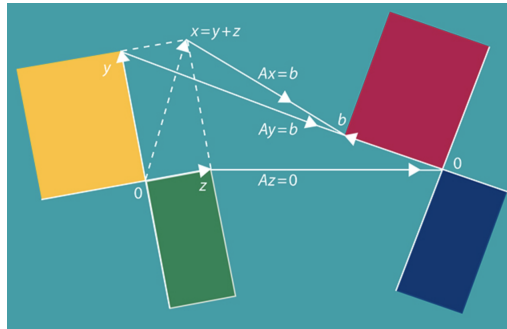# Linear Algebra Reference

Sidharth Baskaran

May 5, 2021

Table of Contents

1 Row and Column Picture

1.1 Row picture

Involves viewing matrix as linear equations graphed on a line or plane. Take the example $A\vec{x} = \vec{b}$ below:

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 8 \\ 9 \\ 10 \end{bmatrix}$$

This can be viewed as the following system:

$$\begin{cases} 1x + 2y + 3z = 8 \\ 3x + 4y + 5z = 9 \\ 4 + 5y + 6z = 10 \end{cases}$$

1.2 Column picture

Involves viewing this setup as a linear combination of column vectors. Take $A\vec{x} = \vec{b}$ again:

$$x \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} + y \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix} + z \begin{bmatrix} 3 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 \\ 9 \\ 10 \end{bmatrix}.$$

1.3 Visualization in Space and Solutions

1.3.1 2D space

In $\mathbb{R}^2$, the equations form a line. Independent column vectors means infinite linear combinations of these to get a set of $\vec{b}$ in $\mathbb{R}^2$. If one column vector is dependent on another, they are parallel and various combinations of $\vec{b}$ are on a line.

1.3.2 3D space

The equations form a plane in $\mathbb{R}^3$. If column vectors independent, infinite linear combination of $\vec{b}$ exist in 3D space. If one vector is a scaled combination of another and the third is independent, then solutions lie on a line. If all three are interdependent, the solution is on a line.

2 Matrix Multiplication

2.1 Row and Column Swapping

Can define elementary row operations in the identity matrix.

2.1.1 Swapping rows

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$$

Modifier B is always on **left**.

2.1.2 Swapping columns

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} b & a \\ d & c \end{bmatrix}$$

Modifier B is always on **right**.

2.2 Elimination steps

Performing elimination:

$$E_{2,1}A + E_{3,2}A = (E_{2,1}E_{3,2})A = U$$

Elimination algorithm:

- $E_{2,1}$ is the pivot. Swap with R_2 if 0 (and $(2,1)$ is nonzero).

- $E_{3,2}$ involves getting $(2,2)$ as a pivot assuming nonzero to get $(3,2)$ as 0

- Result is invertible and non-singular, where U is upper-triangular

Matrix multiplication is not necessarily commutative but always associative.

2.3 Matrix Multiplication Facts

If A is an $m \times n$ matrix and B is $n \times p$, then $AB = C$ must be $m \times p$. Standard method would be to take dot products by row and column. By column: Columns of C are combinations of columns of A. By row: Rows of C are combinations of rows of B.

2.4 Example (Row)

$$\begin{bmatrix} 2 & 7 \\ 3 & 8 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 6 \end{bmatrix} + \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \begin{bmatrix} 0 & 0 \end{bmatrix}$$

3 Factorization into A=LU

3.1 Notation

E_{21} is the location at row 2 and column 1, used to eliminate this value.

3.2 Inverse

$$AA^{-1} = I = A^{-1}A$$

Matrix multiplication is not commutative:

$$\left(AB^{-1}\right)\left(BA^{-1}\right) = AIA^{-1} = I$$

Transpose inverse fact:

$$\boxed{\left(A^{-1}\right)^T A^T = I}$$

3.3 Concept

Given $E_{21}A = U$, where U is upper-triangular, $E_{21}^{-1}A = E_{21}^{-1}U$ gives:

$$\boxed{A = LU \text{ where } L = E_{21}^{-1}}$$

4 Linear Transformations

4.1 Rules and Notation

Domain is the input space and codomain is the output space.

- $T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w})$

- $T(c\vec{v}) = cT(\vec{v})$

Thus,

$$\boxed{T(c\vec{v} + d\vec{w}) = cT(\vec{v}) + dT(\vec{w})}$$

where $c, d \in \mathbb{R}$ and $\vec{v}, \vec{w} \in \mathbb{R}^n$.

Notation given from $\mathbb{R}^m$ to $\mathbb{R}^n$:

$$T : \ \mathbb{R}^m \to \mathbb{R}^n.$$

4.2 Nonlinear examples

- $S : \ R^2 \to \mathbb{R}^2$ where $S \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x^2 \\ y^2 \end{bmatrix}$

- $f : \ \mathbb{R} \to \mathbb{R}$ where $f(x) = mx + b$

A transformation represented by the product of some matrix A and the column vector input $\vec{x}$ is always a linear transformation.

5 Inverse Matrices

5.1 Basic Facts

- If square matrix A is invertible (or inverse exists), then $\boxed{A^{-1}A = AA^{-1} = I}$.

- Can test invertibility of matrix using elimination, i.e. the $n \times n$ matrix A must have n nonzero pivots.

- If $\det(A) \neq 0$, then A is invertible.

5.2 Computing inverses

Can compute inverses with Gauss-Jordan, eliminating $[A\ I]$ to $[I\ A^{-1}]$. If a matrix is invertible, then solution to $A\vec{x} = \vec{b}$ is $\vec{x} = A^{-1}\vec{b}$.

A 2×2 matrix is only invertible if $ad - bc \neq 0$:

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Matrix inversion occurs in reverse order:

$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}.$$

6 Linear Transformations and Inverse Matrices

6.1 Example with transformation

The 2×2 matrix A with property $R_\theta(\vec{v}) = A\vec{v}$ rotates the vector by θ. Using the unit circle to find the coordinates using the basis vectors $\vec{e}_1$ and $\vec{e}_2$:

$$R_\theta(\vec{e}_1) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

$$R_\theta(\vec{e}_2) = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

This results in A:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Finding the inverse of this is simply rotating back by θ, so finding R_θ^{-1}:

$$\begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

7 Linear Transformations in Geometry

7.1 Rotations

Any matrix of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ where $a^2 + b^2 = 1$. Thus, $\theta = \tan^{-1}(\frac{b}{a})$, or by any other trigonometric relation.

