MATH 355 All Notes

1 Systems of Linear Equations

2 Questions originate from a system:

- Existence - is there a solution
- Uniqueness - is there one / multiple solutions.

Can be represented as a matrix:

$$x + 2y = 5, 3x - 7y = 41 \rightarrow \begin{bmatrix} 1 & 2 & 5 \\ 3 & -7 & 41 \end{bmatrix}$$

where $\begin{bmatrix} 1 & 2 \\ 3 & -7 \end{bmatrix}$ is the coefficient matrix.

1.1 Elementary Row Operations

3 Types, which are reversible:

- Multiply a row by a non-zero scalar(Scaling)
- Add / subtract multiples of one row to / from another(Replacement)
- Swap positions of two rows(Interchange)

Can be used to solve a system of equations. However, anytime the row $\begin{bmatrix} 0 & 0 & ... & 0 & \mid \textit{not } 0 \end{bmatrix}$ appears, no solution exists.

Matrices in row-reduced echelon form(RREF) means no more row operations are needed. Matrices are in RREF if:

- Any row that is all 0's are at the bottom
- For any row that is not all 0's, the first non-0 entry must equal 1(1's are called pivots)
- If you have a pivot 1, it is the only non-zero entry in its column
- Any pivot in a lower row is to the right of any pivot in a higher row(Pivots go down and right).

All augmented matrices can be reduced to a unique RREF, with the number of rows beng the number of equations, and the number of columns being the number of variables / indeterminants, where basic variables occur when there are pivots and free variables occur when there are non-pivots. For example, for the matrix:

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 2 \end{bmatrix}$$

x_1, x_2 are pivots, so they are basic variables, but x_3 is a non-pivot, so it is a free variable. A linear system has infinitely many solutions if, in the RREF of its augmented matrix:

- There is no row of the form $\begin{bmatrix} 0 & 0 & ... & 0 & \mid \textit{not } 0 \end{bmatrix}$
- There exists a free variable

A linear system has a unique solution if, in the RREF of its augmented matrix:

- There is no row of the form: $\begin{bmatrix} 0 & 0 & ... & 0 & \mid \textit{not } 0 \end{bmatrix}$
- There is no free variable

If the number of variables > the number of equations, then either there is no solution or there are infinitely many solutions.

1.2 Interpreting RREF

A linear system is consistent(it has atleast one solution) if and only if the right most column of the augmented matrix is NOT a pivot column. If the system is consistent then:

- It has a unique solution if there are NO free variables
- It has infinitely many solutions if there is one variable

Can express basic variables in terms of free variables, e.g. $x_1 = t, x_2 = 2t$.

1.3 RREF Algorithm

We can reduce matrices into RREF form by using this algorithm:

- Begin with left-most non-zero column
- Select a pivot in the pivot column and move it to the top
- Use row replacement to create 0s below this pivot
- Move to next column, shortening column, removing top row
- Apply steps 1 - 3, moving onto next column
- Normalize each pivot to 1, eliminating all non-zero entries above.

1.4 Vector Spaces Over a Field

A field is a set of elements where you can do addition and multiplication in a "nice" way. For field β, the following must be true:

- $\forall a, b \in \beta, a + b \in \beta$
- $\forall a, b \in \beta, a + b = b + a$
- $\forall a, b, c \in \beta, (a + b) + c = a + (b + c)$

- $0 \in \beta, 0 + a = a$
- $\forall a \in \beta, \exists - a, a + -a = 0$
- $\forall a, b \in \beta, a \cdot b \in \beta$
- $(a \cdot b) \cdot c) = a \cdot (b \cdot c)$
- $\exists 1 \in \beta, 1 \cdot a = a$
- $\forall a \in \beta, \exists a^{-1}, a \cdot a^{-1} = 1$
- $a(b + c) = ab + ac$

A vector space(V) is a set with two operations:

- Internal Operation(+): $\forall v, w \in V, v + w \in V$
- External Operation($\cdot$): $\forall a \in Field, v \in V, a \cdot v \in V$

We can describe $\mathbf{R}^n$ as the set of ordered $n-$tuples of real numbers, can be described as

$$\mathbf{R}^n = \{ \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \cdots \\ a_n \end{bmatrix} \}$$

Where $a_1, a_2, ...a_n \in \mathbf{R}$. $\mathbf{R}^n$ is a vector space over the scalar field R

2 Linear Combinations

Let $v_1, v_2, ..., v_n \in V_n$ be vectors, a linear combination is $a_1v_1, a_2v_2, ...a_nv_n$ where $a_1, a_2...a_n \in Field$.

We can tell if a vector is a linear combination of other vectors if there is a $c_1v_1 + c_2v_2 = v_3$, where v_1, v_2 are the other vectors and v_3 is a linear combination. In this case, we can create an augmented matrix where each column is a vector, where v_3 is the last column and find the RREF form to find $c_1, c_2, ...$. If no such solution exists, then the vector is not a linear combination.

2.1 Span

Let V be a vector space over field F. For vectors $v_1, v_2, ..v_n \in V$, the span of $v_1, v_2, ...v_n$ is defined to be ALL linear combinations of $v_1, ...v_n$, i.e.:

$$\{c_1v_1 + c_2v_2 + ...c_nv_n | c_i \in F\} \tag{1}$$

Denoted by Span, where 0 is always in any Span

2.2 Linear Combination of Vectors

Let V be a vector span over F. A linear combination of $v_1, v_2, ..v_n \in V$ is $c_1v_1 + c_2v_2 + ... + c_nv_n; c_1, ...c_n \in F$

2.3 Linear Dependence

For any $v_1, ...v_n \in V_1$ if $c_1v_1 + c_2v_2 + ... + c_nv_n = 0$ iff $c_1 = c_2 = ... = c_n = 0$, $v_1, v_2, ...$ are linearly independent, if not, then linearly dependent.

If any vector is 0, then the entire set is linearly dependent as you can multiply the 0 vector by anything and get 0.

We know that for two vectors, v_1, v_2, the vectors are linearly dependent if $v_1 = \frac{-c_2}{c_1}v_2$, or if the vectors are colinear.

If $v_1, ..., v_n \in V$ are linearly dependent, then atleast(but NOT EVERY) one of the vectors can be written as a linear combination of the rest(one of the vectors is in the span of the others).
Let $v_1, ..., v_n \in \mathbf{R}^n$. If $m > n$, then $v_1, ..., v_n$ are linearly dependent.

3 Linear Transformations

A linear transformation between vector spaces V, W is a map $T : V \to W$, i.e. for any $v \in VT(v) \in W$ which satisfies, where V is the domain and W is the codomain, can be described by $T(v) = Av$ where A is the standard matrix.

- $T(v \pm v') = T(v) \pm T(v')$
- $\forall v \in V, c \in F, T(c \cdot v) = c \cdot T(v)$

We can find A by plugging in $e_1, e_2, ...e_n$ into T, where the result of $T(e_i)$ will be the corresponding column of the standard matrix.

Any Linear Transformation $T : \mathbf{R}^n \to \mathbf{R}^m$ can be written as $T(v) = Av$ for some matrix A size $m \times n$. Justified by the fact that since $T : V \to W$, $T(c_1v_1 + c_2v_2 + ... + c_nv_n) = c_1T(v_1) + c_2T(v_2) + ... + c_nT(v_n) \to W$, which can be defined as $Ax = W$

We can then find A by finding the image of the standard basis, essentially finding

$$[T(\begin{bmatrix} 1 \\ 0 \\ ... \\ 0 \end{bmatrix}), T(\begin{bmatrix} 0 \\ 1 \\ ... \\ 0) \end{bmatrix}, ..., T(\begin{bmatrix} 0 \\ 0 \\ ... \\ 1 \end{bmatrix})]$$

Let $T_1 : V \to W$ and $T_2 : W \to U$ be a linear transformation. Then $T_2 \circ T_1 : V \to U$ is a linear transformation.

