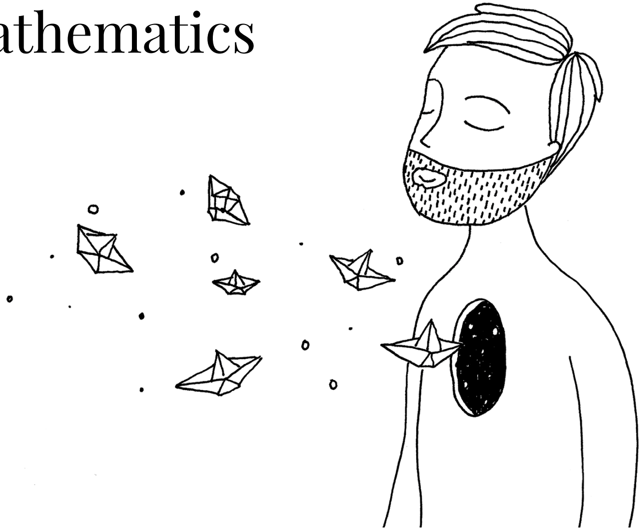
4509 - Bridging Mathematics

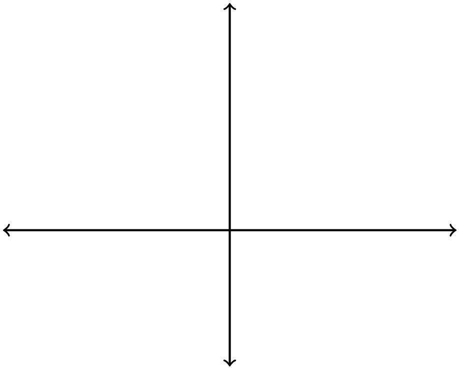
Matrices



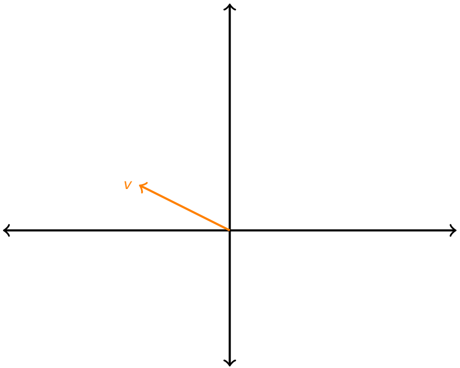
PAULO FAGANDINI

NOVA SCHOOL OF BUSINESS & ECONOMICS

Introduction

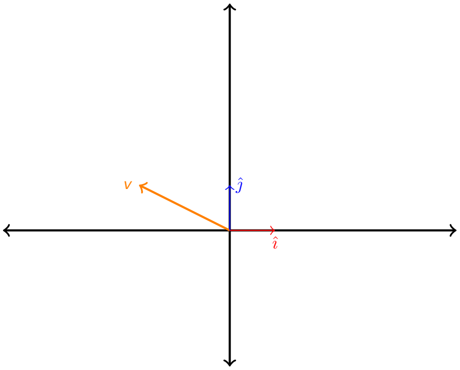


Introduction



$$v = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

Introduction



$$v = -2 \times \hat{\imath} + 1 \times \hat{\jmath} = -2 \times \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \times \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

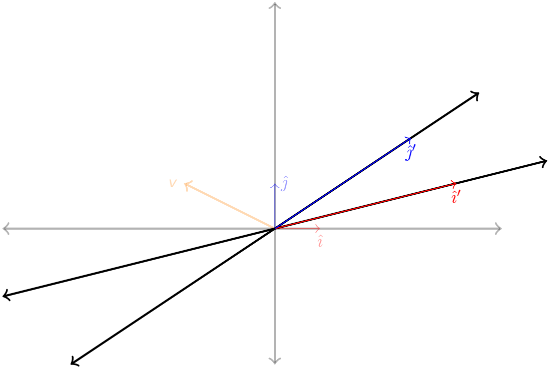
Introduction

Being not very rigorous, we can define a linear transformation as a transformation on every vector on the plane that must satisfy 2 things:

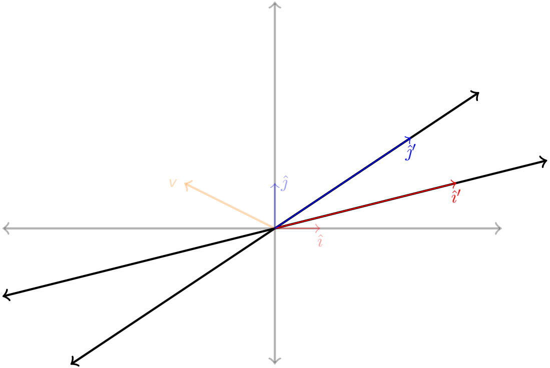
1. Lines must be transformed into lines
2. The origin must remain in the same place

We will deal with the formal definition and rigor later...

Introduction

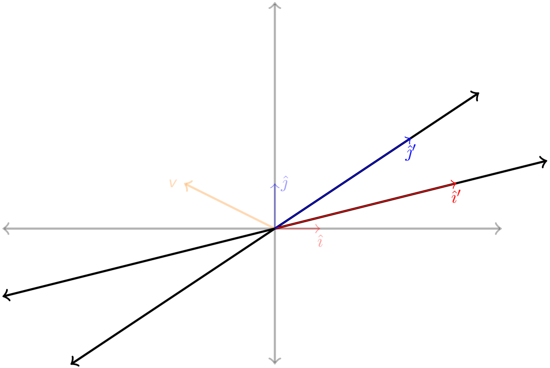


Introduction



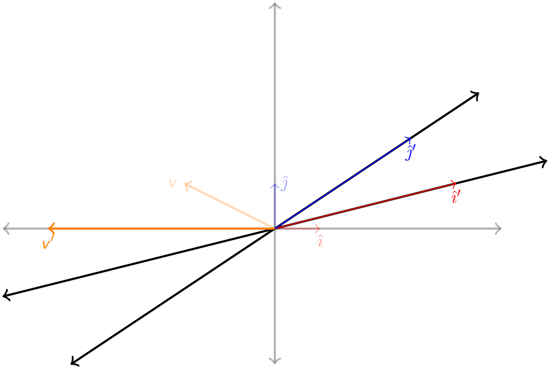
$$\hat{\imath}' = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \quad \hat{\jmath}' = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Introduction

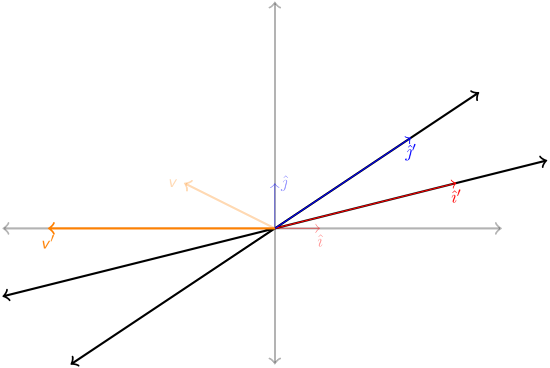


$$v = -2 \times \hat{i}' + 1 \times \hat{j}' = -2 \times \begin{bmatrix} 4 \\ 1 \end{bmatrix} + 1 \times \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -5 \\ 0 \end{bmatrix}$$

