Matrices and Linear Algebra: Solving $Ax = b$

We will discuss different ways to solve the same problem:

How do we find a solution for x_1 and x_2 that solves the following equation?

```
      A           x           b
   ----------------------------
   |  2   1 | . |  x₁ | = |   5 |
   | -2   1 |   |  x₂ |   |  -3 |

```

We will propose four different methods, and show how they relate to each other. These methods are:

1. System of Linear Equations (SLE)
2. Geometrical Interpretation
3. Gauss-Jordan Elimination (or Gaussian Elimination)
4. Inverse Matrix

Interesting enough, we will also be able to use each method in conjunction to the others. For example, we will see how to use SLE, Gaussian Elimination and geometrical interpretation to find the Inverse Matrix. Let's start!

1. System of Linear Equations (SLE)

Using matrix multiplication, we can rewrite the problem in terms of two linear equations:

```
   |  2   1 | . |  x₁ | = |   5 |    ==>   2x₁ + x₂ = 5
                |  x₂ |
```

and

```
   | -2   1 | . |  x₁ | = |  -3 |    ==>   -2x₁ + x₂ = -3
                |  x₂ |
```

That is, we need to find x_1 and x_2 that satisfy both equations:

```
   (i)   2x₁ + x₂ =  5
   (ii) -2x₁ + x₂ = -3
```

There are two common ways to solve this:

- Eliminating one variable from both equations

- Isolating and substituting one variable into the other equation

Let's see both:

a) Eliminating x_1

We know that adding a constant to both sides of an equation doesn't change the validity of the expression. So, let's add 5 to both sides `(ii)`:

```
(iii) -2x₁ + x₂ + 5 = -3 + 5
```

We also know from `(i)` that $2x_1 + x_2 = 5$. Let's use it to substitute in `(iii)`:

```
(iv) -2x₁ + x₂ + (2x₁ + x₂) = -3 + 5
```

Notice that we chose such values that would eliminate x_1 from the equation. This is equivalent of adding `(i)` to `(ii)`.

Simplifying both sides:

```
(v)   2x₂ = 2   ==>   x₂ = 1
```

Now that we know the value of x_2, we can substitute back in the initial equations and find x_1:

```
(vi)   2x₁ + 1 =  5    ==>  2x₁ =  4   ==>  x₁ =  2
(vii) -2x₁ + 1 = -3    ==> -2x₁ = -4   ==>  x₁ =  2
```

Notice that both result in the same solution, which is $x_1 = 2; x_2 = 1;$

b) Isolating and Substituting x_2

Since x_2 is the easiest variable to isolate, we will isolate it:

```
(i)   2x₁ + x₂ =  5
(ii) -2x₁ + x₂ = -3
```

From `(i)`:

```
(iii)   x₂ =  -2x₁ + 5
```

Substituting `(iii)` into `(ii)`:

```
(iv) -2x₁ + (-2x₁ + 5) = -3   ==>   -4x₁ = -8   ==>   x₁ = 2
```

Substituting `(iv)` back into `(iii)`:

```
(v)  x₂ =  -2(2) + 5 =  -4 + 5 = -1
```

As we expected, the solution is $x_1 = 2$; $x_2 = 1$;

2. Geometrical Interpretation

By plotting the problem into a Cartesian Plan, we can find the solution. Again, we will see two different methods to plot this problem, and solve both:

- Line equations (the row problem)
- Vectors (the column problem)

a) Line equations (the row problem)