7.2 Scaling and dilation

Horizontal scaling affects the x-component:

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} kx \\ y \end{bmatrix} \implies A = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$$

Vertical scaling affects the y-component:

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ ky \end{bmatrix} \implies A = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$$

Dilation is scaling by k for both x and y:

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} kx \\ ky \end{bmatrix} \implies A = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$$

7.3 Normalizing a vector

Can make any vector into a unit vector parallel to the original:

$$\boxed{\vec{u} = \frac{\vec{v}}{||\vec{v}||}}$$

The magnitude of a unit vector is always 1 ($||\vec{u}|| = 1$).

7.4 Projections

x-axis:

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}$$

y-axis:

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}$$

Mathematically:

$$||\text{proj}_l(\vec{v})|| = ||\vec{v}|| \cos\theta$$

The dot product:

$$\boxed{\vec{a} \cdot \vec{b} = ||\vec{a}||\,||\vec{b}|| \cos\theta}$$

Unit vector $\vec{u}$ given by the following because the line l can be represented by $\begin{bmatrix} 1 \\ m \end{bmatrix}$:

$$\vec{u} = \frac{1}{\sqrt{1 + m^2}} \begin{bmatrix} 1 \\ m \end{bmatrix}$$

The projection matrix, onto a line of slope m:

$$\text{proj}_l(\vec{v}) = (\vec{u} \cdot \vec{v})\vec{u} = \begin{bmatrix} \frac{v_1 + v_2 m}{1+m^2} \\ \frac{v_1 m + v_2 m^2}{1+m^2} \end{bmatrix}$$

Thus,

$$A = \begin{bmatrix} \frac{1}{1+m^2} & \frac{m}{1+m^2} \\ \frac{m}{1+m^2} & \frac{m^2}{1+m^2} \end{bmatrix}$$

General projection matrix given $a^2 + b^2 = 1$:

$$\boxed{A = \begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix}}$$

7.5 Reflections

Given by:

$$\text{refl}_l(\vec{v}) = 2\text{proj}_l(\vec{v}) - \vec{v}$$

Has the matrix A:

$$A = \begin{bmatrix} \frac{1-m^2}{1+m^2} & \frac{2m}{1+m^2} \\ \frac{2m}{1+m^2} & \frac{m^2-1}{1+m^2} \end{bmatrix}$$

If $a^2 + b^2 = 1$:

$$\boxed{A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}}$$

7.6 Shear

Horizontal:

$$A = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$$

Vertical:

$$A = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$$

8 Inverse of a Linear Transformation

8.1 Definition in I/O space

- Each item in output receives **at most** 1 input $\implies$ injectivity

- Each item in output receives **at least** 1 input $\implies$ surjectivity

- If **both conditions are satisfied** $\implies$ bijectivity

Invertibility is therefore synonymous with bijectivity.

8.2 Conclusions

Injectivity concludes that $\text{rank}(A) = m$, where A is $n \times m$. This is because there must be a leading one in each column.

Surjectivity concludes that the last row in $\text{rref}(A)$ is $0\ 0 \cdots 0\ 1$. Thus there must be no rows of 0 in $\text{rref}(A)$, so all invertible matrices are square. Also an invertible matrix is **nonsingular** and an invertible matrix is **singular**.

9 The Matrix Product

9.1 Composition

Can define the following (linear) transformation:

$$\boxed{T_C(\vec{x}) = T_A(T_B(\vec{x})) = (T_A \circ T_B)(\vec{x})}$$

Following diagram represents the composition:

$$T_A \circ T_B$$
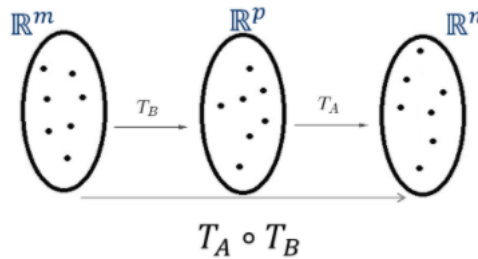
Would imply that A is $n \times p$, B is $p \times m$, and AB is $n \times m$. Can define the following:

$$\boxed{\text{The } i^{th} \text{ column of the matrix } AB \text{ is the matrix-vector product } A(i^{th} \text{ column of the matrix } B)}$$

9.2 Proofs

Claim: The product of 2 invertible matrices must be an invertible matrix.
Proof: Given that $(AB)(AB)^{-1} = I_n$:

$$(AB)(AB)^{-1} = I_n$$
$$A(B(AB)^{-1}) = I_n$$
$$A^{-1}A(B(AB)^{-1}) = A^{-1}I_n$$
$$B(AB)^{-1} = A^{-1}$$
$$B^{-1}B(AB)^{-1} = B^{-1}A^{-1}$$
$$(AB)^{-1} = B^{-1}A^{-1}$$

Claim: If $(AB)^{-1}$ exists, then A and B are both invertible.
Proof: Given that $(AB)(AB)^{-1} = I_n$ and $(AB)^{-1}(AB) = I_n$:

$$A(B(AB)^{-1}) = I_n$$
$$((AB)^{-1}A)B = I_n$$
$$\boxed{\therefore \exists\, A^{-1},\, B^{-1} \in \mathbb{R}^n}$$

9.3 Properties

- Associativity: $(AB)C = A(BC)$
- Distribution: $A(B + C) = AB + BC$
- Respects scalar multiplication: $(kA)B = k(AB) = A(kB)$

10 Transposes, Permutations, Spaces

10.1 Permutations

Function to make row exchanges. Elimination with row exchanges:

$$A = LU \implies PA = LU$$

Works for any invertible A.

$$P = \text{identity with reordered rows (exchanges)}$$

Count of possible reorderings ($n \times n$ permutations): $n! = n(n-1)\cdots 3(2)(1)$.

$$\boxed{P^{-1} = P^T \text{ and } P^T P = I}$$

Defining a transpose, or flip over diagonal:

$$(A^T)_{ij} = A_{ji}$$

For symmetric matrices, transpose does not cause change; $A^T = A$. If two rectangular matrices R^T and R give a square matrix, then $R^T R$ is always symmetric.

$$\boxed{(R^T R)^T = R^T R^{TT} = R^T R}$$

10.2 Vector Spaces and Subspaces

Examples: $\mathbb{R}^2$ is all vectors in 2D space, $x - y$ plane: $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$. $\mathbb{R}^3$ is all vectors with 3 components. All combinations of vectors in $\mathbb{R}^n$ yield a result in that space $\mathbb{R}^n$.

$$\boxed{\mathbb{R}^n \text{ is all column vectors with } n \text{ components.}}$$

The origin exists to allow for scalar multiplication and addition of vectors. Every vector space has a $\vec{0}$.