Introduction

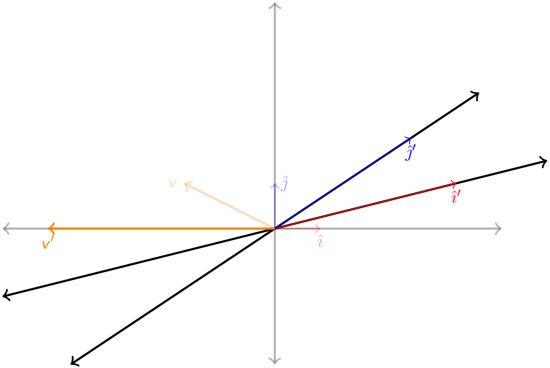


Introduction



$$w = \begin{bmatrix} x \\ y \end{bmatrix} \text{ lands on } x \times \begin{bmatrix} 4 \\ 1 \end{bmatrix} + y \times \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4x + 3y \\ 1x + 2y \end{bmatrix}$$

Introduction



$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} = x \times \hat{\imath} + y \times \hat{\jmath}$$

$$\begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4x + 3y \\ 1x + 2y \end{bmatrix} = x \times \hat{\imath}' + y \times \hat{\jmath}'$$

Introduction

What about another vector in the same "direction" than v? say

$$z = \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}$$

...

Introduction

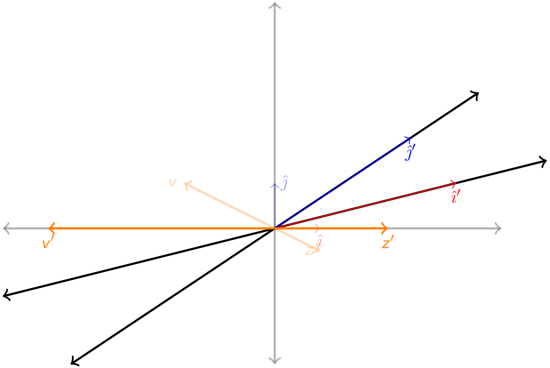
What about another vector in the same "direction" than v? say

$$z = \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}$$

...

$$\begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -0.5 \end{bmatrix} = \begin{bmatrix} 2.5 \\ 0 \end{bmatrix}$$

Introduction



Introduction

So transforming two vectors in the same line, they both end up also in the same line... keep this in mind.

Introduction

Could we take back $\hat{\imath}'$ to $\hat{\imath}$ and $\hat{\jmath}'$ to $\hat{\jmath}$?

Introduction

Could we take back $\hat{\imath}'$ to $\hat{\imath}$ and $\hat{\jmath}'$ to $\hat{\jmath}$?
Sure, if we apply the transform

$$\begin{bmatrix} 0.4 & -0.6 \\ -0.2 & 0.8 \end{bmatrix}$$

Introduction

Could we take back $\hat{\imath}'$ to $\hat{\imath}$ and $\hat{\jmath}'$ to $\hat{\jmath}$?
Sure, if we apply the transform

$$\begin{bmatrix} 0.4 & -0.6 \\ -0.2 & 0.8 \end{bmatrix}$$

$$\begin{bmatrix} 0.4 & -0.6 \\ -0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Applying the inverse!

Introduction

The important thing is that: if the vectors are linearly dependent, then we cannot invert the matrix, we just saw that two vectors that reside on the same line, end up in the same (although probably a different one) line.

Introduction

There are a couple of interesting vectors on the whole space when we apply this linear transformation...

Take for example the following vector: $e = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

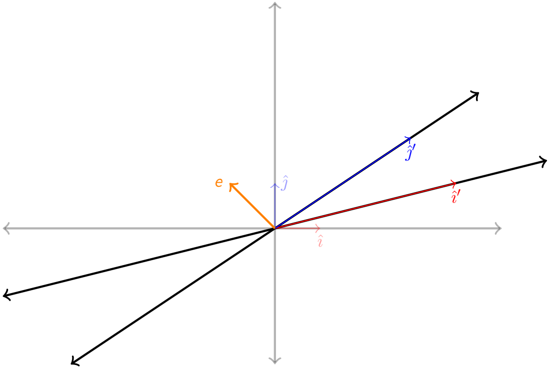
Introduction

There are a couple of interesting vectors on the whole space when we apply this linear transformation...