The equations of our SLE can be plotted in the Cartesian plan as two lines:
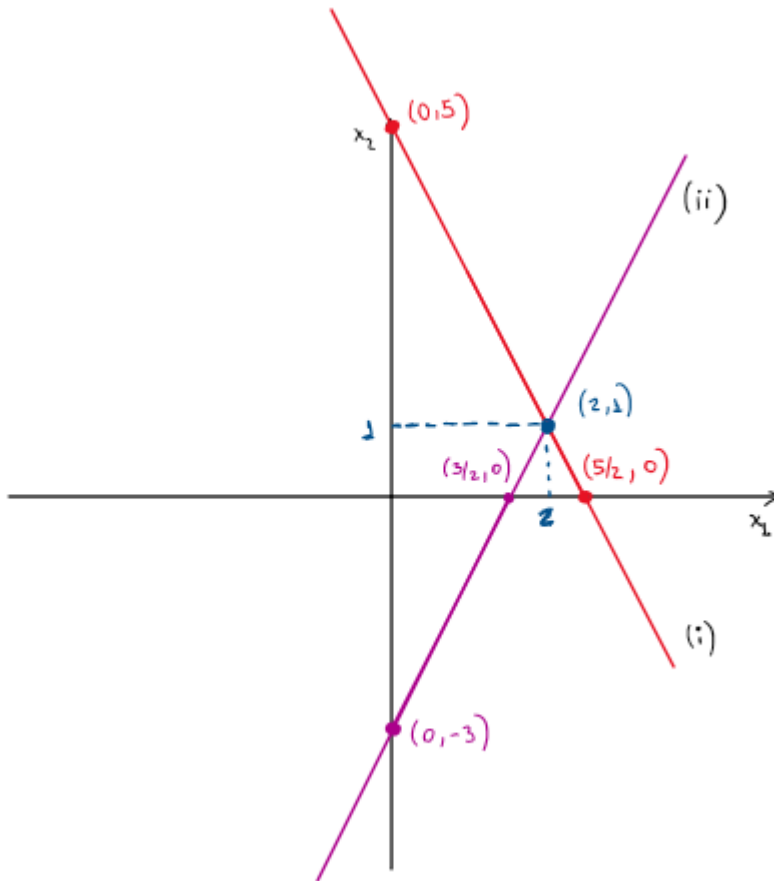
```
(i)   2x₁ + x₂ =  5
(ii) -2x₁ + x₂ = -3
```

One easy way to plot them is by finding the intercepts, that is, the value in which the line intercepts the axes:

For `(i)`:

```
When x₁ = 0, then x₂ =  5
When x₂ = 0, then 2x₁ = 5   ==> x₁ = 5 / 2
```

For `(ii)`:

```
When x₁ = 0, then   x₂ = -3
When x₂ = 0, then -2x₁ = -3   ==> x₁ = 3 / 2
```

From the image above, we see in red all points that satisfy `(i)`, and in purple all points that satisfy `(ii)`. We also see in blue the only point that satisfy both equations, `(2,1)`, which is the solution to our problem. We are going to discuss in the next section why we name this interpretation as the "row problem".

b) Vectors (the column problem)

Notice that in the previous method, we thought about our problem `Ax` = `b` in term of the rows of `A` and `b`:

```
(i)  |  2  1 |     ==>    2x₁ + x₂ =  5       | 5 |
     ---------                                ------
(ii) | -2  1 |     ==>   -2x₁ + x₂ = -3       | -3 |
```

For this next geometrical interpretation, we will look at the columns of `A` and `b`:

```
 c₁    c₂                       c₁           c₂          b
---------------------------------------------------------------
|  2 || 1 |     ==>    x₁ |  2 | + x₂ |  1 | = |  5 |
| -2 || 1 |     ==>       | -2 |      |  1 | = | -3 |
```