10.2.1 Subspaces

If a vector space is defined as 1st quadrant in $\mathbb{R}^2$, then multiplying by a negative scalar k removes the result from that space, so it is not **closed** under that operation, so this is not a vector space. Vector space must be closed under linear combinations. Thus, subspace in $\mathbb{R}^2$ is all multiples of that vector, a line and the line must go through $\vec{0}$. Every subspace must contain $\vec{0}$.

Subspaces of $\mathbb{R}^2$:

- All of $\mathbb{R}^2$

- Any line through $\vec{0}_2$ or L

- Just $\vec{0}_2$ or Z

Similarly, for $\mathbb{R}^3$ can have $\mathbb{R}^3$, plane, line, $\vec{0}_3$.

Given $A = \begin{bmatrix} 1 & 3 \\ 2 & 3 \\ 4 & 1 \end{bmatrix}$, all linear combinations of these columns form a subspace. This is called **column space**, $C(A)$. This would form a plane in $\mathbb{R}^3$. Thus, the column space is a subspace.

11 Image and Kernel

11.1 Defining Image and Kernel

11.1.1 Image

Of a function, the set of vectors in the codomain hit by the domain. The image of a function $f : \mathbb{R}^m \to \mathbb{R}^n$ is:

$$\boxed{\text{Im}(f) = \{\vec{y} \in \mathbb{R}^n | \exists \, \vec{x} \in \mathbb{R}^m \text{ s.t. } f(\vec{x}) = \vec{y}\}}$$

Similar in concept to the range of a function in non-linear context.

11.1.2 Kernel

Set of vectors in the domain that are mapped to $\vec{0}$ in the codomain. Kernel of a function $f : \mathbb{R}^m \to \mathbb{R}^n$ is:

$$\boxed{\mathrm{Ker}(f) = \{\vec{x} \in \mathbb{R}^m | f(\vec{x}) = \vec{0}\}}$$

Analogous to the roots/zeros of a polynomial.

11.2 Examples

Following linear transformation's image lives in $\mathbb{R}^2$:

$$T(\vec{x}) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \vec{x}$$

Following linear transformation's image forms a plane that is $\mathbb{R}^2$:

$$T(\vec{x}) = \begin{pmatrix} 1 & -1 \\ 0 & 2 \\ 2 & 2 \end{pmatrix} \vec{x}$$

11.3 Span

If A is an $n \times m$ matrix, then image of $T(\vec{x}) = A\vec{x}$ is set of all vectors in $\mathbb{R}^n$ that are linear combinations of column vectors of A.

Thus, the span of a set of n vectors is all linear combinations of those vectors:

$$\boxed{\mathrm{span}(\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n) = \{\sum c_i \vec{v}_i | c_i \in \mathbb{R}\}}$$

So the span of column vectors of A is the image of the associated linear transformation.

11.4 Kernel

Kernel amounts to finding solutions to $A\vec{x} = \vec{0}$. Kernels are closed under linear combinations. Kernel can never be empty set, it always holds true that $T(\vec{0}) = \vec{0}$.

11.5 Invertible Linear Transformations

Main conclusions about image and kernel:

- Kernel is always (trivially) $\{\vec{0}\}$, else would imply dependance and therefore singularity in the associated matrix A so $\boxed{\mathrm{ker}(T) = \{\vec{0}\}}$

- Image is always the space $\mathbb{R}^n$ if associated matrix A is $n \times n$, so $\boxed{\mathrm{Im}(T) = \mathbb{R}^n}$

12 Subspaces and Basis

12.1 Image and Kernel

If $T : \mathbb{R}^m \to \mathbb{R}^n$ then $\text{Im}(T) \subset \mathbb{R}^n$ and $\ker(T) \subset \mathbb{R}^m$ because the associated matrix A is $n \times m$ in dimension.

Both are closed under linear combinations:

- If $\vec{y_1}, \vec{y_2} \in \text{Im}(T)$ then $a\vec{y_1} + b\vec{y_2} \in \text{Im}(T)$ as well
- If $\vec{x_1}, \vec{x_2} \in \text{Ker}(T)$ then $a\vec{x_1} + b\vec{x_2} \in \text{Ker}(T)$ as well

12.2 Subspaces

Collection of vectors in $\mathbb{R}^n$ is called a subspace in $\mathbb{R}^n$ if collection is nonempty and closed under linear combinations. Examples (and counterexamples):

- $W = \left\{ \begin{pmatrix} 3s \\ 2 + 5s \end{pmatrix} \mid s \in \mathbb{R} \right\} \subset \mathbb{R}^2$ is not a subspace because $\vec{0}$ is not contained within the set, so not closed under linear combinations

- $W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_2 \end{pmatrix} \mid 2x_1 + x_2 - x_3 = 0 \right\}$ is a subspace due to matrix representation and the image of this matrix containing $\vec{0}$ due to $T(\vec{0} = \vec{0})$

$\mathbb{R}^2$ is not a subspace of $\mathbb{R}^3$ because though a plane can be drawn in $\mathbb{R}^3$, its components will be of the form $\begin{bmatrix} x \\ y \\ k \end{bmatrix}$, where k is fixed. Since $\mathbb{R}^3$ vectors always have 3 coordinates, they can't represent $\mathbb{R}^2$. $\mathbb{R}^2$ can only be represented by $\mathbb{R}^2$ vectors. Thus, R^n is not a subspace of $\mathbb{R}^{n+1}$.