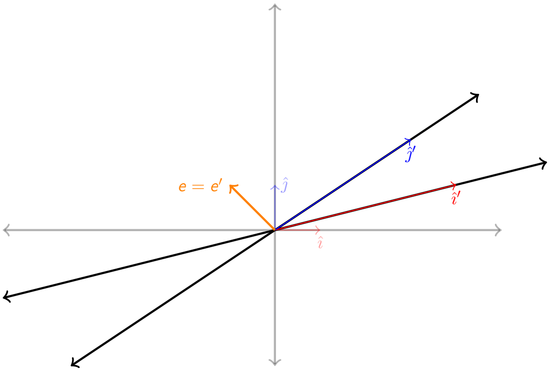
Take for example the following vector: $e = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Introduction

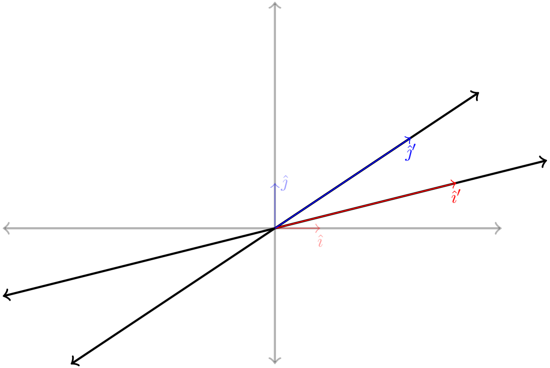


Introduction



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

Introduction



$$e.v._1 = \left[-1, 1\right], \quad \lambda_1 = 1$$

Introduction

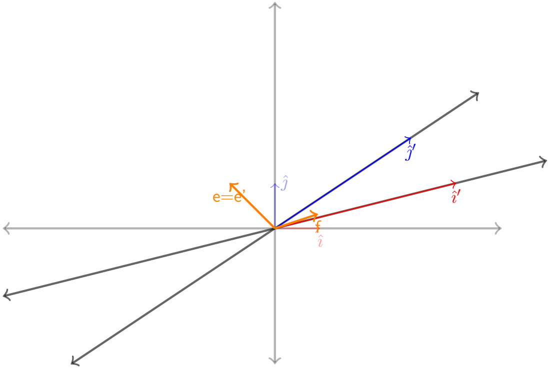
Or the vector: $f = \begin{bmatrix} 0.9487 \\ 0.3162 \end{bmatrix}$

Introduction

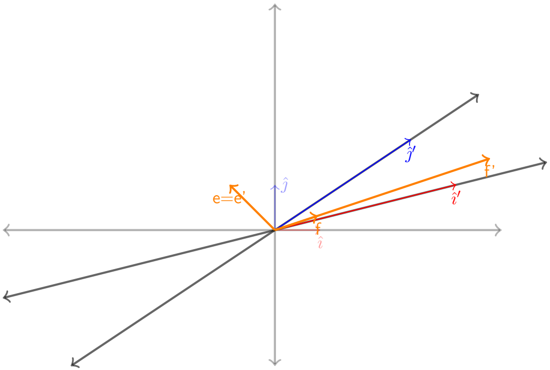
Or the vector:$f = \begin{bmatrix} 0.9487 \\ 0.3162 \end{bmatrix}$

$$\begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0.9487 \\ 0.3162 \end{bmatrix} = \begin{bmatrix} 4.7434 \\ 1.5811 \end{bmatrix}$$

Introduction

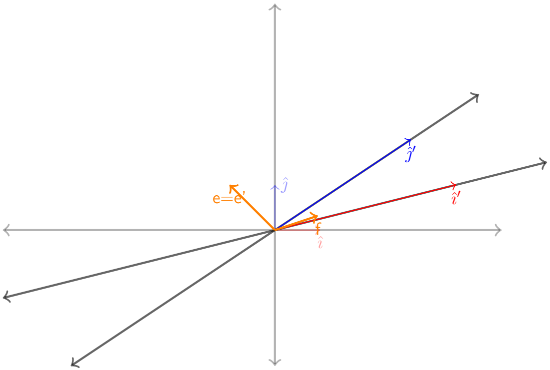


Introduction



$$5 \times \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.9487 \\ 0.3162 \end{bmatrix} = \begin{bmatrix} 4.7434 \\ 1.5811 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0.9487 \\ 0.3162 \end{bmatrix}$$

