Crash Course in Linear Algebra

Vector Basic[1]

Vector can be represent as

$$\vec{A} = A_1\hat{e}_1 + A_2\hat{e}_2 + \dots$$

Where $\hat{e}_1, \hat{e}_2, \dots$ are called **basis**. It just a fancy word for unit vector in which we use to define our vector.

Think about it as $\hat{x}$ and $\hat{y}$ you have seen in physics. We normally use $\hat{x}$ and $\hat{y}$ since it has two nice properties that

1. It is a **unit vector**; meaning that the length of it is 1.

$$\hat{e}_i \cdot \hat{e}_i = 1 \; \forall i$$

2. Each of the basis are **orthorgonal** to each other. This means

$$\hat{e}_i \cdot \hat{e}_j = 0 \text{ if } i \neq j$$

To sound more fancy the two can be written succintly using kronecker delta function as

$$\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$$

where

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

The basis that satisfy the two above equation is called **orthornormal** basis and it is the basis like this that allows us to do dot product very easily.

$$\begin{aligned} \vec{A} \cdot \vec{B} =& (A_1\hat{e}_1 + A_2\hat{e}_2) \cdot (B_1\hat{e}_1 + B_2\hat{e}_2) \\ =& A_1B_1(\hat{e}_1 \cdot \hat{e}_1) + A_1B_2(\hat{e}_1 \cdot \hat{e}_2) + A_2B_1(\hat{e}_2 \cdot \hat{e}_1) + A_2B_2(\hat{e}_2 \cdot \hat{e}_2) \\ =& A_1B_1 \times 1 + A_1B_2 \times 0 + A_2B_1 \times 0 + A_2B_2 \times 1 \\ =& A_1B_1 + A_2B_2 \end{aligned} \tag{1}$$

without this property dot product would be a bit more complicated.

[1] This is not a rigorous treatment for any of the concepts. For rigorous treatment, look elsewhere. What we are trying to do here is to get the intuition across.

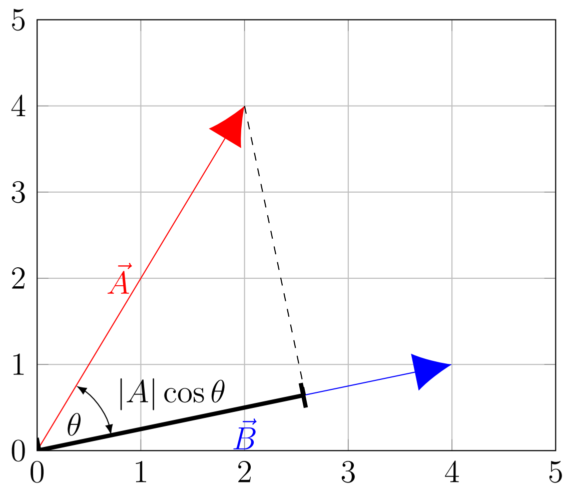
Dot Product is Projection Length

Let us get an intuitive feeling about the dot product. The first time you learn dot product it is defined as the product of the two length and the angle between the two. This is called geometrical definition of dot product.

$$\vec{A} \cdot \vec{B} = |\vec{A}||\vec{B}| \cos\theta$$

Typically when we have computer do calculation we normally use the algebraic definition of the dot product show in Equation 1. The two definition are equivalent. The proof is a bit tedious; if you are curious you can find it on Internet[2].

The geometrical definition of dot product has a very nice intepretation. **The dot product is the length of projection of one vector onto another**. Everytime you see a dot product I want you to have this automatically in your head and everytime you need to do a projection I want you to think dot product is what I need. Let us see why this is the case.



The picture above illustrate the following view of the definition

$$\vec{A} \cdot \vec{B} = \left(|\vec{A}| \cos\theta\right) |\vec{B}|$$

where $|\vec{A}| \cos\theta$ is just the length of projection of $\vec{A}$ onto $\vec{B}$. This basically means the length of $\vec{A}$ in the direction of $\vec{B}$.