Claim: Span of a set of vectors in $\mathbb{R}^n$ is a subspace of $\mathbb{R}^n$.
Proof:
Let $S = \{\vec{v_1}, \vec{v_2}, \cdots \vec{v_m}\}$. Let $\vec{w}, \vec{y} \in \text{span}(S)$. Thus, $\vec{w} = \sum c_i \vec{v_i}$ and $\vec{y} = \sum d_i \vec{v_i}$ where $d_i, c_i \in \mathbb{R}$.

$$a\vec{w} + b\vec{y} = a \sum c_i \vec{v_i} + b \sum d_i \vec{v_i}$$
$$= \sum ac_i \vec{v_i} + \sum bd_i \vec{v_i}$$
$$= \sum (ac_i + bd_i) \vec{v_i} \in \text{span}(S)$$

List of subspaces in $\mathbb{R}^2$ would be $\mathbb{R}^2$, $\{t\vec{v} \mid t \in \mathbb{R}\}$, $\{\vec{0}\}$.

12.3 Intersection and Union

If V and W are collections of vectors in $\mathbb{R}^n$:

- $V \cap W = \{\vec{x} \mid \vec{x} \in V \text{ and } \vec{x} \in W\}$ is the intersection
- $V \cup W = \{\vec{x} \mid \vec{x} \in V \text{ or } \vec{x} \in W\}$ is the union

12.4   Redundant Vectors

If for some transformation T there exists the following:

$$\text{im}(T) = \text{span}\left\{\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}2\\2\\2\end{pmatrix}, \begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}2\\3\\4\end{pmatrix}\right\}$$

There are redundant vectors in this case. The minimum number of vectors in the span is 2, for $\vec{0}$ cannot be produced then. With 3 vectors in $\mathbb{R}^3$, any one can be the result of linear combinations of the other 2. So, it would be appropriate to say that:

$$\text{im}(T) = \left\{\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\2\\3\end{pmatrix}\right\}$$

These are then **linearly independent**. This set forms a **basis** for that set of vectors. Thus, the basis can be found for any matrix. The basis of I_n is then $\{\vec{e}_1, \vec{e}_2 \cdots \vec{e}_n\}$.

12.5   Intersection and Union

> If V and W are subspaces of $\mathbb{R}^n$, then $V \cap W$ is a subspace of $\mathbb{R}^n$ and $V \cup W$ is **not** a subspace of $\mathbb{R}^n$.

An intersection is the items contained in both sets, so $\vec{0} \in V \cap W$. If $\vec{v}, \vec{w} \in V \cap W$, then $\vec{v}, \vec{w} \in V$ and $\vec{v}, \vec{w} \in W$. This means that $\vec{v} + \vec{w} \in V$ and $\vec{v} + \vec{w} \in W$ so $\vec{v} + \vec{w} \in V \cap W$. Similarly, if some $k\vec{v} \in V \cap W$ where $k \in \mathbb{R}$ then $k\vec{v} \in V$ and $k\vec{v} \in V$. Thus, $V \cap W$ is a subspace of $\mathbb{R}^n$.

The union is the items contained in either set. If $V = \text{span}(\vec{e}_2)$ and $W = \text{span}(\vec{e}_1)$, then let $\vec{y} = \begin{bmatrix}0\\y\end{bmatrix} \in V$ and $\vec{x} = \begin{bmatrix}x\\0\end{bmatrix} \in W$. Thus, $\vec{y}, \vec{x} \in V \cup W$. However, $a\vec{x} + b\vec{y} \notin V \cup W$ where $a, b \in \mathbb{R}$.

13   Basis of a Kernel

13.1   Example

Let $A = \begin{bmatrix}1 & 4 & 7\\2 & 5 & 8\\3 & 6 & 9\end{bmatrix}$. Finding the basis:

$$\text{rref}(A) = \begin{bmatrix}1 & 0 & -1\\0 & 1 & 2\\0 & 0 & 0\end{bmatrix}$$

Since there is redundance, a possible expression of the basis of A:

$$\mathfrak{B} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \right\}$$

When finding kernel, must solve $A\vec{x} = \vec{0}$. So with $[\mathrm{rref}(A)|\vec{0}]$:

$$\vec{x} = \begin{bmatrix} t \\ -2t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Thus,

$$\ker(A) = \left\{ t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \,\middle|\, t \in \mathbb{R} \right\}$$

Subsequently, the basis of the kernel of A can be represented as $\mathfrak{B} = \left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}$.

Equivalent statements for $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \dots, \vec{v}_m\}$ being linearly independent:

- None of the vectors are redundant

- Only relation is trivial

- Kernel of $\begin{pmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_m \\ | & | & & | \end{pmatrix}$ is trivial

- Rank of $\begin{pmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_m \\ | & | & & | \end{pmatrix}$ is m

- If $m = n$ then $\begin{pmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_m \\ | & | & & | \end{pmatrix}$ reduces to I_n

14   Dimension

14.1   Rank and independence

If $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_m\}$ is a collection if independent vectors then

$$\begin{pmatrix} | & | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \dots & \vec{v}_m \\ | & | & | & & | \end{pmatrix}$$

must have a rank of m. This is because row reducing the matrix corresponds to the following relation:

$$c_1 \vec{v}_1 + c_2 \vec{v}_2 + c_3 \vec{v}_3 + \cdots + c_m \vec{v}_m = \vec{0}$$

Also, $m \leq n$ where n is the number of rows in each column vector, in order to have linear independence for this set.

14.2 Dimension

Considering an xy-plane in $\mathbb{R}^3$:

$$V = \left\{ \begin{bmatrix} s \\ t \\ 0 \end{bmatrix} \;\middle|\; s, t \in \mathbb{R} \right\}$$

The basis of this set contains 2 vectors (e.g. dimension of 2), with example being:

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

> If V is a subspace of $\mathbb{R}^n$ and $\mathcal{B}$ and $\mathcal{C}$ are two bases of V, then $\mathcal{B}$ and $\mathcal{C}$ contain the same number of vectors.

