MATH 115 - Linear Algebra for Engineers

Kevin Carruthers

Fall 2012

Vectors

Vectors have a magnitude and a direction, are denoted by vector arrows, and are said to be in $\mathbb{R}^n$

Two Main Operations

You can perform two main operations with vectors: **vector addition**, which is the algebraic tail to tip addition of vectors, and **scalar multiplication**, which is $t\vec{v}$ where $t \epsilon \mathbb{R}$ and $\vec{v} \epsilon \mathbb{R}^n$

Linear Combinations

The linear combinatio of a set of vectors $(\vec{v_0}, \dots \vec{v_n}$ is any vector which can be obtained from these vectors through vector addition and scalar multiplication. It has the form $a_0\vec{v_0} + \cdots + a_n\vec{v_n}$ where $a_0, \dots a_n \epsilon \mathbb{R}$

Example: for $\vec{a} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$$3\vec{a} - 2\vec{b} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

Any vector in $\mathbb{R}^2$ is a linear combination of this set.

Dot Product

In $\mathbb{R}^n$, the dot product of $\vec{u} = \begin{bmatrix} a_0 \\ \dots \\ a_n \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} b_0 \\ \dots \\ b_n \end{bmatrix}$ is a scalar defined to be

$$\vec{a} \circ \vec{b} = a_0 b_0 + \cdots + a_n b_n$$

thus if $\vec{a} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$

$$\vec{u} \circ \vec{v} = 3 * 2 + 5 * 6 = 36$$

Properties

- $\vec{u} \circ \vec{v} = \vec{v} \circ \vec{u}$
- $t\vec{u} \circ \vec{v} = t(\vec{u} \circ \vec{v})$, for any scalar t
- $\vec{u} \circ (\vec{v} + \vec{w}) = \vec{u} \circ \vec{v} + \vec{u} \circ \vec{w}$

Magnitude

The **magnitude** of a vector is its length.

Example: for $\vec{v} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}$

$$||v|| = \sqrt{3^2 + 4^2} = 5$$

More generally

$$||v|| = \sqrt{\vec{v} \circ \vec{v}}$$

and

$$||v||^2 = \vec{v} \circ \vec{v}$$

Unit Vector

The **unit vector** is a vector of length one. Given $\vec{v}$, the unit vector with the same direction is

$$\frac{\vec{v}}{||v||}$$

This is called **normalization**.

Distance Between Points

For P and Q, the **distance** between them is $||PQ||$

Angle Between Two Vectors

For $\vec{u}$ and $\vec{v}$, the **angle** between them is θ.

Deriving from the **cosine law** $c^2 = a^2 + b^2 - 2ab\cos\theta$, we get

$$\begin{aligned}
||\vec{u}-\vec{v}||^2 &= ||u||^2 + ||v||^2 - 2||u||||v||\cos\theta \\
&= (\vec{u}-\vec{v}) \circ (\vec{u}-\vec{v}) \\
&= \vec{u}\circ\vec{u} - \vec{u}\circ\vec{v} - \vec{v}\circ\vec{u} + \vec{v}\circ\vec{v} \\
&= ||u||^2 - 2\vec{u}\circ\vec{v} + ||v||^2 \\
-2||u||||v||\cos\theta &= -2\vec{u}\circ\vec{v} \\
\cos\theta &= \frac{\vec{u}\circ\vec{v}}{||u||||v||} \\
\theta &= \cos^{-1}\left(\frac{\vec{u}\circ\vec{v}}{||u||||v||}\right)
\end{aligned}$$

Definition: two vectors $\vec{u}$ and $\vec{v}$ are orthogonal if $\vec{u}\circ\vec{v} = 0$

Lines

In $\mathbb{R}^2, y = mx + b$, but in $\mathbb{R}^n$ we must two vectors to represent a line. The equation for a line in $\mathbb{R}^n$ follows the same forms as in $\mathbb{R}^2$, but uses any vector on the line as its intercept, and any vector parallel to the line as its slope.

Example:

$$\vec{x} = (3,1,4,1) + t(2,0,3,2), tin\mathbb{R}$$

Parametric Form

In **parametric form**, we solve for the values of each variable in the resulting vector. For the above equation we have

$$\vec{x} = \begin{cases} x_0 = 3 + 2t \\ x_1 = 1 \\ x_2 = 4 + 3t \\ x_3 = 1 + 2t \end{cases}$$

where $tin\mathbb{R}$

Planes (In $\mathbb{R}^3$ only)

Every plane in $\mathbb{R}^3$ has a **normal vector** orthogonal to the plane. $\vec{n}$ must be orthogonal to $\vec{PX}$, where P and X are both points on the plane, so

$$\vec{n} \circ \vec{PX} = 0$$

For any line in $\mathbb{R}^3$ of the form $ax + by + cz = k$, $\vec{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$

Projections

$\text{proj}_{\vec{u}}(\vec{v})$ is the **projection** of $\vec{v}$ onto $\vec{u}$

1. $\text{proj}_{\vec{u}}(\vec{v})$ is a scalar multiple of $\vec{u}$
2. $\text{proj}_{\vec{u}}(\vec{v})$ and $\vec{v} - \text{proj}_{\vec{u}}(\vec{v})$ are orthogonal

$$\begin{aligned} \text{proj}_{\vec{u}}(\vec{v}) &= \frac{\vec{u} \circ \vec{v}}{||u||^2}\vec{u} \\ \text{perp}_{\vec{u}}(\vec{v}) &= \vec{v} - proj_{\vec{u}}(\vec{v}) \\ \text{perp}_{\vec{u}}(\vec{v}) \circ \text{proj}_{\vec{u}}(\vec{v}) &= 0 \end{aligned}$$

