:

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y \\ y+z \end{pmatrix}.$$

a) Show that the map is a linear transformation.

b) Find the matrix representation of T, $[T]$.

Solution. a) To show linearity we check the addition and scalar multiplication conditions:

addition condition we have that

$$T\left(\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}\right) = T\begin{pmatrix} x_1+x_2 \\ y_1+y_2 \\ z_1+z_2 \end{pmatrix} = \begin{pmatrix} (x_1+x_2)+(y_1+y_2) \\ (y_1+y_2)+(z_1+z_2) \end{pmatrix}$$

$$\underset{\text{separate the vectors}}{=} \begin{pmatrix} x_1+y_1 \\ y_1+z_1 \end{pmatrix} + \begin{pmatrix} x_2+y_2 \\ y_2+z_2 \end{pmatrix} = T\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + T\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}.$$

multiplication condition we see that

$$T\left(\lambda\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = T\begin{pmatrix} \lambda x \\ \lambda y \\ \lambda z \end{pmatrix} = \begin{pmatrix} \lambda x+\lambda y \\ \lambda y+\lambda z \end{pmatrix} = \begin{pmatrix} \lambda(x+y) \\ \lambda(y+z) \end{pmatrix} = \lambda\begin{pmatrix} x+y \\ y+z \end{pmatrix} = \lambda T\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Both conditions are valid, and as such T is a linear transformation.

b) We have that

$$T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and as such

$$[T] = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

$\square$

7 Determinants and inverses

coming soon... for now, see lecture slides.

The inverse of a square matrix A is a matrix A^{-1} such that the products AA^{-1} and $A^{-1}A$ are both equal to the identity matrix $\mathbb{1}$.

Finding the inverse of a matrix A is useful because it gives us a very quick way to write down solution to systems of equations $A\mathbf{x} = \mathbf{b}$. Just multiply both sides by A^{-1} and we get $\mathbf{x} = A^{-1}\mathbf{b}$. This shows us that, for a square system of equations (the number of equations equals the number of unknowns), there is a unique solution if and only if the matrix has an inverse.

Important fact: The inverse of A exists if and only if the determinant of A is nonzero.

8 Eigenvalues and eigenvectors

An eigenvector for a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is a nonzero vector $\mathbf{v}$ such that $T(\mathbf{v})$ is a scalar multiple of $\mathbf{v}$. I.e.,

$$T(\mathbf{v}) = \lambda\mathbf{v}$$

for some scalar λ which is called the corresponding eigenvalue. (We exclude the zero vector because it satisfies this trivially for any value of λ.) Notice that if $\mathbf{v}$ is an eigenvector, then any scalar multiple of $\mathbf{v}$ is also an eigenvector.

If T is a shear transformation of $\mathbb{R}^2$, then a vector pointing along the shearing axis is an eigenvector with eigenvalue 1. Not every linear transformation has eigenvectors. For instance, a rotation of $\mathbb{R}^2$ has none (unless we decide to work with complex numbers, but that is a story for another day...)

Suppose we have a linear transformation given by a matrix A and we want to find an eigenvector. We are looking for a solution to the equation

$$A\mathbf{v} = \lambda\mathbf{v}$$

where both the vector $\mathbf{v}$ and the scalar λ are unknown. Subtracting $\lambda\mathbf{v}$ from both sides, we can rearrange this as

$$(A - \lambda\mathbb{1})\mathbf{v} = 0.$$

If we forget around the fact that we don't know what λ is yet, this looks like a system of linear equations. It has one obvious solution: $\mathbf{v} = 0$. But that is not a very interesting solution because eigenvectors are supposed to be nonzero vectors. So, in order to find a nonzero $\mathbf{v}$, we need this system of equations to have infinitely many solutions. This happens if and only if $(A - \lambda\mathbb{1})$ does not have an inverse, which happens if and only if its determinant is 0.

Thus, the procedure to find an eigenvector is to first find the possible eigenvalues by solving the equation

$$\det(A - \lambda \mathbb{1}) = 0$$

for λ. Then, once you have a possible value of λ, you go back and look for a non-zero $\mathbf{v}$ that gives a solution to

$$(A - \lambda \mathbb{1})\mathbf{v} = 0.$$

Solving $\det(A - \lambda \mathbb{1}) = 0$ amounts to finding the roots of a polynomial $\lambda^n + \cdots = 0$. Unless the numbers happen to be very carefully selected, it is best to get a computer to do this. But in the 2×2 case we can easily do it by hand using the quadratic formula. Given a polynomial $a\lambda^2 + b\lambda + c$, the roots are given by the formula

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

The quantity $b^2 - 4ac$ has the fancy name *discriminant*.

1. If it is negative then we are taking the square root of a negative number and we are going to need to use complex numbers.
2. If it is zero then the choice of $\pm$ is irrelevant and there is only one root. In fact, the polynomial can be written as $a(\lambda + b/(2a))^2$.
3. If the discriminant is positive then there are two distinct roots, one given by taking the $\pm$ to be $+$, and the other given by choosing it to be $-$.

Example 14. Let us try to find the eigenvectors and eigenvalues of $A = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}$. First we compute

$$0 = \det(A - \lambda \mathbb{1}) = \det \begin{pmatrix} 3-\lambda & 2 \\ 2 & 3-\lambda \end{pmatrix} = (3-\lambda)^2 - 4 = \lambda^2 - 6\lambda + 5.$$

This is a quadratic equation, so we use the quadratic formula:

$$\lambda = \frac{6 \pm \sqrt{36 - 4 \cdot 5}}{2} = 3 \pm 2 = 1 \text{ or } 5.$$

For each of these possible values for λ, we can now look for an eigenvector.

Let us first take $\lambda = 5$. We want to find a nonzero solution to the equation

$$(A - 5\mathbb{1})\mathbf{v} = 0.$$

Writing this out, we have

$$\begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This is the system of linear equations

$$-2v_1 + 2v_2 = 0$$
$$2v_1 - 2v_2 = 0$$

The second equation is just -1 times the first, so if we find a solution to the first then it is automatically a solution to the second equation as well. The first equation tell us that $v_2 = v_1$, and nothing more. So $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is a solution, and any other solution is a scalar multiple of this. Thus we have found an eigenvector with eigenvalue 5.

If we instead take $\lambda = 1$, we arrive at the system of equations

$$2v_1 + 2v_2 = 0$$
$$2v_1 + 2v_2 = 0$$

and the solutions to this system are all the vectors in the span of $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.