Dimension of a subspace is number of vectors in the basis.

14.2.1 Example

Considering the following matrix:

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 & 2 \\ 1 & 2 & 0 & 2 & 3 \\ 1 & 2 & 0 & 3 & 4 \\ 1 & 2 & 0 & 4 & 5 \end{bmatrix}$$

By discounting the redundant vectors, a possible basis for $\text{Im}(A)$:

$$\mathcal{B}_{\text{image}} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \right\}$$

So the dimension of $\text{Im}(A)$ is 2. Finding a basis for $\ker(A)$ is the same as solving $A\vec{x} = \vec{0}$:

$$\ker(A) = \left\{ \begin{bmatrix} -2s - w \\ s \\ t \\ -w \\ w \end{bmatrix} \;\middle|\; s, t, w \in \mathbb{R} \right\} = \left\{ s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + w \begin{bmatrix} -1 \\ 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} \;\middle|\; s, t, w \in \mathbb{R} \right\}$$

So the basis for ker(A):

$$\mathcal{B}_{\text{kernel}} = \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\}$$

And dimension of ker(A) is 3. However, it is shown that rref(A) gives dimension of **image and kernel**.

14.3 Rank-Nullity Theorem

- If T is a linear transformation, then
  $\dim(\text{Im}(T)) + \dim(\ker(T)) =$ dimension of domain of T

- If A is a matrix, then rank(A) + nullity(A) = number of columns of A

- In a linear system,
  number of leading variables + number of free variables = total number of variables

Considering non-invertible matrices A and B, let AB be invertible. It must hold true that $\ker(B) = \{\vec{0}\}$. If the dimensions of B are $p \times n$, Im(B) is a subspace of $\mathbb{R}^p$ has dimension n. This means that it is a vertically rectangular matrix with $n \leq p$. Thus, A is $n \times p$ so it is horizontally rectangular.

15 Coordinates

15.1 Coordinate vectors

For example, the basis of xy plane can be:

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\}$$

To form $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ with this basis, can do $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$. The coefficients used form the following vector:

$$\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

Known as $\mathcal{B}$-**coordinate vector**. Notation:

$$\left[\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right]_{\mathfrak{B}} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 2 \end{bmatrix}$$

Generally, given $\mathfrak{B} = \{\vec{v}_1, \vec{v}_2, \vec{v}_3, \dots, \vec{v}_m\} \subset \mathbb{R}^n$ is linearly independent, then $[\vec{v}_i]_{\mathfrak{B}} = \vec{e}_i \in \mathbb{R}^m$.

This is because row-reducing the matrix of $\mathfrak{B}$ gives $\text{rref}(A)$ where A is this matrix. Given the same $\mathfrak{B}$, can find the components of $\vec{w}$:

$$[\vec{w}]_{\mathfrak{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$$

$$\vec{w} = c_1 \vec{v}_1 + c_2 \vec{v}_2 + \cdots + c_m \vec{v}_m$$

Thus,

$$\vec{w} = \begin{bmatrix} | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$$

Matrix is called change of basis matrix S. A standard basis is given as $\vec{e}_1, \vec{e}_2, \cdots$. A nonstandard basis is not of this form.

15.2 B-matrix

If A is $n \times n$ and $T(\vec{x}) = A\vec{x}$ where $T : \mathbb{R}^n \to \mathbb{R}^n$, then there exists a matrix B such that $[T(\vec{x})]_{\mathfrak{B}} = B[\vec{x}]_{\mathfrak{B}}$. This is called the $\mathfrak{B} - matrix$. If $\vec{v}_i \in \mathfrak{B}$, then $[\vec{v}_i]_{\mathfrak{B}} = \vec{e}_i$.

This means that $[T(\vec{v}_i)]_{\mathfrak{B}} = B[\vec{v}_i]_{\mathfrak{B}} = B\vec{e}_i$, so **the ith column of B must be** $[T(\vec{v}_i)]_{\mathfrak{B}}$.

Multiple ways to calculate $\mathfrak{B}$-matrix of T, considering T to be a projection onto $y = \frac{x}{3}$:

Project onto the line $y = \dfrac{x}{3}$

Start here with a vector $\vec{x}$ in $\mathbb{R}^2$ $T(\vec{x})$

Convert into the non-standard basis $\mathfrak{B}$

Multiply by S^{-1} Multiply by S^{-1}

Convert into the non-standard basis $\mathfrak{B}$

$[\vec{x}]_{\mathfrak{B}}$ $[T(\vec{x})]_{\mathfrak{B}}$

Project onto the line $y = \dfrac{x}{3}$

Multiply by B
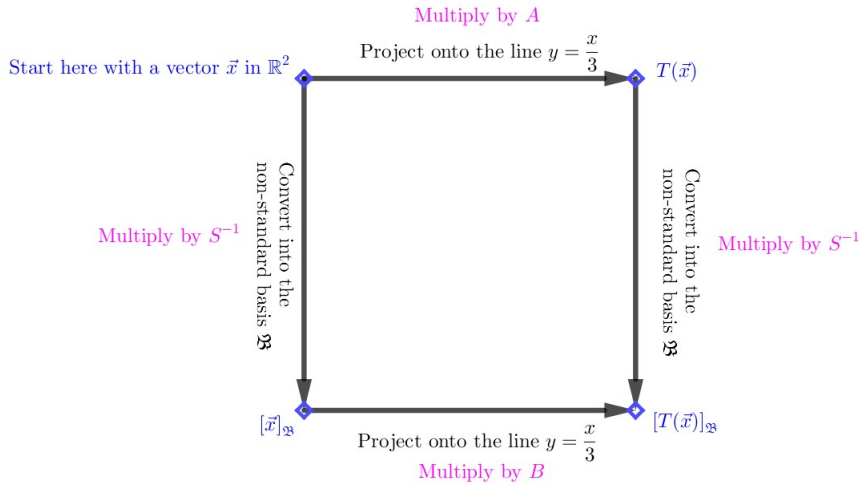
Means that multiple ways to get to $[T(\vec{x})]_{\mathfrak{B}}$. When following $\vec{x}$ and going right and down:

$$S^{-1}(A\vec{x}) = S^{-1}A\vec{x} = [T(\vec{x})]_{\mathfrak{B}}$$

Going down and right:

$$B\left(S^{-1}\vec{x}\right) = BS^{-1}\vec{x} = [T(\vec{x})]_{\mathfrak{B}}$$

Thus,

$$S^{-1}A = BS^{-1}$$

$$\boxed{S^{-1}AS = B}$$

If this is satisfied, then A is similar to B or $A \sim B$.