Shortest Distance

The tip of $\text{proj}_{\vec{u}}(\vec{v})$ is the **closest point** to $\vec{v}$

Example: The shortest distance between the point P and the line $\vec{QR}$ is

$$||\text{perp}_{\vec{QR}}(\vec{QP})||$$

Projection onto a Plane

For a point P and a plane containing point Q, the **projection** onto that plane is

$$\text{perp}_{\vec{n}}(\vec{QP})$$

Vector Algebra

The definition of $\mathbb{R}^n$ is $\{(a_0, \dots a_n) \mid a_0, \dots a_n in \mathbb{R}\}$

There are 10 properties of $\mathbb{R}^n$. For any $\vec{x}, \vec{y}, \vec{z} in \mathbb{R}^n$. $s, t in \mathbb{R}$

1. $\vec{x} + \vec{y} in \mathbb{R}^n \leftarrow$ **closure of addition**
2. $\vec{x} + \vec{y} = \vec{y} + \vec{x} \leftarrow$ **commutivity**
3. $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z}) \leftarrow$ **associativity**
4. There exists $\vec{0} in \mathbb{R}^n$ such that $\vec{0} + \vec{x} = \vec{x}$
5. For every $\vec{x} in \mathbb{R}^n$, there exists $-\vec{x} in \mathbb{R}^n$ such that $\vec{x} + (-\vec{x} = \vec{0})$
6. $t\vec{x} in \mathbb{R}^n \leftarrow$ **closure under scalar multiplication**
7. $s(t\vec{x}) = (st)\vec{x}$
8. $(s + t)\vec{x} = s\vec{x} + t\vec{x}$
9. $t(\vec{x} + \vec{y}) = t\vec{x} + t\vec{y}$
10. $1 * \vec{x} = \vec{x}$

Any algebraic structure that satisfies these 10 properties will "act" like $\mathbb{R}^n$ and be called **vector spaces**. We can apply things in $\mathbb{R}^n$ to these structures.

Subspaces

Subspaces are vector spaces within $\mathbb{R}^n$. To check if a subset of $\mathbb{R}^n$ is a **subspace**, we only need to check properties 1, 4, and 6; the other properties are inherited.

Definition: A non-empty subset S of $\mathbb{R}^n$ is a subspace of $\mathbb{R}^n$ if for all $\vec{x}, \vec{y} in S, t in \mathbb{R}$

$$\vec{x} + \vec{y} in S \text{ and } t\vec{x} in S$$

Property 4 follows from the fact that the set is non-empty.

Generally, any line containing $\vec{0}$ is a subspace and any line that does not contain $\vec{0}$ is not a subspace (not closed under scalar multiplication). Generally, any plane through the origin is a subspace.

Spanning Sets

Recall: every element of $\mathbb{R}^n$ in a plane can be written as a linear combination of $\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\}$

Definition: the **span** of $\{\vec{v_0}, \dots \vec{v_k}\}$ is denoted $\text{span}\{\vec{v_0}, \dots \vec{v_k}\}$ is the set of all linear combinations of $\vec{v_0}, \dots \vec{v_k}$

$$\text{span}\{\vec{v_0}, \dots \vec{v_k}\} = \{a_0\vec{v_0}, \dots a_k\vec{v_k} \mid a_0, \dots a_k in \mathbb{R}\}$$

$\text{span}\{\begin{bmatrix}1\\2\end{bmatrix}\}$ is the line through 0 with distance $\begin{bmatrix}1\\2\end{bmatrix}$. If $\vec{x} in \text{span}\{\begin{bmatrix}1\\2\end{bmatrix}\}$, then $\vec{x} = a_0 \begin{bmatrix}1\\2\end{bmatrix}, a_0 in \mathbb{R}$

Theorem 0.1. *If $\vec{v_0}, \dots \vec{v_k} in \mathbb{R}^n$, then $\text{span}\{\vec{v_0}, \dots \vec{v_k}\}$ is a subspace of $\mathbb{R}^n$*

In $\mathbb{R}^2$ we have $\mathbb{R}^2 = \text{span}\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\} = \text{span}\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\}$ because $\begin{bmatrix}1\\1\end{bmatrix}$ is a linear combination of the other two vectors, and is thus redundant.

Theorem 0.2. *If $\vec{v_0}, \dots \vec{v_k} in \mathbb{R}^n$ and $\vec{v_k}$ is a linear combination of $\vec{v_0}, \dots \vec{v_{k-1}}$, then $\text{span}\{\vec{v_0}, \dots \vec{v_k}\} = \text{span}\{\vec{v_0}, \dots \vec{v_{k-1}}\}$*

Linear Independence

Definition: The set $\{\vec{v_0}, \dots \vec{v_k}\}$ is **linearly dependent** if there exists $a_0, \dots a_k$ not all zero such that $a_0\vec{v_0} + \cdots + a_k\vec{v_k} = \vec{0}$. Otherwise (ie, if $a_0, \dots a_k$ are all zero) the set is **linearly independent**.