This means that[3]

$$\vec{A} \cdot \vec{B} = (\text{length projection of } \vec{A} \text{ onto } \vec{B}) \times (\text{length of } \vec{B}) \tag{2}$$

The real use of this interpretation comes in when $\vec{B}$ is a unit vector like (eg. those basis vector $\hat{e}_1, \hat{e}_2, \ldots$). The above equation becomes

---

[2] `http://www.mit.edu/~hlb/StantonGrant/18.02/details/tex/lec1snip2-dotprod.pdf`

[3] You can do the same exercise and convice yourself that it means the other way too.

$$\vec{A} \cdot \vec{B} = (\text{length projection of } \vec{B} \text{ onto } \vec{A}) \times (\text{length of } \vec{A})$$

$$\vec{A} \cdot \hat{e}_i = \text{length of } \vec{A} \text{ in the direction of } \hat{e}_i \tag{3}$$

$$= \text{component of } \vec{A} \text{ in } \hat{e}_i \text{ direction} \tag{4}$$

$$= A_i \tag{5}$$

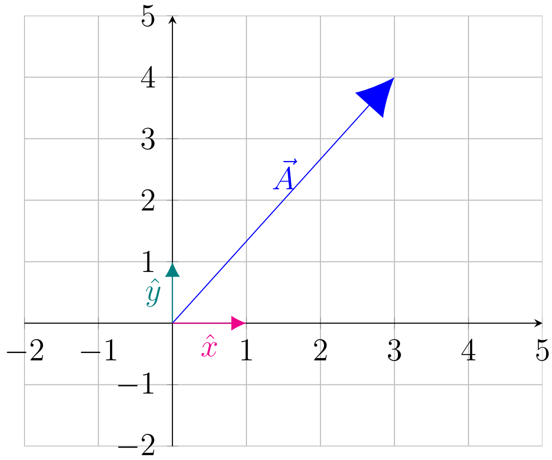
The last line is very important. It says that if we have a vector and we want to find the component of it for basis $\hat{e}_i$ all we need to do is to compute dot product.

One more useful information about the dot product is about orthorgonality(perpendicularity). When the two vectors are perpendicular the dot product is zero.

$$\vec{A} \cdot \vec{B} = |\vec{A}||\vec{B}| \cos 90° = 0$$

Vector Components and Vector Basis

Let us elaborate a little bit more on Equation 5. Consider the vector in the picture below



We normall **represent** this vector as

$$\vec{A} = 3\hat{x} + 4\hat{y}$$

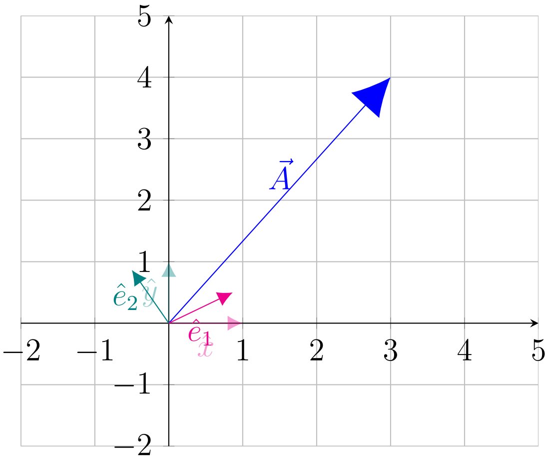
and the reason for that is the length of $\vec{A}$ in $\hat{x}$ direction is 3 and the length of $\vec{A}$ in $\hat{y}$ direction is 4. The length in each of basis direction are the **components** of $\vec{A}$. For example, the component of $\vec{A}$ in $\hat{x}$ direction is 3.

However, $3\hat{x} + 4\hat{y}$ is not the only way to represent $\vec{A}$. We can use other basis too. For example, let us consider another set of orthornormal basis $\{\hat{e}_1, \hat{e}_2\}$ which is a little bit tilted compared to standard $\hat{x}$ and $\hat{y}$.

$$\hat{e}_1 = \frac{\sqrt{3}}{2}\hat{x} + \frac{1}{2}\hat{y}$$

$$\hat{e}_2 = -\frac{1}{2}\hat{x} + \frac{\sqrt{3}}{2}\hat{y}$$

You can verify that these two are orthornormal basis. $\hat{e}_1$ and $\hat{e}_2$ along with $\vec{A}$ are shown in the figure below