Introduction



$$e.v._2 = \begin{bmatrix} 0.9487 \\ 0.3162 \end{bmatrix}, \quad \lambda_2 = 5$$

Introduction

What happens with these vectors?

	ev_1	ev_2
$A\times$	$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 4.7434 \\ 1.5811 \end{bmatrix}$

Introduction

What happens with these vectors?

	ev_1	ev_2
$A\times$	$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 4.7434 \\ 1.5811 \end{bmatrix}$
$A^2\times$	$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 23.7171 \\ 7.9057 \end{bmatrix}$

Introduction

What happens with these vectors?

	ev_1	ev_2
$A\times$	$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 4.7434 \\ 1.5811 \end{bmatrix}$
$A^2\times$	$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 23.7171 \\ 7.9057 \end{bmatrix}$
$A^3\times$	$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 118.585 \\ 39.528 \end{bmatrix}$

Introduction

What happens with these vectors?

	ev_1	ev_2
$A\times$	$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 4.7434 \\ 1.5811 \end{bmatrix}$
$A^2\times$	$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 23.7171 \\ 7.9057 \end{bmatrix}$
$A^3\times$	$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 118.585 \\ 39.528 \end{bmatrix}$
λ	1	5

Definition

A real **matrix** is a rectangular array of real numbers.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{nm} \end{pmatrix}$$

Where $a_{ij} \in \mathbb{R}$. A is said to be an element of $\mathbb{R}^{m \times n}$

A vector would be then a matrix with only 1 column!

Let $A, B \in \mathbb{R}^{m \times n}$. Let $C \in \mathbb{R}^{n \times l}$. Finally, let $\alpha \in \mathbb{R}$.

1. $[A + B]_{ij} = a_{ij} + b_{ij}$
2. $[A \cdot C]_{ik} = \sum_{j=1}^{n} a_{ij} \cdot c_{jk}$, and it has a dimension $m \times l$
3. $[\alpha A]_{ij} = \alpha a_{ij}$

Definition

Let $A \in \mathbb{R}^{m \times n}$, A's **transpose**, denoted $A^t \in \mathbb{R}^{n \times m}$ is such that its elements are:

$$a_{ij}^t = a_{ji}$$

Definition

Matrix $A \in \mathbb{R}^{m \times n}$ is said to be **squared** if $n = m$

Definition

Matrix A is said to be **symmetric** if $A^t = A$

Definition

Matrix A is said to be **antisymmetric** if $A^t = -A$

Definition

The **Identity** is a squared matrix $I_n \in \mathbb{R}^{n \times n}$ that has $I_{ij} = 0$ if $i \neq j$, and $I_{ij} = 1$ if $i = j$.

The identity has a nice property: $AI_n = I_m A = A$ for any $A \in \mathbb{R}^{m \times n}$.

Definition

Matrix A is **invertible**, if there is another matrix A^{-1} such that $A \cdot A^{-1} = A^{-1} \cdot A = I$

Conjecture

Given $A, B, C \in \mathbb{R}^{n\times n}$

1. $A + B = B + A$
2. $A(BC) = (AB)C$
3. $A(B + C) = AB + AC$
4. $(A + B)^t = A^t + B^t$
5. $(AB)^t = B^t A^t$
6. $(A^t)^t = A$
7. *If* A *and* B *are invertible, then* AB *and* BA *are invertible as well. Furthermore* $(AB)^{-1} = B^{-1}A^{-1}$
8. *If* A *is invertible, then* $(A^t)^{-1} = (A^{-1})^t$

Quick quiz, 15 min, prove points 7 and 8. You can use points 1-6 as true and given.