Basis

Definition: A **basis** B of a subspace S is a linearly independent subset of S such that $S = \text{span}\{B\}$

Systems of Linear Equations

A linear equation is in the form

$$a_0x_0 + \cdots + a_nx_n = b$$

where

$$a_1, \dots a_n, b in \mathbb{R} \text{ and } x_0, \dots x_n in \mathbb{R}^n$$

A solution is a vector $\vec{x}$ that satisfies the equation.

Matrices

A **matrix** A is m x n if it has m rows and n columns. The entry at row i and column j is the ij-th entry, denoted by $(A)_{ij}$ or a_{ij}. When $m = n$, it is a square matrix. The diagonal entries of a square matrix are $a_{11}, a_{22}, \dots$

Reduced Row-Echelon Form (RREF)

A matrix is in **row-echelon form (REF)** if

1. The firt nonzero entry in each row is a **leading one** (1)
2. Rows of zeros are at the bottom of the matrix
3. Each leading one is to the right of the leading ones in all the rows above it

In addition, if each column containing a leading one has 0's everywhere else, then it is in **RREF**.

To get the complete solution from RREF

1. Assign each non-leading zero a parameter
2. Write solutions to leading variables in terms of these parameters

Rank

The **rank** of a matrix is the number of leading ones in its RREF.

Consistency

A system of linear equations is **consistent** if it has at least one solution. Otherwise it is inconsistent.

RREF Facts

1. Every augmented matrix can be reduced to RREF by **elementary row operations** (vector addition, scalar multiplication, and row-swapping)
2. The RREF of an augmented matrix is always unique
3. $\text{rank}\{A\} \leq \min\{m, n\}$ where m and n are the number of rows and columns in the matrix
4. A system is inconsistent if and only if there is a row of the form $\begin{bmatrix} 0 & \dots & 0 & 1 \end{bmatrix}$ in its RREF
5. If a system is consistent, then the number of parameters in the solution set is the number of variables (columns) n - $\text{rank}\{A\}$
6. A consistent system has a unique solution when $n = \text{rank}\{A\}$ (i.e. no parameters)
7. A consistent system has infinitely many solutions when $n > \text{rank}\{A\}$

Homogeneous System

A system is **homogeneous** if all the constants in the right-most column are 0. By taking each var = 0, we get a solution to any homogeneous system. Any homogenous system has either only the trivial solution or infinitely many solutions

To guarantee consistency of a set spanning $\mathbb{R}^n$, the span must contain at least n vectors, where n is the number of rows.

Bases

In $\mathbb{R}^n$ any basis has size n.
Theorem 0.3. *If S is a subspace, then any basis for S has the same size.*

Definition: the dimension of a subspace is the size of its basis.

$$\dim\{S\} = k$$

Example: the dimension of any plane is 2.

Note: $\{\vec{0}\}$ is not a basis for $\{\vec{0}\}$ because it is not linearly independant.

Special Matrices

- In the **Zero Matrix** every entry is 0 and the matrix is denoted by 0 or 0_{mn}
- A matrix is **diagonal** if every off-diagonal entry is 0
- In the **Identity Matrix** every diagonal entry is 1 (the identity matrix is diagonal)
- An **Upper Triangular Matrix** is a square matrix where anything below the diagonal is 0
- A **Lower Triangular Matrix** is a square matrix where anything above the diagonal is 0

Two basic operations

1. **Matrix addition** $\rightarrow$ if A and B are two matrices of the same size, then we define $A + B$ by $(A + B)_{ij} = (A)_{ij} + (B)_{ij}$
2. **Scalar Multiplication** $\rightarrow$ if A is a matrix and t is a scalar, then we define tA as $(tA)_{ij} = t(A)_{ij}$

The set of all m x n matrics together with these two operations satisfy the 10 properties of $\mathbb{R}^n$ (eg. closure of addition, closure of scalar multiplication, commutativity, ...), so this is a vector space.

Transpositions

The **transpose** of a m x n matrix is an n x m matrix where $(A^T)_{ij} = (A)_{ji}$

Properties of a Transpose

1. $(A^T)^T = A$
2. $(kA)^T = k(A^T)$
3. $(A + B)^T = A^T + B^T$

Definition: A square matrix A is symmetric if $A^T = A$. It is skew-symmetric if $A^T = -A$

Matrix Multiplications

Definition: let A be an a by b matrix and B be a b by c matrix. Then AB is an a by c matrix defined by

$$(AB)_{ij} = A_i * B_j$$

Non-comutativity: $AB \neq BA$. Order of multiplication matters.

Cencellation law: If $AC = BC$, $A \neq B$

Properties of Matrix Multiplication

1. If A is m x n, then $IA = AI = A$
2. $A(BC) = (AB)C = ABC$
3. $A(B + C) = AB + AC$
4. $(B + C)A = BA + CA$
5. $k(AB) = (kA)B = A(kB)$
6. $(AB)^T = B^T A^T$

Linear Mappings

A function is a **linear mapping** if for any $\vec{x}, \vec{y} in \mathbb{R}^m$, $f(\vec{x}) in \mathbb{R}^n$ and $t in \mathbb{R}$

1. $f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$
2. $f(t\vec{x}) = tf(\vec{x})$

A function $f(\vec{x}) = [f]\vec{x}$ is a **matrix mapping** if for any $\vec{x} in \mathbb{R}^m$, $[f]_{mxn}$, and $f(\vec{x}) in \mathbb{R}^m$. All matrix mappings are linear mappings and vice-versa.