Our goal is to represent $\vec{A}$ in terms of $\hat{e}_1$ and $\hat{e}_2$. That means we need to **find the component** of $\vec{A}$ along $\hat{e}_1$ and $\hat{e}_2$. This is very easy to do. All we need to do is to compute $A_1 = \vec{A} \cdot \hat{e}_1$ and $A_2 = \vec{A} \cdot \hat{e}_2$ which are[4]

$$A_1 = \vec{A} \cdot \hat{e}_1 = (3\hat{x} + 4\hat{y}) \cdot \left(\frac{\sqrt{3}}{2}\hat{x} + \frac{1}{2}\hat{y} \right) = \frac{3\sqrt{3}}{2} + 2 = 4.59 \tag{6}$$

$$A_2 = \vec{A} \cdot \hat{e}_2 = (3\hat{x} + 4\hat{y}) \cdot \left(-\frac{1}{2}\hat{x} + \frac{\sqrt{3}}{2}\hat{y} \right) = -\frac{3}{2} + 2\sqrt{3} = 1.96 \tag{7}$$

Thus,

$$\vec{A} = 4.59\hat{e}_1 - 1.96\hat{e}_2 = 3\hat{x} + 4\hat{y} \tag{8}$$

I put two equality in the same line for a very important reason. **Both of the representations are representing exactly the same object.** We are expressing exactly that same quantity in two different basis. One can go back and forth between the two. There is no information loss here. Later on we will find that some basis are more useful than another.

Matrix Multiplication as A Bunch of Dot Products

Here is a little known fact with matrix multiplication which will be useful for getting an insight when you see a complicated looking equation. First, we normally represent a vector using a column matrix of components. For example,

$$\vec{A} = 3\hat{e}_1 + 4\hat{e}_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}_{\{\hat{e}_1, \hat{e}_2\}}$$

The basis subscript for the matrix indicate that we use the component in that basis to write down the matrix. Normally this is clear on which basis we are talking about so people usually drop it. When it's unclear which basis we are talking about we will write it down explicitly.

Similarly, the transpose of the vector[5] can be represented by row matrix.

[4]Equation 1 work only if you express the two vectors in the exactly the same orthornormal basis in this case we express both vectors in terms of $\{\hat{x}, \hat{y}\}$

[5]Fancy name for it is dual vector.

$$\vec{A}^T = \begin{bmatrix} 3 & 4 \end{bmatrix}_{\{\hat{e}_1, \hat{e}_2\}}$$

This means the dot product can be written as the matrix multiplication of transpose and another matrix. For example, Let

$$\vec{A} = 3\hat{e}_1 + 4\hat{e}_2 \tag{9}$$
$$\vec{B} = 5\hat{e}_1 + 6\hat{e}_2 \tag{10}$$

The dot product of the two vectors $(3\hat{e}_1 + 4\hat{e}_2) \cdot (5\hat{e}_1 + 6\hat{e}_2)$ is given by

$$\vec{A} \cdot \vec{B} = A^T B \tag{11}$$
$$= \begin{bmatrix} 3 & 4 \end{bmatrix} \times \begin{bmatrix} 5 \\ 6 \end{bmatrix} \tag{12}$$
$$= [39] \tag{13}$$
$$= 39 \tag{14}$$

This should be viewed as the dot product and matrix multiplication is essentially the same thing. There is no reason we should stop at just 1. Let us consider a larger matrix multiplication.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \times \begin{bmatrix} 7 \\ 8 \end{bmatrix} = \begin{bmatrix} 7 \times 1 + 8 \times 2 \\ 7 \times 3 + 8 \times 4 \\ 7 \times 5 + 8 \times 6 \end{bmatrix} \tag{15}$$

You can see that each line looks a lot like a dot product that is

$$\begin{bmatrix} — & \vec{A}_1^T & — \\ — & \vec{A}_2^T & — \\ — & \vec{A}_3^T & — \\ & \vdots & \\ — & \vec{A}_n^T & — \end{bmatrix} \times \begin{bmatrix} | \\ \vec{B} \\ | \end{bmatrix} = \begin{bmatrix} \vec{A}_1 \cdot \vec{B} \\ \vec{A}_2 \cdot \vec{B} \\ \vec{A}_3 \cdot \vec{B} \\ \vdots \\ \vec{A}_n \cdot \vec{B} \end{bmatrix} \tag{16}$$

So, when you see matrix multiplication, think that it is just a bunch of dot product. This is also why np.dot in numpy does matrix multiplication.