The set of all Linear Transformations from V to W forms a vector space.

For linear transformation $T : V \to W$, V is the domain of the linear transformaiton, while W is the codomain of the linear transformation,

- $T(v) = w$ for some $v \in \mathbf{R}^n$
- $Av = w$ is consistent
- $w \in$ Span of columns of A

$T : V \to W$ is called onto if $\forall w \in W, \exists v \in V$ S.T. $T(v) = W(Im(T) =$ W, essentially if all of W has a source) or if the columns of A span $\mathbf{R}^n$, or if $Ax = v$ has a solution for every v, or if A row reduces to have a pivot in each row.
$T : \mathbf{R}^n \to \mathbf{R}^m$ cannot be onto if $n < m$

T is called one-to-one(or into) if $\forall w \in W, \exists$ at most one V S.T $T(v) = u$ - all of V maps to a unique value in W. $T\ into \iff$ columns of A are linearly independent, or if $Ax = 0$ has only the $x = 0$ solution or if A row reduces to have a pivot in each column.
$T : \mathbf{R}^n \to \mathbf{R}^m$ cannot be into if $n > m$

$dim(Im(T)) = dim(V) - dim(ker(T))$
Rank is the number of linearly independent columns / rows(same result) in a matrix.

We also know that every map $f : A \to B$ can be written as $g \circ h$, where $h : A \to C, g : C \to B$ such that h is onto and g is 1-1. This is true as we can let $C = Im(f)$.

If T is both onto and $1 - 1$, then T is a bijection. If $T : V \to W$ is a bijection, then V and W have the same dimensions, and there exists a $T^{-1} : W \to V$ where $T \circ T^{-1} : W \to W$ and $T^{-1} \circ T : V \to V$, with both being identity maps. $T^{-1}(W) =$ unique $v \in V$ such that $T(v) = w$

If $T : \mathbf{R}^n \to \mathbf{R}^n$, then T is 1-1, T is onto, T is invertible are all equivalent since same direction, meaning that $Ax = 0$ has no non-trivial solution and that the columns of A span $\mathbf{R}^n$.

$T \circ T \circ T$ is invertible if T is invertible, but $T_1 + T_2$ is not invertible if T_1, T_2 are invertible

The basis is a set of linearly independent vectors $\{v_i | i \in I\}$ such that $\forall v \in V, V = \sum_{i \in I} c_i v_i$ where V is a vector space over F. For this course, I is a finite set such that $\{v_1, v_2, ..., v_n\} n \in \mathbf{N} V = Span\{v_i | i \in 1\} = \sum_{i \in I} c_i v_i$(set of linearly independent vectors that span a vector space, any vector in the space can be represented as a linear combination of the basis).

Dimensions - If V can be spanned by finitely many vectors, $dim(V) = \#\{basis\}$, where $dim(V) =$ the maximum number of linearly independent vectors.

Kernel(T) $= \{v \in V | T(v) = 0\}$, set of all vectors where $T(v) = 0$, where T is one-to-one iff $Ker(T) = 0$.

T is one-to-one $\iff$ columns of A are linearly independent
T is onto $\iff$ columns of A spans $\mathbf{R}^n$

4 Matrix Multiplication

For $A : m \times n, B : n \times l, AB = A(B_1, B_2, ..., B_l) = (AB_1, AB_2, ...AB_l) \to m \times l$, or $AB_{ij} = \sum_{k=1} nA_{ik}B_{kj}$, where we are multiplying corresponding values in the rows of A to the columns of B. For Identity Matrix I_n, which is a matrix of zeros with the left to right diagonal being 1, $AI = A = IA$. Identity Matrices are a special type of diagonal matrix, where all digits are 0 except for the left to right diagonal are 0, where the product of two diagonal matrices is a diagonal matrix whose diagonal is the product of the corresponding values of both matrices.