Now, instead of having a SLE, we have a linear combination of two vectors c_1 = $(2,-2)$ and c_2 = $(1,1)$, resulting in the vector b = $(5, -3)$. So, if by adding and scaling c_1 and c_2 we get the vector b, than the combinations x_1 and x_2 used to scale c_1 and c_2, respectively, are the solution to the problem.
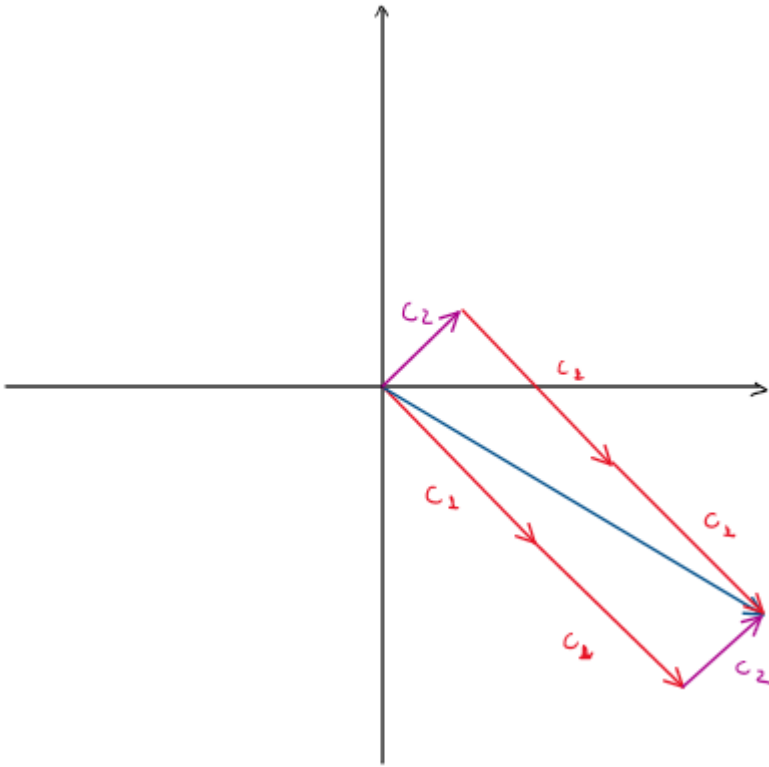
From the image above, we see in red the vector c_1, in purple the vector c_2 and in blue the vector b. It is also possible to notice that (only) when we add c_1 scaled by 2 and c_2 (scaled by 1), we get b. As expected, $x_1 = 2$ and $x_2 = 1$.

3) Gauss-Jordan Elimination

Using an *augmented matrix*, we can perform three elementary row operations to reduce the matrix to the *Reduced Row Echelon Form (RREF)*. These operations are:

a) Switch the position of two rows b) Multiply a row by any non-zero constant c) Add a scalar multiple of one row to any other row

```
(i)    |  2  1 |  5 |    1/2(i) --> a)
(ii)   | -2  1 | -3 |    1/2(ii) --> a)

(iii) |  1  1/2 |  5/2 |
(iv)  | -1  1/2 | -3/2 |    (i) + (ii) --> c)

(v)    |  1  1/2 |  5/2 |    (i) - 1/2(ii) --> c)
(vi)   |  0   1  |   1  |

(vii)  |  1  0 |  2 |    RREF(A)
(viii) |  0  1 |  1 |
```

From performing the same operations in both sides of the equation - that is, Ax and b - the valitidy of the expression holds true. So, we can rewrite the problem as:

```
| 1  0 | . |  x₁ | = |  2 |
| 0  1 |   |  x₂ |   |  1 |
```

Which is equivalent of $x_1 = 2; x_2 = 1$, as expected.

4) Inverse Matrix

Before learning multiple methods to find the *Inverse Matrix*, let's first understand what is and why do we want to find the *Inverse Matrix*.

Let's take one step back before we move forward:

a) Solving the problem with just 1 equation and 1 variable

How do we solve a linear equation with just one variable, x?

The most straight-forward way is by isolating the variable:

For example:

```
Solving Ax = b, where A = 3; b = 12;

3x = 12
3x / 3 = 12 / 3
x = 4
```

In (ii), we isolate x to find the solution. The way we do so is by diving both sides by A. In a more general solution:

```
Ax = b
Ax / A = b / A
x = b / A

for A, x and b are scalars, and A ≠ 0.
```

Now back to our original problem:

b) Solving the problem with 2 equations and 2 variables

```
    A           x          b
---------------------------
| 2  1 | . |  x₁ | = |  5 |
| -2  1 |   |  x₂ |   | -3 |
```

Could we try to solve this problem using the same method that we used in the previous section?

Let's rewrite the right side so there is a matrix and a vector:

```
| 2  1 | . | x₁ | = | 1  0 ||  5 |
| -2  1 |   | x₂ |   | 0  1 || -3 |
```

Now let's try to divide both sides by A:

```
| 2  1 | . | x₁ | = | 1  0 ||  5 |
| -2  1 |   | x₂ |   | 0  1 || -3 |            ┐(´•_•`)┌ WHAT❓❓
---------           ---------
| 2  1 |            | 2  1 |
| -2  1 |           | -2  1 |
```

And then:

```
| x₁ | = | 1  0 ||  5 |
| x₂ |   | 0  1 || -3 |                        ╲(0_0 )╱ Whaaat❓❓
         ---------
         | 2  1 |
         | -2  1 |
```

Can we solve it like this?