16 Determinants

16.1 Introduction to Determinant

Can define the 2×2 determinant as a function $D : \mathbb{M}_{2\times 2} \to R$. It can be observed that 2×2 matrix A is only invertible if $D(A) = ad - bc \neq 0$.

16.2 Cross-Product

Given $\begin{bmatrix} a \\ c \end{bmatrix}$ and $\begin{bmatrix} b \\ d \end{bmatrix}$, the cross product is defined as the $\mathbb{R}^3$ vector $D(A)\vec{e}_3 = (ad - bc)\vec{e}_3$. The direction of this vector is the sign of $\det(A)$.

Can visualize using right hand rule: if sweeping index into middle is appropriate for the vectors, then the direction of thumb is cross-product direction (positive). Otherwise, sign is negative.

16.2.1 Algorithm

For $\vec{v} = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} w_x \\ w_y \\ w_z \end{bmatrix}$:

$$\vec{v} \times \vec{w} = c_z \vec{e}_1 + c_y \vec{e}_2 + c_z \vec{e}_z$$

Following through, to calculate each component ignore the desired row and perform cross-product on remaining matrix:

$$c_x = \begin{bmatrix} v_y \\ v_z \end{bmatrix} \times \begin{bmatrix} w_y \\ w_z \end{bmatrix} = v_y w_z - v_z w_z$$

The y component is done as $bc - ad$ compared to $ad - bc$.

$$c_y = \begin{bmatrix} v_x \\ v_z \end{bmatrix} \times \begin{bmatrix} w_x \\ w_z \end{bmatrix} = w_x v_z - w_z v_x$$

$$c_z = \begin{bmatrix} v_y \\ v_z \end{bmatrix} \times \begin{bmatrix} w_y \\ w_z \end{bmatrix} = v_x w_y - v_y w_x$$

16.3 Determinant Theory

Considering $A = \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \\ | & | & | \end{bmatrix}$, say it is singular such that $\vec{v}_3 \in \text{span}\{\vec{v}_1, \vec{v}_2\}$. Because it is assumed that $\vec{v}_1, \vec{v}_2$ are linearly independent, then $\text{span}\{\vec{v}_1, \vec{v}_2\}$ is perpendicular to $\vec{v}_1 \times \vec{v}_2$ by definition. Thus, $\boxed{(\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3 = 0}$. If $\vec{v}_1, \vec{v}_2$ are not linearly independent, then this is still 0 because the cross-product (area of parallelogram made by vectors) is still 0.

If $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are linearly independent, then $(\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3 \neq 0$.

$$\boxed{D(A) = \det(A) = (\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3}$$

The sub-matrices used when computing cross-products are **minors**. Can rewrite determinant:

$$\boxed{\det A = a_{1,3} |A_{1,3}| - a_{2,3} |A_{2,3}| + a_{3,3} |A_{3,3}|}$$

Must use following rule for the sign of constant terms $a_{m,n}$ (dot product):

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$$

16.4 Rules

Determinant of $n \times n$ follows recursively:

$$\boxed{\det A = a_{1,1} |A_{1,1}| - a_{1,2} |A_{1,2}| + a_{1,3} |A_{1,3}| + \cdots \pm a_{1,n} |A_{1,n}|}$$

Rules:

- Swapping rows multiplies determinant by -1

- Multiplying row by m scales determinant by m

- Replacing row with sum of row and multiple of another does not change determinant

- If A and B are $n \times n$, then $\det(AB) = \det(A)\det(B)$

- Cramer's rule: If $A\vec{x} = \vec{b}$ is a linear system with invertible A then $\vec{x}$ components can be determined from $x_i = \frac{\det(A_{-b,i})}{\det(A)}$ where $A_{-b,i}$ replaces i^{th} column of A with $\vec{b}$

17 Intro to Dynamical Systems

17.1 Dynamical Systems and Eigenvectors

In general, a discrete dynamical system can be modeled as:

$$\boxed{\vec{x}(t+1) = A\vec{x}(t)}$$

where the transformation undergone by the system is $\vec{x}(t) \to \vec{x}(t+1)$ with matrix A. Additionally, note that $\vec{x}(t) = \begin{bmatrix} c(t) \\ r(t) \end{bmatrix}$ where $c(t)$ and $r(t)$ are some closed formulas.

Finding $\vec{x}(t)$ for an arbitrary integer $t > 0$:

$$\boxed{\vec{x}(t) = A^t \vec{x}(0) = A^t \vec{x}_0}$$

Repeat definition of eigenvector from below:

> If A is an $n \times n$ matrix, an eigenvector of A is a nonzero vector $\vec{v}$ that has the property that $\vec{v}$ and $A\vec{v}$ are parallel. Same as saying that $A\vec{v} = \lambda\vec{v}$, so λ is an eigenvalue.

17.2 Dynamical Systems Example

Following equations model transformation from t to $t+1$:

$$\boxed{\begin{aligned} c(t+1) &= 0.86c(t) + 0.08r(t) \\ r(t+1) &= -0.12c(t) + 1.14r(t) \end{aligned}}$$

Is discrete dynamical linear system: changed over discrete time interval and dynamic as variables change according to t. As a matrix-vector equation:

$$\begin{pmatrix} 0.86 & 0.08 \\ -0.12 & 1.14 \end{pmatrix} \begin{pmatrix} c(t) \\ r(t) \end{pmatrix} = \begin{pmatrix} c(t+1) \\ r(t+1) \end{pmatrix}$$

$\vec{x}(t) = \begin{pmatrix} c(t) \\ r(t) \end{pmatrix}$ is the **state vector** at time t. $\vec{x}(0) = \begin{pmatrix} c_0 \\ r_0 \end{pmatrix}$ is the **initial state vector**.