Additionally, If $A : m \times n, B : n \times p$ and A is the standard matrix for $T : \mathbf{R}^n \to \mathbf{R}^m$ and B is the standard matrix for $S : \mathbf{R}^p \to \mathbf{R}^n$, then AB is the standard matrix for $T \circ S : \mathbf{R}^p \to \mathbf{R}^m$

For Matrix Multiplication:

- $A + B = B + A$
- $A + (B + C) = (A + B) + C$
- $a(bA) = (ab)(A)$
- $a(A + B) = aA + aB$
- $AB \neq BA$
- $A(BC) = (AB)C$

4.1 Multiplying Vectors by Matrices

Result is a vector, essentially a linear combination of the columns of the matrix.

For $A = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 2 & 4 \\ 2 & 4 & -3 & 0 \end{bmatrix}, v = \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix}, Ax = v:$

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = x = \mathbf{R}^4$$

$$a_1 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + a_2 \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix} + a_3 \begin{bmatrix} 0 \\ 2 \\ -3 \end{bmatrix} + +a_4 \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix} \stackrel{?}{=} \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix}$$

Is v a linear combination of the columns of A, can use system of equations given by augmented matrix:

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

4.2 Matrix Inverse

$AB = AC$ does not imply that $B = C$. Cannot always find a matrix A' such that $A' \cdot A = I_n$
If an inverse does exists, there are two possibilities:

- $A' \cdot A = I_n \rightarrow$ Left-Inverse
- $A \cdot A' = I_m \rightarrow$ Right-inverse

Left and Right inverses may not be unique, as the left inverse may not always be equal to the right inverse, but they must have size $n \times m$ if the other matrix is size $m \times n$ However,if A is a square matrix, then the left inverse is also the right inverse of A. If A is a square, then the inverse is unique. A matrix with an inverse is invertible. Additionally, matrix multiplication cannot increase rank.

A matrix is invertible if and only if its determinant is not 0, and all columns of an invertible matrix are linearly independent. Additionally, if a matrix is invertible, the following statements are equivalent:

- A is an invertible matrix
- $A \xrightarrow{RREF} I_n$
- A has n pivot positions
- The equation $Ax = 0$ has only the trivial solution
- The columns of A are linearly independent
- The linear transformation $x \rightarrow Ax$ is one-to-one
- The equation $Ax = b$ has at least one solution for each $b \in R^n$
- The columns of A span $\mathbf{R}^n$
- The linear transformation $x \rightarrow Ax$ maps $\mathbf{R}^n$ onto $\mathbf{R}^n$
- There is an $n \times n$ matrix C such that $CA = I_n$
- There is an $n \times n$ matrix D such that $AD = I_n$
- A^T is invertible

If all matrices are square:

- $(A^{-1} + B^{-1}) \neq (A + B)^{-1}$
- $A^{-1}B^{-1} = (BA)^{-1}$

If $Ax = B$ is a linear system, where A is an $n \times n$ invertible matrix, then $x = A^{-1}B$. Additionally, we know that matrices A, B are similar if there exists a matrix P such that $AP = PB$

4.3 Elementary Matrices

Matrices which, when multiplied with another matrix, produce a elementary row operation. Can be found by applying the corresponding row operation to I_n.

All Elementary Matrices are invertible. If A is an $n \times n$ matrix, it can be reduced to RREF by applying $o_1, o_2, ...o_n$, elementary row operations, or multiplying the matrix by $E_1, E_2, ..., E_n$, which are the corresponding elementary matrices. Therefore $E_n E_{n-1} E_{n-2} ... E_1 A = RREF$, where if A is invertible, $RREF = I_n$.