NO! We cannot! Division by matrix is not defined, so dividing A and I by A are not valid operations. However, not all is lost! If we find one (defined) operation that we could perform in both sides that isolates x, then we find the solution!

Idea: If we find any matrix M that when multiplied by A we get I, then we can perform the same operation in both sides of the equation to isolate x:

```
M . | 2  1 | . | x₁ | = M . | 1  0 | . |  5 |
    | -2  1 |   | x₂ |       | 0  1 |   | -3 |
```

In the left side we have M.Ax. By definition of our proposed M, :

```
         A           I
   -------------------------
M . | 2  1 |  = | 1  0 |
    | -2  1 |    | 0  1 |
```

In the right side, we have `M.I.b`. And by definition of `I`:

```
        I
   ----------------
M . |  1   0 | = M
    |  0   1 |
```

Therefore:

```
| 1  0 | . | x₁ | = M . |  5 |   ==>   | x₁ | = M. |  5 |
| 0  1 |   | x₂ |       | -3 |         | x₂ |       | -3 |
```

We managed to isolate `x`. By multiplying matrix `M` by the vector `b = (5,-3)`, we find x_1 and x_2, and solve our problem. Now, we just need to find `M`, if it exists.

c) Definition of the Inverse Matrix

You probably guessed by now that `M` is not any matrix. The *Inverse Matrix* of any given matrix `A` is the matrix that when multiplied by `A`, results in `I`.

Instead of calling it `M`, the typical notation of the inverse of `A` is `A⁻¹`, or `inv(A) = A⁻¹`. By definition, `A⁻¹. A = A . A⁻¹ = I`.

We will not discuss the several properties of an *Inverse Matrix* here. Instead, we will dive into methods for finding the inverse. The idea is that by solving and thinking of it differently, it will consolidate your mathematical intuition and reasoning about linear algebra, in general.

d) Finding the Inverse Matrix

Let's discover how to find the inverse matrix by using the following methods:

- SLE
- Gauss-Jordan
- Geometrical Interpretation (Matrix Composition)
- Cramer's Rule

This seems just as hard as solving the problem in the first place! Well, it is! However, once you find the inverse once, you can reuse it to find the solution to any vector `b`:

```
| x₁ | = A⁻¹ . | b₁ |
| x₂ |         | b₂ |

where b₁, b₂ are any real numbers.
```

So, there will be ocasions where solving a system of linear equations by finding an inverse matrix will be just too much work. But, in other cases, it will be a time saver! The trick is to be able to differentiate these occasions, and choose the best method to tackle your problem.

For that, let's learn some methods:

I) SLE

We can solve two SLE to find the inverse matrix in order of us to solve another SLE. It seems just not helpful at all, but as we mentioned, you can use the inverse for solving multiple SLE (and some other reasons).

Let's start with our definition `A . A⁻¹ = I`:

```
A⁻¹ . |  2  1 |  = |  1  0 |
       | -2  1 |     |  0  1 |
```

We know by the properties of an inverse matrix that `A⁻¹` has the same dimensions as `A`, that is, `A⁻¹` is a 2x2 matrix. So let's write `A⁻¹` using variables:

```
| a  b | . |  2  1 |  = |  1  0 |
| c  d |   | -2  1 |     |  0  1 |

 By performing the multiplication of matrices:

| 2a-2b  a+b |  = |  1  0 |
| 2c-2d  c+d |  = |  0  1 |
```

From the equality above:

```
(i)    2a - 2b = 1
(ii)    a +  b = 0

(iii) 2c - 2d = 0
(iv)    c +  d = 1
```

Using any methods explained in the first section, we solve for all variables:

```
(i) + 2(ii)  ==>   4a = 1   ==>   a = 1/4
(ii)  ==> 1/4 + b = 0   ==>   b = -1/4

(iii) + 2(iv)  ==>   4c = 2 ==> c = 1/2
(iv)  ==>  1/2 + d = 1   ==>   d = 1/2
```

Therefore:

```
A⁻¹ = | 1/4  -1/4 |
      | 1/2   1/2 |
```

As we know from solving SLE, we can either find one solution, infinite solutions or no solution at all. This means not always there will be an inverse.