Calculating arbitrary state vector:

$$\begin{pmatrix} 0.86 & 0.08 \\ -0.12 & 1.14 \end{pmatrix}^t \vec{x}_0 = \vec{x}(t)$$

In this example, $c(t) = (100)1.1^t$ and $r(t) = (300)1.1^t$, so next state vector is 1.1 times the current. However, for $\vec{x}_0 = \begin{pmatrix} 1000 \\ 1000 \end{pmatrix}$ there exists no such scalar pattern. Can use the basis of 2 (scalar pattern respected) vectors:

$$\mathcal{B} = \left\{ \begin{pmatrix} 100 \\ 300 \end{pmatrix}, \begin{pmatrix} 200 \\ 100 \end{pmatrix} \right\}$$

Writing this state vector as a lin. combination:

$$\begin{pmatrix} 1000 \\ 1000 \end{pmatrix} = 2 \begin{pmatrix} 100 \\ 300 \end{pmatrix} + 4 \begin{pmatrix} 200 \\ 100 \end{pmatrix}$$
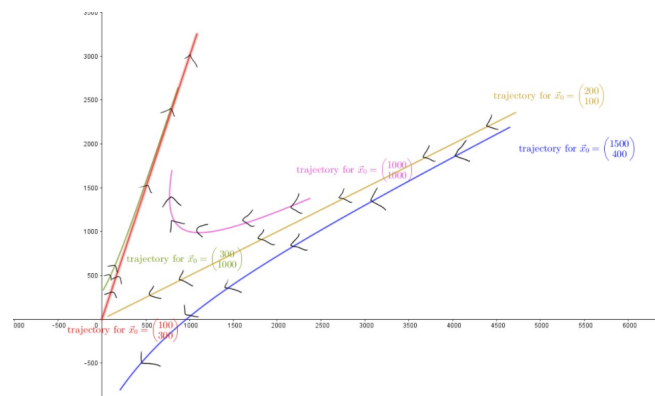
After applying coefficient matrix to both sides and simplifying:

$$\begin{pmatrix} 0.86 & 0.08 \\ -0.12 & 1.14 \end{pmatrix}^t \begin{pmatrix} 1000 \\ 1000 \end{pmatrix} = 2(1.1)^t \begin{pmatrix} 100 \\ 300 \end{pmatrix} + 4(0.9)^t \begin{pmatrix} 200 \\ 100 \end{pmatrix}$$

Thus,

$$c(t) = 200(1.1)^t + 800(0.9)^t$$
$$r(t) = 600(1.1)^t + 400(0.9)^t$$

Different trajectories for various initial state vectors:

Called a **phase portrait** for a discrete dynamical system. Indicates performing of system based on initial states. The 2 state vectors in the basis are **eigenvectors**.

> If A is an $n \times n$ matrix, an eigenvector of A is a nonzero vector $\vec{v}$ that has the property that $\vec{v}$ and $A\vec{v}$ are parallel. Same as saying that $A\vec{v} = \lambda\vec{v}$, so λ is an eigenvalue.

If there exists an $n \times n$ matrix A with $\lambda = 0$, then kernel of A must be nontrivial because $A\vec{v} = \vec{0} = 0\vec{v}$, therefore A is singular.

18   Eigenvalue of a Matrix

18.1   Eigenvalue for rotation transformation

Claim: If $0 < \theta < 2\pi$ then transformation $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ only has an eigenvector when $\theta = \pi$ (when $\lambda = -1$).

Proof: The matrix is $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Following by the definition of an eigenvector:

$$A\vec{v} = \lambda v \iff$$
$$A\vec{v} - \lambda\vec{v} = \vec{0} \iff$$
$$A\vec{v} - \lambda(I\vec{v}) = \vec{0} \iff$$
$$A\vec{v} - (\lambda I)\vec{v} = \vec{0} \iff$$
$$(A - \lambda I)\vec{v} = \vec{0} \iff$$
$$\det(A - \lambda I) = 0$$

Thus,

$$\det(A - \lambda I) = 0$$
$$\det\left(\begin{array}{cc} \cos\theta - \lambda & -\sin\theta \\ \sin\theta & \cos\theta - \lambda \end{array} \right) = 0$$
$$\lambda^2 - 2\lambda\cos\theta + \cos^2\theta + \sin^2\theta = 0$$
$$\lambda^2 - 2\lambda\cos\theta + 1 = 0$$

The discriminant of this quadratic $(b^2 - 4ac)$ is $4\cos^2\theta - 4$, so for a real solution $4\cos^2\theta - 4 \geq 0$. It then follows that:

$$4\cos^2\theta - 4 \geq 0$$
$$\cos^2\theta \geq 1$$
$$\cos^2\theta = \pm 1$$
$$\theta = \pi$$

Note that because $(A - \lambda I)\vec{v} = \vec{0}$ implies a nontrivial kernel for $A - \lambda I$, $\det(A - \lambda I) = 0$.

18.2   Characteristic Polynomials

Characteristic polynomial is for $\det(A - \lambda I)$ with variable λ:

$$\boxed{P_A(\lambda) = \det(A - \lambda I)}$$

General polynomial for $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$:

$$p_A(\lambda) = \det(A - \lambda I)$$
$$= \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix}$$
$$= (a - \lambda)(d - \lambda) - bc$$
$$= \lambda^2 - (a + b)\lambda + (ad - bc)$$

Ends up that $\text{tr}(A) = a + d$ and $\det(A) = ad - bc$:

$$\boxed{p_A(\lambda) = \lambda^2 - \text{tr}\, A\lambda + \det A}$$

18.2.1 General formula

In general, if A is an $n \times n$ matrix, then

$$\boxed{p_A(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} \text{tr}\, A\lambda^{n-1} + \cdots + \det A}$$

Conjectures:

- By FTLA, degree n polynomial will have n complex roots so at least n real eigenvalues

- If all n roots are real, then $\text{tr}(A)$ is sum of eigenvalues and determinant is the product of them

- Since roots are either real or in complex conjugate pairs, $(a + bi$ or $a - bi)$ then when n is odd A has at least 1 real eigenvalue

19 Eigenvector of a Matrix

19.1 Eigenspace

Kernel of a matrix always forms subspace of domain. If λ is an eigenvalue for A, kernel of $A - \lambda I$ is the **eigenspace** associated with λ and this is denotes as $E_\lambda = \ker(A - \lambda I)$.