Solution

7 Start with AB, multiply by A^{-1} from the left, you are left with $A^{-1}AB = IdB = B$. Now multiply by B^{-1}, so you get $B^{-1}A^{-1}AB = B^{-1}IdB = B^{-1}B = Id$. Then $(B^{-1}A^{-1})(AB) = Id$ so it must be that $B^{-1}A^{-1} = (AB)^{-1}$. To complete the proof, you need to show that you can do the same from the "right".

8 Start with $(A^{-1}A)^t = Id^t = Id$, use property 5 and you get $(A^{-1}A)^t = A^t(A^{-1})^t = Id$, then $(A^{-1})^t$ must be the inverse (again the only thing that is missing is to show that it works if you start with $(AA^{-1})^t$ as well which is trivial.

Conjecture

The set of the matrices in $\mathbb{R}^{m \times n}$, together with the sum and scalar multiplication is a vector space.

Conjecture

A squared matrix $A \in \mathbb{R}^{n \times n}$ is invertible if and only if all of its columns are linearly independent.

Definition

Matrix $A \in \mathbb{R}^{m \times n}$ is **upper triangular** if it has the following shape:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2m} & \dots & a_{2n} \\ 0 & 0 & \ddots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & a_{mm} & \dots & a_{mn} \end{bmatrix}$$

That is, it has zeroes below its main diagonal.

Conjecture

The set of the upper triangular matrices in $\mathbb{R}^{n \times m}$, with the sum and scalar multiplication is a vector subspace of $\mathbb{R}^{n \times m}$.

Definition

Matrix A is **lower triangular** if A^t is upper triangular.

Definition

Matrix A is **diagonal** if it is upper and lower triangular at the same time.

Definition

The **rank** of a matrix A, denoted by $rank(A)$ is the maximum number of linearly independent rows or columns of A.

A convenient way to write down a system of equations:

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

It would be $AX = B$, where,

$$
A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}
$$

Definition

Given a system of equations $AX = B$,

- $\hat{X} \in \mathbb{R}^n$ is a **particular solution** of the system if $A\hat{X} = B$.
- X_0 is an **homogeneous solution** if $AX_0 = 0$.

Note that for $\lambda \in \mathbb{R}$, $A(\hat{X} + \lambda X_0) = A\hat{X} + \lambda AX_0 = B + 0 = B$.

Definition

The **kernel** of $A \in \mathbb{R}^{m \times n}$ is defined as:

$$Ker(A) := \{X \in \mathbb{R}^n | AX = 0\}$$

Conjecture

$Ker(A) \subseteq \mathbb{R}^n$ is a vector subspace of $\mathbb{R}^n$

Definition

The dimension of $Ker(A)$ is called the **nullity** — or nullspace —, and it is denoted by $Null(A)$. If $Ker(A) = \{0\}$, then $Null(A) = 0$.

Note that a system of equations as the one shown before, has unique solution only if $Null(A) = 0$.

Conjecture

Matrix $A \in \mathbb{R}^{n \times n}$ is invertible if and only if $Null(A) = 0$.

Conjecture

$A \in \mathbb{R}^{n \times n}$ is invertible if and only if the system $AX = B$ as a unique solution, for any $B \in \mathbb{R}^n$.

Definition

The **image** of $A \in \mathbb{R}^{n \times n}$ is defined as:

$$Im(A) := \{Y \in \mathbb{R}^n | \exists X \in \mathbb{R}^n, Y = AX\} \equiv \{AX | X \in \mathbb{R}^n\}$$

Conjecture

Let $A \in \mathbb{R}^{n \times n}$. $Im(A)$ is a vector subspace of $\mathbb{R}^n$.

Definition

The dimension of $Im(A)$ is called the **range** of A. Let's denote it as $R(A)$.