- When there is no inverse, we say that the matrix is *singular*.
- In the example above, A has an inverse A⁻¹, therefore we say that A is *invertible*.

II) Gauss-Jordan

Using *Gauss-Jordan* to find an inverse is almost exactly the same process as explained to solve the SLE. The only difference is that instead of augmenting the matrix A by b, we will be augmenting it by I:

```
(i)    |  2  1 |  1  0 |   1/2(i)   --> b)
(ii)   | -2  1 |  0  1 |   1/2(ii)  --> b)

(iii) |  1  1/2 |  1/2  0 |
(iv)  | -1  1/2 |  0  1/2 |   (i) + (ii) --> c)

(v)   |  1  1/2 |  1/2   0 |   (i) - 1/2(ii) --> c)
(vi)  |  0   1  |  1/2  1/2 |

(vii) |  1  0  |  1/4  -1/4 |   RREF(A)
(viii)|  0  1  |  1/2   1/2 |
```

Notice that the steps we used to reduce A to RREF were exactly the same as before, what changed was only what was in the right side of the *augmented matrix*. Well, this makes sense:

```
Starting from:

    A    .    x    =    I    .    b
---------------------------------------
|  2  1 | . |  x₁ | = |  1  0 | . |  5 |
| -2  1 |   |  x₂ |   |  0  1 |   | -3 |
```

and performing the steps we performed above, it results in:

```
    I    .    x    =       A⁻¹       .    b
----------------------------------------------------------------
|  1  0 | . |  x₁ | = |  1/4  -1/4 | . |  5 | = |  8/4 | = |  2 |
|  0  1 |   |  x₂ |   |  1/2   1/2 |   | -3 |   |  2/2 |   |  1 |
```

Which is $x_1 = 2; x_2 = 1$, exactly the solution we found before.

III) Geometrical Interpretation (Matrix Composition)

We can represent the vectors of matrix A in a cartesian plan.
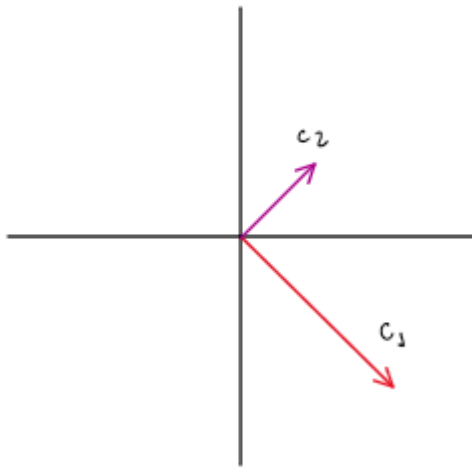
The objective is to find the inverse matrix, which would transform A into I. So, let's try to find a composition of simple matrices that does the same. Luckily, in this example it will be easy.

First thing, let's verify that the vectors $c_1 = (2,-2)$ and $c_2 = (1,1)$ are orthogonal. We can use the *dot product* to do it.

As we know, the *dot product* of orthogonal vectors is 0:

```
|  2 | . |  1 | = 2 - 2 = 0  ==>  orthogonal
| -2 |   |  1 |
```

We can also verify what is the angle between c_1 and the desired position that we want to place c_1, which is $e_1 = (1,0)$.

We will also use the geometrical and the algebraic definitions of the *dot product* for that.

Dot Product

```
Algebraic Def
-------------
|  2 | . |  1 | = 2.1 + (-2).0 = 2
| -2 |   |  0 |
```

So, $c_1.e_1 = 2$.

By the geometrical definition, we find:

```
Geometrical Def
-----------
```

```
c₁.e₁ = ||c₁|| ||e₁|| cosθ
```

So finding the length of c_1 and e_1 will solve for $\cos\theta$. This can also be done with the *dot product*.