Non Trivial Solution to System of Homogeneous Linear Equation

Let us consider the following system of equations

$$2x + 3y = 0 \tag{17}$$
$$4x + 2y = 0 \tag{18}$$

The equations above are linear and homogeneous. The homogeneous is from the 0 on the right handside. If you try to solve this equation you will find that the only solution is the trivial solution: $x = 0$ and $y = 0$. In fact most of the system of homogeneous linear equations has only trivial solution.

Let us consider another system of equations

$$1x + 2y = 0 \tag{19}$$
$$2x + 4y = 0 \tag{20}$$

Once you solve these equations you will find that there are infinitely many solutions. For example, $x = 2y = -1$ or $x = 4, y = -2$ will do. You may also notice that they are essentially the same equation.

The system equation can be represented with matrix equation; for example, the equatinos above can be written as

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \times \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or succintly

$$\mathbf{A}\vec{x} = 0$$

where **A** is a matrix.

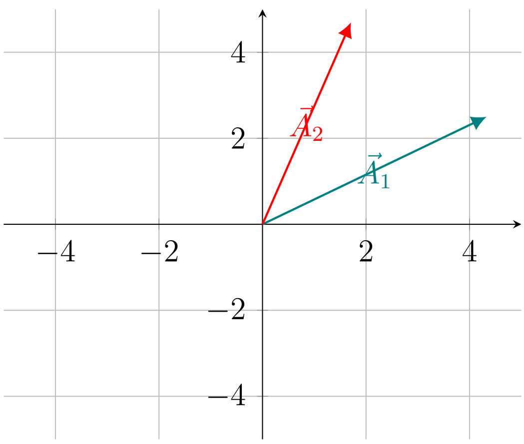
One can actually prove that these are the only two situations we can have for system of equations

- The only solution is trivial solution.
- Or there are infinitely many solutions.

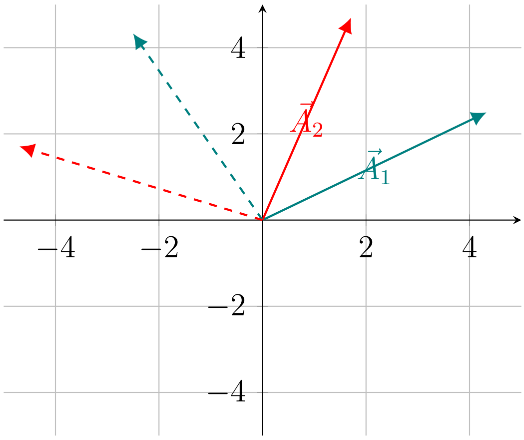
Our goal here is to find the condition for matrix **A** such that it has infinitely many solutions. This is easy if we view the matrix multiplication as a dot product.

$$\begin{bmatrix} — & \vec{A}_1^T & — \\ — & \vec{A}_2^T & — \\ — & \vec{A}_3^T & — \\ & \vdots & \\ — & \vec{A}_n^T & — \end{bmatrix} \times \begin{bmatrix} | \\ \vec{x} \\ | \end{bmatrix} = \begin{bmatrix} \vec{A}_1 \cdot \vec{x} \\ \vec{A}_2 \cdot \vec{x} \\ \vec{A}_3 \cdot \vec{x} \\ \vdots \\ \vec{A}_n \cdot \vec{x} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{21}$$

The above equation implies that for the equation to be true $\vec{x}$ has to be perpendicular to all the $\vec{A}_i$. Let us try to understand this for 2 dimension problem. Let $\vec{A}_1$ and $\vec{A}_2$ be two vector illustrated below.