Solving Mappings

$$\begin{aligned}
f(\vec{x}) &= f(x_0\vec{e_0} + \cdots + x_n\vec{e_n}) \\
&= x_0 f(\vec{e_0}) + \cdots + x_n f(\vec{e_n}) \\
&= (f(\vec{e_0}), \ldots f(\vec{e_n}))(x_0, \ldots x_n)
\end{aligned}$$

The standard matrix of f is $[f] = \begin{bmatrix} f(\vec{e_0}) & \ldots & f(\vec{e_n}) \end{bmatrix}$

If $f : \mathbb{R}^n \to \mathbb{R}^m$ is linear, f is an m x n matrix.

Linear Combinations

Definition: let $f, g : \mathbb{R}^n \to \mathbb{R}^m$ be linear. We define $f + g : \mathbb{R}^n \to \mathbb{R}^m$ by

$$(f + g)(\vec{x}) = f(\vec{x}) + g(\vec{x})$$

Definition: if $tin\mathbb{R}$, we define $tf : \mathbb{R}^n \to \mathbb{R}^m$ by

$$(tf)(\vec{x}) = tf(\vec{x})$$

Note: the set of all linear mappings $f : \mathbb{R}^n \to \mathbb{R}^m$ forms a vector space.

Compositions of Mappings

for $f(x) = \cos x, g(x) = 1 - x^2$

$$g \circ f(x) = 1 - \cos^2 x$$

The **codomain** of f is the same as the domain of g

Definition: let $f : \mathbb{R}^n \to \mathbb{R}^m$, and $g : \mathbb{R}^m \to \mathbb{R}^p$ be linear. We define $g \circ f : \mathbb{R}^n \to \mathbb{R}^p$ by $g \circ f(\vec{x}) = g(f(\vec{x}))$. If f, g are linear, $g \circ f$ is as well

$$\begin{aligned}
g \circ f(\vec{x}) &= g(f(\vec{x})) \\
&= g([f]\vec{x}) \\
&= [g][f]\vec{x} \\
[g \circ f] &= [g][f]
\end{aligned}$$

Suppose we want to take $\vec{x}$, rotate it $\frac{\pi}{4}$ around x_2, then project it onto $x_1 + x_2 + x_3 = 0$. For $f =$ rotation and $g =$ projection, this is

$$g \circ f(\vec{x}) = [g][f]\vec{x}$$

Geometric Mappings

- Rotation in $\mathbb{R}^2$
  - Let $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ be the mapping that rotates the input vector an angle θ counter-clockwise around the origin. This is a linear mapping
- Rotation in $\mathbb{R}^3$
  - Rotate an angle of θ around the x_3-axis in the x_1, x_2 direction. This is a linear mapping
- Reflection over a line in $\mathbb{R}^2$
  - Let $f : \mathbb{R}^2 \to \mathbb{R}^2$. $f(\vec{x}) = \vec{x} - 2\ \mathrm{perp}_{\vec{d}}(\vec{x}) = \mathrm{proj}_{\vec{d}}(\vec{x}) - \mathrm{perp}_{\vec{d}}(\vec{x})$. This is a linear mapping

Bases (again)

For any subspace S of dimension k, any set of k linearly independant vectors in S form a **basis** (ie. span S).

Example: in the plane $P : x + y + x = 0$, which has a dimension of 2, two random linear independant vectors are $\begin{bmatrix} 3 \\ 1 \\ -4 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ -5 \end{bmatrix}$. These vectors form a basis .

Proof: if $\{\vec{v_1}, \dots \vec{v_k}\} in S$ is linearly independant, but not a basis, there must be some $\vec{w} in S$ not in span$\{\vec{v_1}, \dots \vec{v_k}\}$. Then span$\{\vec{v_1}, \dots \vec{v_k}, \vec{w}\}$ is linearly dependant because it contains more then k vectors. However, since $\vec{w}$ is not in the span$\{\vec{v_1}, \dots \vec{v_k}\}$, we know that span$\{\vec{v_1}, \dots \vec{v_k}, \vec{w}\}$ is linearly independant. As this is a contradiction, $\vec{w}$ must not exist, and $\{\vec{v_1}, \dots \vec{v_k}\} in S$ must be a basis.

Inverses

Definition: let A be a square matrix and B be the **inverse** of A such that $BA = I$ and $AB = I$.where $AB = I$, $BA = I$, and B is a unique matrix.

Finding an Inverse

To find A^{-1}, we solve $\left[\begin{array}{c|c} A & I \end{array}\right]$. For any invertible matrix

$$\begin{aligned} A\vec{x} &= \vec{b} \\ \vec{x} &= A^{-1}\vec{b} \end{aligned}$$

Note: This means that if A can be row-reduced to I, then A^{-1} exists.

Properties of an Invertible Matrix

A is invertible if and ony if A has rank n (where A is nxn). If A is invertible, then A^{-1} is invertible and $(A^{-1})^{-1} = A$

If A, B are nxn invertible matrices and $tin\mathbb{R}$, then

- $(tA)^{-1} = \frac{1}{t}A^{-1}$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^T)^{-1} = (A^{-1})^T$

Determinants

Definition: The **determinant** of an nxn matrix A is

$$\text{det}A = \sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}a_{ij}$$

where

$$C_{ij} = (-1)^{i+j}\text{det}A(i,j)$$

For any 2x2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $\text{det}A = ad - bc$

Example: find the determinant of $\begin{bmatrix} 0 & -1 & 3 \\ 3 & 1 & 5 \\ -3 & 2 & 0 \end{bmatrix}$

$$\begin{aligned} \text{det}A &= C_{11}a_{11} + C_{12}a_{12} + C_{13}a_{13} \\ &= 0(1*0 - 5*2) + -(-1)[3*0 - 5*(-3)] + 3[3*2 - 1*(-3)] \\ &= 33 \end{aligned}$$

The determinant of A is denoted $|A|$

Upper Triangular Matrices

The determinant of an **upper triangular matrix** is equal to the product of the numbers along its diagonal.