We know that the *dot product* of a vector with itself is equal to the squared length of the vector:

```
||c₁||² = |  2 | . |  2 | = 2.2 + (-2).(-2) = 8
          | -2 |   | -2 |

||c₁|| = √8

||e₁||² = |  1 | . |  1 | = 1
          |  0 |   |  0 |

||c₁|| = √1 = 1
```

So, $||c_1|| = \sqrt{8}$; $||e_1|| = 1$.

Now we can use this values to find $\cos\theta$:

```
c₁.e₁ = = 2 = √8 cosθ  ==>  cosθ = 2/√8 = √2/2

θ = 45°
```

So, $\theta = 45°$.

We could verify also that the angle between c_2 and $e_2 = (0,1)$ is 45°, but we don't need to. Since both pair of vectors are orthogonal, we know that the angle between them will be the same.

Being so, we will rotate c_1 and c_2 by 45° anti-clockwise. Let's review the formula for a rotation matrix:

```
Rotation Matrix
----------------
| cosθ -sinθ |
| sinθ  cosθ |
```

We can deduct the formula using the *unit circle*:
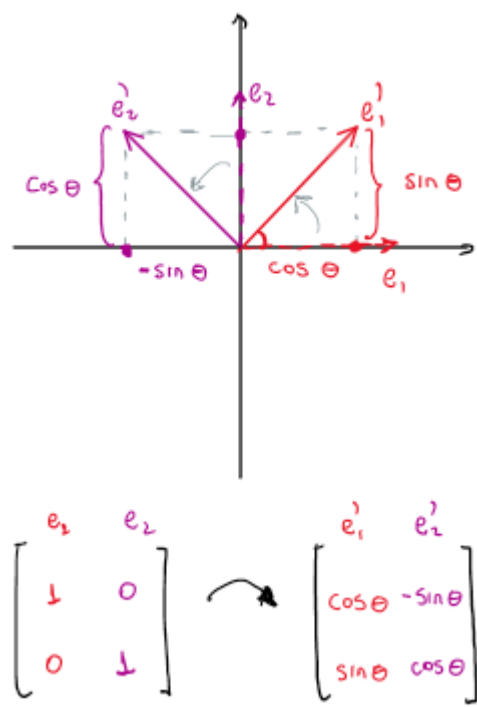
Therefore, for $\theta = 45°$ we get:

```
| √2/2   -√2/2 |
| √2/2    √2/2 |
```

This is the result of our rotation:
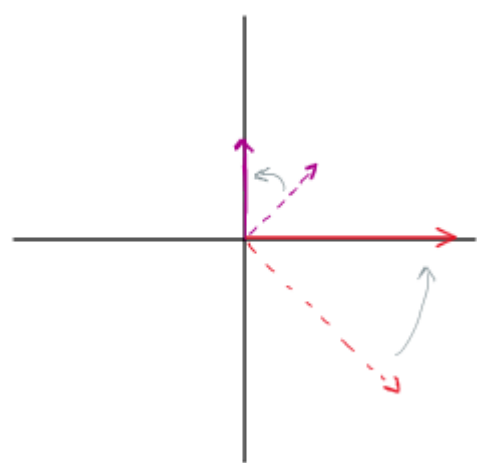
Now, we need to transform c_1 and c_2 into *Unit Vectors*. As we know, *Unit Vectors* have length equal 1. The way to to do it is by scaling each vector by $1/||c_1||$ and $1/||c_2||$, respectively.

Let's find $||c_1||$ and $||c_2||$:

```
C₁ = |  2 |    C₂ = |  1 |
     | -2 |         |  1 |

Using the dot product:
```

```
||c₁||² = |  2 | . |  2 | = 2.2 + (-2).(-2) = 8
           | -2 |   | -2 |

||c₂||² = |  1 | . |  1 | = 1 + 1 = 2
           |  1 |   |  1 |

Or alternatively by using Pythagoras:

          _____
||c₁|| = √2² + (-2)² = √8

          _____
||c₂|| = √1² + 1² = √2
```

Simplifying:

```
||c₁|| = 2√2
||c₂|| = √2

1/||c₁|| = 1 / 2√2 = √2/4
1/||c₂|| = 1 /  √2 = √2/2
```