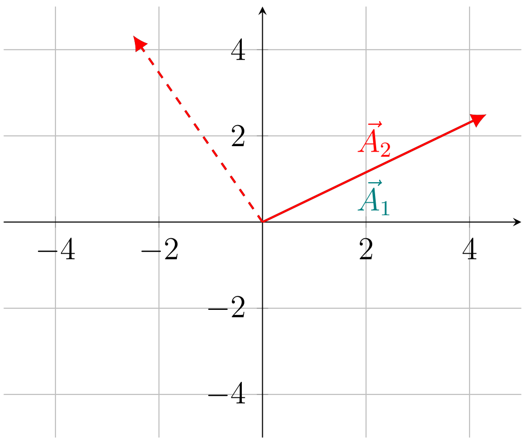


The system of equation $\mathbf{A}\vec{x} = 0$ requires us to find the vector $\vec{x}$ that is perpendicular to **both** $\vec{A}_1$ and $\vec{A}_2$ at the same time. Yet, the vector that is perpendicular to $\vec{A}_1$ and $\vec{A}_2$ do not point in the same direction as illustrated below.



In this case, the only vector that would satisfy $\mathbf{A}\vec{x} = 0$ would be $\vec{x} = 0$; the trivial solution.

So, to get nontrivial solution we need to get the dash teal and red line to point in the same direction. To do that, we need $\vec{A}_1$ and $\vec{A}_2$ to point in the same direction as illustrated below. In general as long as there is **no "volume/area"** cover by all the vector $\vec{A}_i$ then we can find at least one vector that satisfy the equation.



In this case any vector $\vec{x}$ along the dashed line would satisfy the equation.

Let look at what this means for matrix **A**. Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

This means that

$$\vec{A}_1 = a\hat{e}_1 + b\hat{e}_2$$
$$\vec{A}_2 = c\hat{e}_1 + d\hat{e}_2$$

Since we require that $\vec{A}_1$ and $\vec{A}_2$ to point in the same direction, this means

$$\vec{A}_1 = k\vec{A}_2 \; \exists k \neq 0$$

Thus we have two equation one for each direction.

$$a = kb \tag{22}$$
$$c = kd \tag{23}$$

Dividing the two we have

$$\frac{a}{c} = \frac{b}{d} \tag{24}$$

or

$$ad - cd = 0 \tag{25}$$

This is a very familiar looking expression. The left hand side is nothing but the determinant of matrix **A**. This is also true in higher dimension since the determinant is the same thing as "area/volume" up to a minus sign. So we have our condition that the equation

$$\mathbf{A}\vec{x} = 0$$

will have non trivial solution if and only if[6]

$$\det \mathbf{A} = 0 \tag{26}$$

This is probably the most important use of determinant. Let us see why this fact is useful.

Eigenvector and Eigenvalue

Let us start with a an observation. Given a square matrix

$$\mathbf{A} = \begin{bmatrix} 6 & 7 \\ 2 & 11 \end{bmatrix}$$

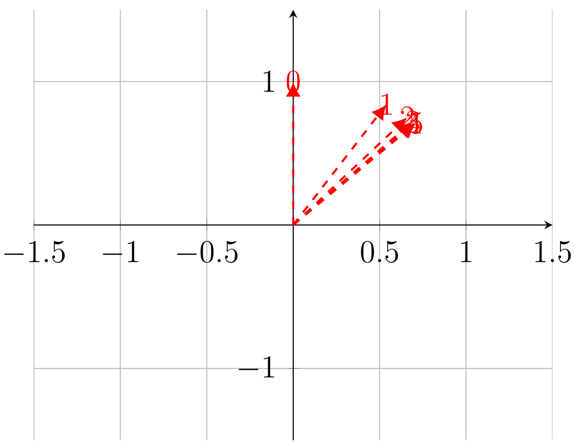
Let try to multiply it with $\vec{x} = (0, 1)$

$$\mathbf{A}\vec{x} = \begin{bmatrix} 7 \\ 11 \end{bmatrix}$$

Then we multiply the result again

$$\mathbf{A}\begin{bmatrix} 7 \\ 11 \end{bmatrix} = \dots$$

Then we keep doing this. We will find that there is a stationary direction which it it seems to get locked in place. The figure below shows the unit vector of the resulting vector after each iteration.