Row or Column Multiplication

For any matrix A which is equal to the matrix B, except for one row or column which has been **multiplied by** k

$$\det A = k \det B$$

Row Swapping

For any matrix A which is equal to the matrix B, except for one row or column which has been **switched with another**

$$\det A = -\det B$$

Row Addition

Row addition does not change the determinant of a matrix.

Example: $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 121 & 1 \end{vmatrix}$

Invertibility

Theorem 0.4. *A is invertible if and only if $\det A \neq 0$*

Determinant of a Product

$$\det(AB) = \det(A)\det(B)$$

Traces

The **trace** of a square matrix A is the sum of its diagonal values.

Eigenvalues and Eigenvectors

For a square matrix A, a non-zero vector $\vec{v}$ is an **eigenvector** of A if

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

for some constant λ, which is called an **eigenvalue** of A. Eigenvectors are the non-zero solutions to

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

Any solutions of

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

are eigenvalues of A.

Eigenspaces

Definition: the **eigenspace** of an eigenvalue λ for A is the set of all eigenvectors of A with eigenvalue λ. This is a subspace since it is the solution set to the homogeneous system $(A - \lambda I)\vec{v} = 0$

Characteristic Polynomials

Definition: the **characteristic polynomial** of A is

$$\det(A - \lambda I)$$

For any nxn matrix A, the degree of its characteristic polynomial is n. A polynomial with degree n has exactly n roots (including complex and repeating roots). Thus, any nxn matrix has n eigenvalues. A multiplicity of a root is the number of times it appears as a root of its polynomial.

If r is a root, $\lambda - r$ is a factor of the characteristic polynomial.

Diagonalization

To find the powers of matrices, it can be helpful to **diagonalize** them, since the square of a diagonal matrix is equal to that matrix with each of its entries squared, et cetera.

Definition: A square matrix A is diagonalizable if there exists an invertible matrix P and a diagonal matrix D such that

$$P^{-1}AP = D$$

Suppose A is diagonalizable, and $P^{-1}AP = D$. Then

$$A^n = PD^nP^{-1}$$

Finding P and D

The matrix D is a diagonal vector containing the eigenvalues of A. The matrix P is a matrix composed of the eigenvectors of A, in the same order as their corresponding eigenvalues appear in D.

Example: for a 2x2 matrix A with eigenvalues 3 and 1, and corresponding eigenvectors $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ 1 \end{bmatrix}$

$$D = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$
$$P = \begin{bmatrix} 2 & -3 \\ -1 & 1 \end{bmatrix}$$
$$P^{-1} = \begin{bmatrix} -1 & -3 \\ -1 & -2 \end{bmatrix}$$

Recurrence

For the Fibonnaci sequence $1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233$ defined as

$$f(n+1) = f(n) + f(n-1)$$

we can see that

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} f(n) \\ f(n-1) \end{bmatrix} = \begin{bmatrix} f(n+1) \\ f(n) \end{bmatrix}$$

as this **recurrence** is a linear relationship.

Through diagonalization, we can find

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{1-\sqrt{5}}{2\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{1+\sqrt{5}}{2\sqrt{5}} \end{bmatrix}$$

Reintroducing vectors and solving for the nth power gives us

$$\begin{aligned} \begin{bmatrix} f(n+1) \\ f(n) \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \begin{bmatrix} f(1) \\ f(0) \end{bmatrix} \\ &= \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}^n \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{1-\sqrt{5}}{2\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{1+\sqrt{5}}{2\sqrt{5}} \end{bmatrix} \begin{bmatrix} f(1) \\ f(0) \end{bmatrix} \end{aligned}$$

Thus $f(n)$ is the second component of that equation

$$f(n) = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right]$$

Similar Matrices

Definition: Two nxn matrices A and B are **similar** if there exists an invertible P such that $P^{-1}AP = B$, or $A \sim B$. A is diagonalizable if it is similar to a diagonalizable matrix.

Theorem 0.5. *If $A \sim B$, then they have the same determinant, characteristic polynomial, eigenvalues, rank, and trace.*

Theorem 0.6. *If A is diagonalizable, $detA$ is the product of the eigenvalues of A and trA is the sum of the eigenvalues of A.*

Orthogonality

Definition: A set $\{\vec{v_1}, \dots \vec{v_k}\}$ is **orthogonal** if $v_i \neq v_j$ whenever $i \neq j$

Example: the standard basis is an orthogonal set, ie. $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

If $\{\vec{v_1}, \dots \vec{v_k}\}$ is an orthogonal set that does not include $\vec{0}$, then it is linearly independant.