Conjecture

Let $A \in \mathbb{R}^{n \times n}$. $R(A)$ is the number of l.i. columns of A.

Conjecture

Consider $A \in \mathbb{R}^{n \times n}$. It holds that $Null(A) + R(A) = n$.

Definition

The quadratic form associated to A is a function $Q_A : \mathbb{R}^n \to \mathbb{R}$ such that for any $X \in \mathbb{R}^n$,

$$Q_A(X) = X^t A X \in \mathbb{R}$$

Conjecture

For any matrix $A \in \mathbb{R}^{n\times n}$, there are always symmetric and antisymmetric matrices S and T such that

$$A = S + T$$

Note: Let $S = \frac{A+A^t}{2}$ and $T = \frac{A-A^t}{2}$. While S is symmetric, T is antisymmetric,

Corollary

A quadratic form can be represented as

$$Q_A(X) = X^t S X$$

with S symmetric.

Quick quiz! 15 min to prove the corollary.

Solution

- Let $A = (S + T)$
- Then $Q_A(X) = X^t(S + T)X = X^tSX + X^tTX$
- But $X^tTX \in \mathbb{R}$, so $(X^tTX)^t = X^tTX$ (a number trasposed is the same number).
- So you end up that $X^tTX = (X^tTX)^t = X^tT^tX$
- But T is anytsymmetric so $T^t = -T$...
- Then $X^tTX = -X^tTX$, so if X^tTX is the number z, you have $z = -z$, that only is true for $z = 0$.
- Then $Q_A(X) = X^tSX$

Definition

Let $A \in \mathbb{R}^{n \times n}$, symmetric. Consider the quadratic form $Q_A(X) = X^t A X$. If for any $X \in \mathbb{R}^n \setminus \{0\}$,

1. $Q_A(X) > 0$, A is **positive definite**,
2. $Q_A(X) \geq 0$, A is **positive semi-definite**,
3. $Q_A(X) < 0$, A is **negative definite**,
4. $Q_A(X) \leq 0$, A is **negative semi-definite**.

Definition

$\lambda \in \mathbb{C}$ is an **eigenvalue** (or characteristic value) of matrix $A \in \mathbb{R}^{n \times n}$ if there is a vector, called **eigenvector**, $X_\lambda \in \mathbb{R}^n \setminus \{0\}$ such that

$$AX_\lambda = \lambda X_\lambda$$

Conjecture

Let $A \in \mathbb{R}^{n \times n}$, and λ_1 and λ_2 two eigenvalues of A, with $\lambda_1 \neq \lambda_2$. If V_1 in the vector subspace associated to λ_1, and V_2 in the vector subspace of λ_2 then V_1 and V_2 are linearly independent.

Conjecture

Given $A \in \mathbb{R}^{n \times n}$ symmetric, then its eigenvalues are real valued.

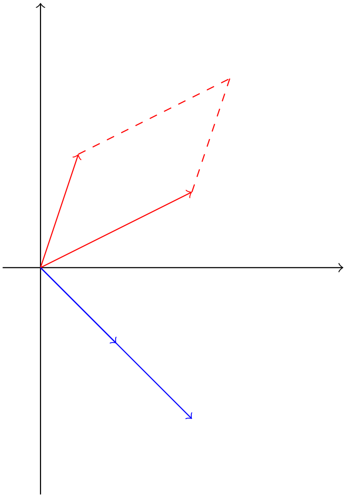
Definition

The **determinant** of a squared matrix A is the hyper-volume of the figure formed by the column vectors of the matrix.

Example

Consider the matrices,

$$A = \begin{pmatrix} 1/2 & 2 \\ 3/2 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}$$



It is easy to see that, given our definition, $det(B) = 0$. It is also easy to show that $det(A) = |1/2 \times 1 - 3/2 \times 2| = 5/2$.