[6]for rigorous proof google up.

These stationary directions are called **Eigenvector**. Each matrix except in some special case has prefered direction. These fixed direction has many applications which we will learn later on. Let us learn how to find one by hand[7]

Let us consider the following matrix

$$\mathbf{A} = \begin{bmatrix} 6 & 7 \\ 2 & 11 \end{bmatrix}$$

if $\vec{x}$ is of a stationary direction then the resulting vector from multiplication would just be the same vector $\vec{x}$ multiply by some constant λ.

$$\mathbf{A}\vec{x} = \lambda\vec{x} \tag{27}$$

Let us insert an identity matrix on the right hand side.

$$\mathbf{A}\vec{x} = \lambda\mathbf{I}\vec{x}$$

Then let's move everything on the right hand side to the left. We have

$$\mathbf{A}\vec{x} - \lambda\mathbf{I}\vec{x} = 0$$

Factoring out $\vec{x}$ we have

$$(\mathbf{A} - \lambda\mathbf{I})\,\vec{x} = 0 \tag{28}$$

This is the same situation as system of homogeneous linear equation we seen before. We know for a fact that it has non trivial solution. This means that the determinant of the matrix on the left has to be 0. That means

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \det\left(\begin{bmatrix} 6-\lambda & 7 \\ 2 & 11-\lambda \end{bmatrix}\right) = 0$$

For 2×2 we can find the determinant by hands.

$$\begin{aligned} 0 = \det\left(\begin{bmatrix} 6-\lambda & 7 \\ 2 & 11-\lambda \end{bmatrix}\right) &= (6-\lambda)(11-\lambda) - 14 && (29) \\ &= 66 - 17\lambda + \lambda^2 - 14 && (30) \\ &= \lambda^2 - 17\lambda + 52 && (31) \\ &= (\lambda - 4)(\lambda - 13) && (32) \end{aligned}$$

Thus

$$\lambda \in \{4, 13\} \tag{33}$$

These are the only two values of λ that can satisfy Equation 27. These special value of λ are called **eigenvalues**[8]. With eigenvalue we can find corresponding eigenvector easily by pluggin it back into Equation 27. Let us first use $\lambda = 4$

[7] Normally you don't do that. We have computer for a reason.

[8] In general, eigenvalues do not have to be real value. It could be complex value.

$$\begin{bmatrix} 6 & 7 \\ 2 & 11 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 4 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{34}$$

Expanding it back to 2 linear equations. The first row gives

$$6x_1 + 7x_2 = 4x_1 \tag{35}$$
$$2x_1 + 7x_2 = 0 \tag{36}$$

and the second row gives

$$2x_1 + 11x_2 = 4x_2 \tag{37}$$
$$2x_1 + 7x_2 = 0 \tag{38}$$

You may notice that Equation 36 and Equation 38 are exactly the same equation. This is not a coincidence. If you recall back, the value of λ was determine so that the two row are pointing in the same direction. That means we can just take one.

Vector that satisfy Equation 36 is any multiple of

$$\vec{x}_{\lambda=4} = \begin{bmatrix} -7 \\ 2 \end{bmatrix} \tag{39}$$

That means that $(-7, 2)$ is an eigenvector corresponding to eigenvalue of $\lambda = 4$.

Let us find eigenvector corresponding to eigvenvalue of $\lambda = 13$. The equation we get will looklike

$$6x_1 + 7x_2 = 13x_1 \tag{40}$$
$$-7x_1 + 7x_2 = 0 \tag{41}$$

vector that satisfy the equation above is a multiple of

$$\vec{x}_{\lambda=13} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{42}$$

The reason for the word multiple of since the only information that matters is the direction not length. We can scale vector with any constant it will still point in the same direciton

In summary, the eigenvalues and eigenvector for matrix $\mathbf{A}$ are

$$[\lambda = 13, \vec{x}_{\lambda=13} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}], [\lambda = 4, \vec{x}_{\lambda=13} = \begin{bmatrix} -7 \\ 2 \end{bmatrix}],$$

So, whenever you see equation like

$$\mathbf{A}\vec{x} = k\vec{x},$$

it's just an eigenvalue problem.