Suppose $\{\vec{v_1}, \dots \vec{v_k}\}$ is an orthogonal basis for a subspace S. Let $\vec{x} inS$. So $\vec{x} = a_1\vec{v_1} + \dots + a_k\vec{v_k}$ for some a_i's. For some i, take the dot product of $\vec{v_i}$ on both sides to get

$$x\vec{v_i} = a_1\vec{v_1} \circ \vec{v_i} + \dots + a_i\vec{v_i} \circ \vec{v_i} + \dots + a_k\vec{v_k} \circ \vec{v_i}$$

$$a_i = \frac{x\vec{v_i}}{||v_i||^2}$$

$$\vec{x} = \frac{xa_1}{||a_i||^2}\vec{v_1} + \dots + \frac{xa_k}{||a_k||^2}\vec{v_k}$$

Example: for $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$

$$\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \frac{7}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{-1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Definition: A set $\{\vec{v_1}, \dots \vec{v_k}\}$ is **orthonormal** if it is orthogonal and $||v_i|| = 1$ for each i (ie. any orthogonal set where each vector has been normalized)

Theorem 0.7. *If $\{\vec{v_1}, \dots \vec{v_k}\}$ is an orthonormal basis for a subspace S and $\vec{k} inS$, then*

$$\vec{x} = (x\vec{v_1})\vec{v_1} + \dots + (x\vec{v_k})\vec{v_k} = kx$$

Orthogonal Matrices

A matrix is **orthogonal** if $A^{-1} = A^T$, or $A^T A = I$. Along it's diagonal, $||v_i|| = 1$, for each i.

If A is orthogonal, so is A^T. Also, each of its rows and columns are **orthonormal**.

Orthogonal Complements

Definiton: let S be the subspace of $\mathbb{R}^n$. The **orthogonal complement** of S, denoted $S^\perp$, is the set of all vectors orthogonal to every vector in S.

$$S^\perp = \left\{ \frac{1}{x} in \mathbb{R}^n \mid \vec{x}\vec{v} = 0 \text{ for all } \vec{v} in S \right\}$$

Example: if P is a plane through the origin

$$P^\perp = \text{span}\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \}$$

$S \cap S^\perp = \{\vec{0}\}$, ie the only vector in both S and $S^\perp$ is $\vec{0}$

Example: let $S = \text{span}\{ \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \}$, find $S^\perp$. For $S \neq S^\perp$

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

$$S^\perp = \text{span}\{ s \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \}$$

Properties of $S^\perp$, $S in \mathbb{R}^n$

1. $S^\perp$ is a subspace (solution set of a homogeneous system)
2. $(S^\perp)^\perp = S$
3. $\dim(S^\perp) = n - \dim(S)$, where dimension is equal to the number of parameters.
4. if $\{v_1, \ldots v_k\}$ is an orthonormal basis for S and $\{x_1, \ldots x_k\}$ is an orthonormal basis for $S^\perp$, then $\{v_1, \ldots v_k, x_1, \ldots x_k\}$ is an orthonormal basis for $\mathbb{R}^n$

Projection onto Subspaces

$$\vec{x} = \text{proj}_{\vec{S}}(\vec{x}) + \text{perp}_{\vec{S^\perp}}(\vec{x})$$

Definition: let $\{v_1, \dots v_k\}$ be an orthonormal basis of the subspace $Sin\mathbb{R}^n$. Then

$$\text{proj}_{\vec{s}}(\vec{x}) = \vec{x}\vec{v_1}\vec{v_1} + \cdots + \vec{x}\vec{v_k}\vec{v_k}$$

If $\{\vec{v_1}, \dots \vec{v_k}\}$ is orthogonal, then

$$\text{proj}_{\vec{s}}(\vec{x}) = \frac{\vec{x}\vec{v_1}}{||v_1||^2}\vec{v_1} + \cdots + \frac{\vec{x}\vec{v_k}}{||v_k||^2}\vec{v_k}$$

Example: find $P = x - 2y + 3z = 0$ projected onto $\vec{x} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$

An orthogonal basis is $\vec{v} = \left\{ \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix} \right\}$ thus

$$\text{proj}_{\vec{p}}(\vec{x}) = \frac{\vec{x}\vec{v_1}}{||v_1||^2}\vec{v_1} = \frac{\vec{x}\vec{v_2}}{||v_2||^2}\vec{v_2}$$

Theorem 0.8. *Let S be a subspace of $\mathbb{R}^n$ and $\vec{x}$ be a vector in $\mathbb{R}^n$. There exists a unique vector $\vec{s}inS$ such that we can find the minimum of $||\vec{x} - \vec{s}||$, and that vector is*

$$\vec{s} = \text{proj}_{\vec{s}}(\vec{x})$$

Gram-Schmidts Procedure

Given a basis $\{w_1, \dots w_k\}$ for a subspace S, **Gram-Schmidtz produce** an orthogonal basis $\{v_1, \dots v_k\}$ for S_i where for each i

$$\text{span}\{w_1, \dots w_i\} = \text{span}\{v_1, \dots v_i\}S_i$$

Suppose $\{w_1, \dots w_k\}$ is a basis for S. We define $S_i = \text{span}\{w_1, \dots w_i\}$. To find an orthogonal basis, we calculate

1. $\vec{v_1} = \vec{w_1}$
2. $\vec{v_n} = \text{perp}_{\vec{S_n}}(\vec{w_n})$ where $\text{perp}_{\vec{S_n}}(\vec{w_n}) = \text{perp}_{\vec{v_n - 1}}(\vec{w_n}) - \cdots - \text{perp}_{\vec{v_1}}(\vec{w_n})$
3. if $\vec{v}inS$, $\vec{v}$ is an orthogonal basis