How to calculate the determinant of a big matrix? Recursively. Let $A \in \mathbb{R}^{n \times n}$. Define A_{ij} as:

$$A_{ij} = \begin{pmatrix} a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{pmatrix}$$

That is, what is left of A after removing row i and column j.

Then,

$$det(A) = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} det(A_{ik})$$

You can choose any i that you prefer.

Conjecture

- *A squared matrix is invertible if and only if its determinant is different from zero.*
- *Take a finite set of matrices $\mathbb{A} \subseteq \mathbb{R}^{n \times n}$, with A_i being the ith element of $\mathbb{A}$ then,*

$$det(A_1 A_2 \dots A_k) = det(A_1) det(A_2) \dots det(A_k)$$

- *If A is invertible, then*

$$det(A^{-1}) = \frac{1}{det(A)}$$

- *For any squared A it holds that $det(A^t) = det(A)$.*

Note that λ is an eigenvalue if

$$AX_\lambda = \lambda X_\lambda \qquad \text{with} \quad X_\lambda \neq 0$$

so λ is an eigenvalue of A if

$$(A - \lambda I)X_\lambda = 0$$

or $X_\lambda \in ker(A - \lambda I)$, which implies that $ker(A - \lambda I) \neq \{0\}$, and therefore $(A - \lambda I)$ must be not invertible! But if $(A - \lambda I)$ is not invertible, then $det(A - \lambda I) = 0$.

Corollary

λ is an eigenvalue of A if and only if $det(A - \lambda I) = 0$.

Definition

Given $A \in \mathbb{R}^{n \times n}$, the **characteristic polynomial** of A is defined as the function $p_A : \mathbb{R} \to \mathbb{R}$ such that

$$p_A(\lambda) = det[A - \lambda I]$$

So, λ is an eigenvalue of A if $p_A(\lambda) = 0$

Conjecture

If A is symmetric, then the eigenvectors of different eigenvalues are orthogonal.

For practical reasons, consider the matrix V as the matrix that has in its columns the eigenvectors of A, and $D(\lambda)$ the diagonal matrix that contains in the column i, the eigenvalue that corresponds to the eigenvector in the column i in V.
Note that:

$$AV = VD(\lambda) \quad \Leftrightarrow \quad A = VD(\lambda)V^{-1}$$

Note that given the properties of matrix multiplication

$$A^{-1} = VD\left(\frac{1}{\lambda}\right)V^{-1}$$

which is one of the fundamental properties of the symmetric matrices.

Conjecture

Given $A \in \mathbb{R}^{n\times n}$, symmetric. It holds that,

$$A = VDV^t$$

With D the diagonal with the eigenvalues of A and V the unit eigenvectors of A.

Conjecture

Let $A \in \mathbb{R}^{n \times n}$, symmetric.

1. *A is positive definite if all the eigenvalues of A are strictly positive.*
2. *A is positive semi definite if all the eigenvalues are nonnegative.*
3. *A is negative definite if all the eigenvalues are strictly negative.*
4. *A is negative semidefinite if all the eigenvalues are non positive.*

Definition

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is linear if, for any $X, Y \in \mathbb{R}^n$, and for any $\alpha \in \mathbb{R}$

$$f(X+Y) = f(X) + f(Y), \quad f(\alpha X) = \alpha f(X)$$

Definition

The **trace** of a square matrix A ($tr(A)$) is the sum of the elements of its diagonal.

Theorem

Let $A \in \mathbb{R}^{n \times n}$, then:

- *the product of the eigenvalues of A is equal to its determinant, that is,*

$$det(A) = \prod_{i=1}^{n} \lambda_i$$

- *the sum of the eigenvalues of A is equal to its trace, that is,*

$$\sum_{i=1}^{n} a_{i,i} = \sum_{i=1}^{n} \lambda_i$$

- *if A is a triangular matrix, then its eigenvalues are the coefficients in the principal diagonal of the matrix, i.e.,*

$$\lambda_i = a_{i,i}$$