The transformation matrix that scales c_1 and c_2 by $\sqrt{2}/4$ and $\sqrt{2}/2$, respectively, is:

```
| √2/4    0  |
|   0   √2/2 |
```

This is the result of our scaling:
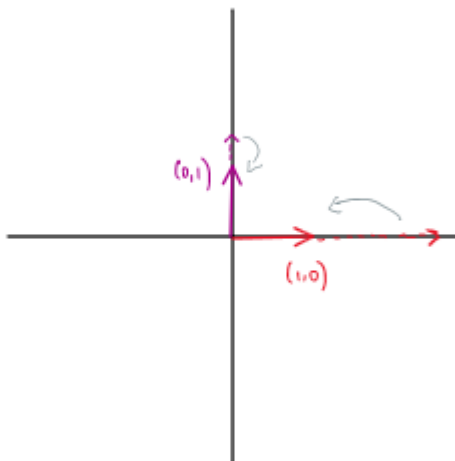
As we intended, we found a matrix composition that when applied to A transform it into I.

It is:

```
| √2/4    0  | . | √2/2  -√2/2 |
|   0   √2/2 | . | √2/2   √2/2 |
```

If so, the result of the multiplication of both matrices is equal to `inv(A)`.

Let's verify it by performing the multiplication:

```
| √2/4   0  | . | √2/2  -√2/2 | = | √2/4 . √2/2     √2/4 .-√2/2 | = | 1/4  -1/4 |
|   0  √2/2 | . | √2/2   √2/2 |   | √2/2 . √2/2     √2/2 . √2/2 |   | 1/2   1/2 |
```

Which we know from the previous methods that it is `A⁻¹`.

IV) Cramer's Rule

Cramer's Rule is not very intuitive and requires remembering so many steps that seems to me not very helpful. I will only briefly comment the general formula and show how we can easily find the inverse for 2x2 matrices. I will try also to discuss whichever intuition we can find by this method:

Cramer's Rule

```
A⁻¹ =  adj(A) / det(A),
```

Where `adj(A)` is the *adjugate matrix* and `det(A)` is the *determinant*.

The *adjugate matrix* is the *transpose* of its *cofactor matrix*. The *transpose* of a matrix is achieved by swapping its columns into rows, and its rows into columns.

The *cofactor matrix* is achieved by finding the n^2 *determinants* of smaller square matrices obtained by removing the respective row and column of each element, and multiplying by the correct signal (`+1` or `-1`).

Let's find these:

```
Cofactor Matrix of A - C(A)
--------------------
A₁₁ = | ■    ■  |        A₁₂ = |   ■     ■ |
      | ■  det(1) |             | det(-2) ■ |

A₂₁ = | ■  det(1) |        A₂₂ = | det(2)  ■ |
      | ■    ■  |               |   ■     ■ |
```

```
Signal = -1 if i+j is odd   ==>   | + - |
         +1 if i+j is even         | - + |
```

```
C(A) = |  1  2 |    Transpose(C(A)) = |  1 -1 |
       | -1  2 |                      |  2  2 |
```

```
det(A) = 1.2 - (-2).1 = 2 + 2 = 4
```

So, our inverse matrix will be:

```
A⁻¹ = 1/4 . |  1 -1 | = |  1/4 -1/4 |
            |  2  2 |   |  1/2  1/2 |
```

Just as we found before.

There are only two things I think it's worth mentioning about this method:

1. Since the determinant is geometrically interpreted as the scaling of the space, it makes sense that the inverse matrix is divided by the determinant. If A expands the space by 4, the inverse will shrink the space to 1/4.
2. Memorizing the shortcut to solve a 2x2 matrix can be helpful to quickly verify the inverse:
    - swap the elements A_{11} and A_{22}
    - change the signal from A_{12} and A_{21}
    - divide by the determinant

```
A = |  2  1 |    ==>  swap A₁₁ and A₂₂            ==> | 1/4 -1/4 | = | 1/4 -1/4 |
    | -2  1 |         change signal A₁₂ and A₂₁       | 2/4  2/4 |   | 1/2  1/2 |
                      divide by det(A)
```