Application: Line fitting, curve fitting

Orthogonal Diagonaliazation

For a "normal" diagonal, where D is a diagonal matrix and P is invertable $P^{-1}AP = D$

Definition: A matrix A is **orthogonally diagonalizable** if there exists an orthogonal matrix Q and a diagonal matrix D such that

$$Q^{-1}AQ = D$$

We need an orthonormal basis of eigenvectors in $\mathbb{R}^n$

Suppose A is orthogonally diagonalizable. SO $Q^TAQ = D$ or $A = QDQ^T$

$$\begin{aligned} A^T &= (QDQ^T)^T \\ &= (Q^T)^T D^T Q^T \\ &= QDQ^T \\ &= A \end{aligned}$$

thus if a matrix is orthogonally diagonalizable, it must be symmetric.

Theorem 0.9 (Principle Axis Theorem)**.** *If A is symmetric, then A is orthogonally diagonalizable.*

Example: $A = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}$

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 2-\lambda & -2 \\ -2 & 5-\lambda \end{vmatrix} \\ &= (2-\lambda)(5-\lambda) - 4 \\ &= (\lambda - 6)(\lambda - 1) \end{aligned}$$

$$\text{For } \lambda = 6, A - 6I = \begin{bmatrix} -4 & -2 \\ -2 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}, \vec{v_1} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\text{For } \lambda = 1, A - 1I = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}, \vec{v_2} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\begin{aligned} D &= \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix} \\ Q &= \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \end{aligned}$$

Theorem 0.10. *If A is symmetric, the eigenvectors corresponding to distinct eigenvalues are orthogonal.*

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

$$\det(A - \lambda I) = -(\lambda+1)^2(\lambda-2)$$
$$\lambda = -1, -1, 2$$

For

$$\lambda = 2, A - 2I = \begin{bmatrix} -2 & -1 & -1 \\ -1 & -2 & -1 \\ -1 & - & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

an eigenvector is

$$\frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

For

$$\lambda = -1, A + 1I = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

a basis for its eigenspace is

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

Gram-Schmidtz this: $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$

Thus A can be diagonalized by P into D

Note: eigenvectors for distinct eigenvalues are already orthogonal. For eigenvectors of an eigenvalue of high multiplicity, use Grom-Schmidtz to orthogonalize them.

Application: Graphing quadratic equations

$$A \rightarrow 2x_1^2 - 4x_1x_2 + 5x_2^2 = 12$$
$$A = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Through orthogonal diagonalization we get

$$A \rightarrow 6y_1^2 + y_2^2 = 12$$

where y_1 and y_2 are specific vectors, thus we can graph A by plotting these vectors on an $x - y$ graph.

Vector Spaces

Example **vector spaces** (sets obeying the 10 required properties):

- $\mathbb{R}^n$
- The set of all mxn matrices
- The set of all polynomials on x of max degree n
- The set of all continuous functions on $[a, b]$

Subspaces of Vector Spaces

Definition: Let $\mathbb{V}$ be a vector space. A non-empty subset S of $\mathbb{V}$ is a **subspace** if for any $\vec{x}, \vec{y} in S, t in \mathbb{R}$

1. $\vec{x} + \vec{y} in S$
2. $t\vec{x} in S$

Subspaces are vector spaces.
Theorem 0.11. *Let $\mathbb{V}$ be a vector space and let $\{v_1, \ldots v_k\} in \mathbb{V}$. Then the span $\{v_1, \ldots v_k\}$ is a subspace of $\mathbb{V}$*

Linear Independence

Definition: Let $\mathbb{V}$ be a vector space, let $\vec{v_1}, \ldots \vec{v_k} in \mathbb{V}$. Then $\{\vec{v_1}, \ldots \vec{v_k}\}$ is **linearly independent** if the only solution to

$$a_1\vec{v_1} + \cdots + a_k\vec{v_k} = 0$$

is the trivial solution.

Basis and Dimension

Definition: Let $\mathbb{V}$ be a vector space. Then a **basis** is a linearly independant set that spans $\mathbb{V}$.
Theorem 0.12. *If S and T are bases for $\mathbb{V}$, then S and T have the same size.*

Definition: The **dimension** of a vector space is the size of its basis.

Let $\mathbb{V}$ be a vector space of dimension n. Then

1. Any set of more than n vectors is linearly dependant
2. Any set of less than n vectors does not span $\mathbb{V}$
3. Any linearly independant set of n vectors is a basis of v

Finding a Basis

Given $S = \text{span}\{\vec{v_1}, \dots \vec{v_k}\}$ find a basis. If there is linear dependence, throw out dependant vectors until you have independence.

Theorem 0.13. *If $\vec{v_k}$ is a non-trivial linear combination of $\{\vec{v_1}, \dots \vec{v_k}\}$, then*

$$\text{span}\{\vec{v_1}, \dots \vec{v_k}\} = \text{span}\{\vec{v_1}, \dots \vec{v_{k-1}}\}$$

Extending a Basis

Given a basis B of a subspace S of $\mathbb{V}$, extend B to a basis for $\mathbb{V}$.

Theorem 0.14. *If $\vec{v_1}, \dots \vec{v_k}$ is linearly independent and $\vec{v_{k+1}}$ is not in the* $\text{span}\{\vec{v_1}, \dots \vec{v_k}\}$, *then $\{\vec{v_1}, \dots \vec{v_k}, \vec{v_{k+1}}\}$ is still linearly independent.*

Usually, we can consider adding vectors from the standard basis.