Can also find that $E_n E_{n-1} ... E_1 = A^{-1}$, meaning that we can find A^{-1} by finding the rref form of $[A|I]$ to $[I|B]$, where B will be A^{-1}. If the statement does not row reduce to I on the left, then A is not invertible.

If A has size 2×2, $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then A is invertible iff $ad - bc \neq 0$. $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

4.4 Matrix Transposes

Transpose of an $m \times n$ matrix A is an $n \times m$ matrix A^T such that $(A^T)_{ij} = A_{ji}$, essentially a reflection across the diagonal.

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

Basis

Definition: A set of Linearly Independent vectors whose span is the entire vector space, or a set of maximal number of linearly independent vectors. Can describe every value in the vector space as a linear combination of the basis.
For example, for $P_n = \{f(x) | deg(fx_1) \leq n\}$, the standard basis is $\{1, x, x^2, ..., x^n\}$.
Remarks: There exists more than 1 set of basis for any non-zero vector space
Every set of basis for a finite dimensional vector space has the same cardinality.

Cardinality of basis = dimension = Number of linearly independent directions

Dimension

Definition: The dimension of V is defined to be the number of vectors in any basis. For example, $dim_{\Re}^{\Re^2} = 2 \implies \{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\} = 2$.

Definition: Let V be a vector space. A vector subspace W is a subset $W \subseteq V$ such that W is itself a vectir space.

Remarks: The dimension of the vector-subspace $\leq$ dimension of the vector space
If $\{v_1, v_2, ..., v_n\}$ is a basis for V, we cannot always find a subset $S' \leq \{v_1, ..., v_n\}$ such that S' is a basis for every vector subspace

Lemmas: $Ker(T) \leq V$ is a vector space
$Im(T) = \{T(v)|v \in V\} \subseteq W$ is a vector space

Rank Nullity Theorem

Let $T \rightarrow V$, which is a linear transformation of finite dimensional vector space. $Ker(T) \xrightarrow{into} V \xrightarrow{onto} Im(T) \xrightarrow{into} W$. $A \xrightarrow{into} B \xrightarrow{onto} C \iff B = A + C$.

Essentially, $Dim(Ker(T)) + Dim(Im(T)) = dim(V)$ or $Dim(Ker(T)) + Rank = dim(V)$, as the rank of T is $dim(Im(T))$ and the nullity of T is $dim(Ker(T))$. To translate into language of matrices:

$T : \Re^n \rightarrow \Re^m$, A_T is the standard matrix. $Im(T) = \{T(v)|v \in \Re^n\} = \{A_T \cdot v|v \in \Re^n\} = \{v_1 A_1 + ... + v_n A_n|v_1, v_2, ..., v_n \in \Re\} = Span\{A_1, ..., A_n\}$, where $A_i = [A_1, ..., A_n\}$.

Rank of any matrix A is the dimension of the Span of $\{A_1, ..., A_n\}$, so for T, rank $A =$ dim(T), where the Image(T) is the span of the columns of A, and a list of columns contains the basis for $Im(T)$, where we keep the pivot columns

$Ker(T) = \{v \in \Re^n|T(v) = 0\} = \{v \in \Re^n|Av = 0\} =$ solution of Av = 0

Dimension of Span $\{A_1, A_2, ...A_n\}$ = dimension of Span of the columns in the RREF of A

Determinant

For 2×2 matrices, the determinant is $ab - cd$, where if $ad - bc = 0$, A is not invertible.

Remark: $|detA|$ is the scaling factor for area under multiplication by A. So if a shape of area 5 is multiplied by matrix A with $detA = 7$, the resulting area would be 35.

Remark: $det(BA) = det(B) \cdot det(A) = det(AB)$.

Propositions: If A is an $m \times n$ matrix, if we:
Switch two columns of A and keep the rest unchanged: $det(A') = -det(A)$
Multiply column a by scalar c: $det(A') = cdet(A)$
Change column to $c_i + ac_j$: $det(A') = det(A)$.