If $A - \lambda I$ has column vectors $\vec{v_1}$ and $\vec{v_2}$, then $E_\lambda = \text{span} \left\{ \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \right\}$ where $c_1\vec{v_1} + c_2\vec{v_2} = \vec{0}$.

19.2 Multiplicity

Dimension of eigenspace E_λ is **geometric multiplicity** of λ. Multiplicity of root λ is **algebraic multiplicity** in characteristic polynomial $p_A(\lambda)$. Therefore, geometric multiplicity $\leq$ algebraic multiplicity, considering the case of $p_A(\lambda) = (\lambda - \lambda_0)^2$.

Thus, this represents a 2×2 matrix which fixes one line and moves every other line, known as a shear. All lines but x-axis move. Has characteristic polynomial $p_A(\lambda) = (\lambda - 1)^2$, so $E_1 = \ker \begin{bmatrix} 0 & k \\ 0 & 0 \end{bmatrix} = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$.

19.3   Eigenbasis

Consists of the eigenvectors of the coefficient matrix. **An $n \times n$ matrix needs n linearly independent eigenvectors to have an eigenbasis**. This means that if A has eigenvectors $\lambda_1 \neq \lambda_2$, then $E_{\lambda_1} \cap E_{\lambda_2} = \{\vec{0}\}$. This is because if some $\vec{v} \in E_{\lambda_1}$, then $A\vec{v} = \lambda_1 \vec{v}$ and $A\vec{v} = \lambda_2 \vec{v}$. Thus, $(\lambda_1 - \lambda_2)\vec{v} = \vec{0}$, so $\vec{v} = \vec{0}$. Furthermore, if E_{λ_1} has basis $\mathfrak{E}_{\lambda_1}$ and E_{λ_2} with $\mathfrak{E}_{\lambda_2}$, $\mathfrak{E}_{\lambda_1} \cup \mathfrak{E}_{\lambda_2}$ is linearly independent as well with the total elements being the sum of the number of elements in each individual basis. Can then be concluded that:

> An $n \times n$ matrix has eigenbasis iff sum of geometric multiplicities of eigenvalues is n.

20   Diagonalization

20.1   Diagonalization and Properties

Diagonal matrix has entries not along the main diagonal be all 0. Example:

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

Characteristic polynomial ends up being

$$p_A(\lambda) = \det \begin{bmatrix} 1 - \lambda & 0 & 0 \\ 0 & 3 - \lambda & 0 \\ 0 & 0 & -1 - \lambda \end{bmatrix}$$

$$p_A(\lambda) = (1 - \lambda)(3 - \lambda)(-1 - \lambda)$$

Matrix similar to diagonal matrix is **diagonalizable**. Thus, A is diagonalizable if there exists an invertible S and diagonal matrix D such that $\boxed{S^{-1}AS = D}$. Because it is known that if $A \sim B$, $p_A(\lambda) = p_B(\lambda)$, whenever A is diagonalizable, the eigenvalues of A will be diagonal entries of any diagonal matrix A is similar to.

Also note that for a diagonal matrix D, D^t consists of all the diagonal entries $\lambda_1^t, \lambda_2^t \cdots, \lambda_n^t$:

$$D^t = \begin{bmatrix} \lambda_1^t & 0 & 0 & 0 \\ 0 & \lambda_2^t & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_n^t \end{bmatrix}$$

It then follows that if $S^{-1}AS = D$:

$$S^{-1}AS = D$$
$$(S^{-1}AS)^t = D^t$$
$$(S^{-1}AS)(S^{-1}AS)(S^{-1}AS)\ldots(S^{-1}AS) = D^t$$
$$S^{-1}A^tS = D$$
$$A^t = SD^tS^{-1}$$

20.2 Diagonalization and Eigenbasis

A square matrix is diagonalizable if it has an eigenbasis.

Let the eigenbasis be $\mathfrak{E} = \{\vec{v}_1, \vec{v}_2, \vec{v}_3, \ldots, \vec{v}_n\}$. To show A is diagonalizable, the $\mathfrak{E}$-matrix can be shown to be diagonal. First, let A's transformation be $T_A(\vec{x}) = A\vec{x}$. The $\mathfrak{E}$-matrix is D. Thus, $\boxed{D[\vec{x}]_\mathfrak{E} = [T_A(\vec{x})]_\mathfrak{E}}$. The first column of D is $[T_A(\vec{v}_1)]_\mathfrak{E}$. However, because $\mathfrak{E}$ is an eigenbasis for A, it must hold true that:

$$A\vec{v}_1 = T_A(\vec{v}_1) = \lambda_1\vec{v}_1$$

Thus, first column of D is $[\lambda_1\vec{v}_1]_\mathfrak{E}$. Since $\lambda_1\vec{v}_1 = \lambda_1\vec{v}_2 + 0\vec{v}_2 + 0\vec{v}_3 + \cdots + 0\vec{v}_n$ as it is part of an

eigenbasis, $[\lambda_1\vec{v}_1]_\mathfrak{E} = \begin{bmatrix} \lambda_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$. If repeated for all vectors in $\mathfrak{E}$, it ends up being that:

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$$

It must also hold true that if A is diagonalizable, it has an eigenbasis. It is known that:

$$S^{-1}AS = D = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$$

so a basis in $\mathbb{R}^n$ consisting of eigenvectors must be found. This is found in the columns of S, the change of basis matrix. The similarity formula can be rearranged to get $AS = SD$. Thus, for the first column of S, $S\vec{e}_1$:

$$\begin{aligned}
A(S\vec{e}_1) &= (AS)\vec{e}_1 \\
&= (SD)\vec{e}_1 \\
&= S(D\vec{e}_1) \\
&= S(\lambda_1\vec{e}_1) \\
&= \lambda_1(S\vec{e}_1)
\end{aligned}$$

This is indeed a basis as S is required to be invertible.