Inner Products

A dot product in $\mathbb{R}^n$ gives us the vector's length and orthogonality.

For $C[a, b]$, the **inner product** is $f, g in [a, b]$ and

$$< f, g >= \int_a^b f(x)g(x) \, \mathrm{d}x$$

If $< f, g >= 0$, f and g are orthogonal.

$$||f||^2 =< f, f >= \int_a^b f^2(x) \, \mathrm{d}x$$

To project $f(x)$ onto $\{\sin x, \cos x\}$ we have

$$\frac{< f(x), \sin x >}{||\sin x||^2} \sin x + \frac{< f(x), \cos x >}{||\cos x||^2} \cos x$$

Complex Numbers

Theorem 0.15 (Fundamental Theorem of Algebra)**.** *The polynomial equation*

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$$

where each $a_i in \mathbb{C}$, $a_n \neq 0$ has at least one root in $\mathbb{C}$.

Definition: A **complex number** z in standard form is $z = a + bi$ where $a, bin\mathbb{R}$. The set of all complex numbers is $\mathbb{C} = \{a + bi \mid a, bin\mathbb{R}\}$

The real part of z is a and the imaginary part is b, thus $\mathbb{R}in\mathbb{C}$.

We define two operations:

1. $a + bi + c + di = (a + c) + (b + d)i$
2. $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$

Division in $\mathbb{C}$

FInd the inverse of $a + bi$.

The inverse of $a + bi$ is

$$(a + bi)^{-1} = \frac{a - bi}{a^2 + b^2}$$

where $a - bi$ is the conjugate $\bar{z}$ and $a^2 + b^2$ is the length, squared.

Complex Conjugate

Definition: If $z = a + bi$, the **conjugate** of z is $\bar{z} = a - bi$

Properties of the Conjugate

1. $z \bar{+} w = \bar{z} + \bar{w}$
2. $z\bar{w} = \bar{z}\bar{w}$
3. $\bar{\bar{z}} = z$
4. $z + \bar{z} = 2a$
5. $z - \bar{z} = 2bi$
6. $z\bar{z} = a^2 + b^2$
7. $z^{-1} = \frac{\bar{z}}{z\bar{z}}$

Complex Plane

We can plot complex numbers in the same way as any two-dimensional number, using Re and Im as our axes.

Modulus

The **modulus** of $z = a + bi$ is $|z| = \sqrt{a^2 + b^2}$

1. $|z| \geq 0$, equality only holds when $z = 0$
2. $|\bar{z}| = |z|$
3. $|zw| = |z||w|$
4. $|z + w| \leq |z| + |w|$

Complex Roots

$$\begin{aligned} a + bi &= r(\cos\theta + i\sin\theta) \\ a \bar{+} bi &= r\big(\cos(-\theta) + i\sin(-\theta)\big) \end{aligned}$$

Example:

$$\begin{aligned} (1+i)^{314} &= \sqrt{2}(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}) \\ &= 2^{157}(\cos\frac{314\pi}{4} + i\sin\frac{314\pi}{4}) \\ &= 2^{157}(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}) \\ &= 2^{157}i \end{aligned}$$

Complex Exponentials

For $f(\theta) = (\cos\theta + i\sin\theta)e^{-i\theta}$

$$\begin{aligned} f'(\theta) &= (-\sin\theta + i\cos\theta)e^{-i\theta} + (\cos\theta + i\sin\theta)(-ie^{-i\theta}) \\ &= e^{-i\theta}(-sin\theta + i\cos\theta - i\cos\theta + \sin\theta) \\ &= 0 \end{aligned}$$

Thus $f(\theta) = C$ for some constant C. $f(0) = 1$, thus $C = 1$ and

$$\cos\theta + i\sin\theta = e^{i\theta}$$

we can reduce this to Euler's Formula

$$e^{i\pi} = -1$$

Roots of Complex Numbers

$z^n = a$ where $a in \mathbb{C}$ has n roots (Fundamental Theorem of Algebra states that every polynomial of degree $n \geq 1$ has at least 1 root in $\mathbb{C}$).

Example:

$$\begin{aligned} z^2 &= i \\ \left(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}}\right)^2 &= i \\ e^{i\frac{\pi}{4}} z &= i \\ e^{i\frac{\pi}{2}} &= i \\ i &= i \end{aligned}$$

Since we can do this with $-\frac{1}{\sqrt{2}}$, we have $z = \frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} - i\frac{1}{\sqrt{2}}$

Example:

$$\begin{aligned} z^n &= a \\ z^n &= re^{i\theta} \\ \left(se^{i\phi}\right)^2 &= re^{i\theta} \\ s^n e^{in\phi} &= re^{i\theta} \end{aligned}$$

So $s^n = r$ and $n\phi = \theta + 2\pi\mathbb{Z}$, thus $s = \sqrt[n]{r}$ and $\phi = \frac{\theta + 2\pi(\mathbb{Z}+n)}{n}$. Every nth root of $\mathbb{Z}$ has the same angle, so we only need $k = 0, \ldots n$

Notes:

1. All roots of $z^n = a$ have the same r, so they are on a circle.
2. The roots are equally spaced on the circle. n roots will divide the circle into n equal pieces.
3. The n roots of a are distinct.