Remark: The Identity matrix has a determinant of 1 and the matrix with the last row which is all 0 has determinant of 0. **Remark:** We can then find the determinant of a matrix by an expansion along a row or column: $det(A) = \sum_{j=1}^{n} a_{ij} c_{ij}$, where c_{ij} is the **cofactor**, which can be found by $c_{ij} = (-1)^{i+j} det(\bar{A_{ij}})$.

Remark: We can also find the determinant of an $n \times n$ matrix where we take all possible products of n entries, one from each row and column. We then attach a sign to each(-1^{i+j}) and then sum them up.

To compute this way, we pick any row or any column, and calculate one of:
$det(A) = a_{1j}c_{1j} + a_{2j}c_{2j} + ... + a_{nj}c_{nj}$
$det(A) = a_{j1}c_{j1} + a_{j2}c_{j2} + ... + a_{jn}c_{jn}$

Remark: The determinant of an upper or lower triangular matrix or a diagonal matrix is the result of the product of the diagonals.

Theorem: $det(A) = det(A^T)$

Theorem: If we are given the matrix $C = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, where A is an $n \times n$ matrix and B is an $m \times m$ matrix, $det(A)det(B) = det(C)$

If a row or column is proportional to another row or column, the determinant is 0, and if a row or column is a linear combination of some other rows / columns, the determinant is 0.

Geometric Interpretation

i. $|det(A)|$ = the scaling factor of A. If a shape defined by vectors is multiplied by a matrix, the determinant of the matrix is the area of the new shape divided by the area of the original shape.

ii. The sign of $det(A)$ determines whether the orientation has changed

Cramer's Rule

If $A = n \times n$ matrix, and $b = n \times 1$ matrix, for linear system $Ax = b$, where A_i is the matrix obtained from A by replacing its ith column by b, $x_i = \frac{det(A_i)}{det(a)}$

Theorem: $A^{-1} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ \vdots & & & \\ \vdots & & & \\ c_{n1} & \dots & \dots & c_{nn} \end{bmatrix}^T$, where $c_{ij} = (-1)^{i+j} det(\bar{A_{ij}})$

Change of Basis

In a coordinate system, $V = (x, y) = x \cdot (1, 0) + y \cdot (0, 1) = xe_1 + ye_2$. Essentially a linear combination of e_1 and e_2 where x, y are weights in a linear combination.

Can other have some other basis, too. Essentially creating new axis. We write the new basis as a linear combination of the original basis and then relate the coefficients to relate the standard basis coefficients to the new basis coefficients, represented by a transition matrix U, where the columns are the new basis vectors

For $V \xrightarrow{T} W$, if we have basis for V and W, we can turn everything to coordinates and get a column of scalars. If we have a basis B for V and a basis C for W. $v \in V$ becomes $[v]_B$, coordinatized relative to B and we can turn $T(v) \in W$ to $[T(v)]_c$, where we can turn $[v]_B$ to $[T(v)]_c$ for some matrix A.

We want $A[v]_B = [T(v)]_c$ for all v in domain, where A is the matrix for T relative to B and C

Building A

If $T : V \to W$, where hte basis of $V = B = \{b_1, b_2, ..., b_n\}$ and basis of $W = C$. A =matrix of T relative to B and C. The ith column of $A = Ae_i = A[b_i]_B = [T(b_i)]_C$.

For example, if $D : P^3 \to P^2$, where $D(f) = f'$. $B = \{t^3, t^2, t, 1\}, C = \{t^2, t, 1\}$

$$A = \begin{bmatrix} | & | & | & | \\ [3t^2]_C & [2t]_C & [1]_C & [0]_C \\ | & | & | & | \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Similarity

$T : V \to V$, with basis B, B' for V, with M, M', the matrix for T with respect to B. $[v]_B \xrightarrow{M} [T(v)]_B$ and $[v]'_B \xrightarrow{M'} [T(v)]_{B'}$, where

$$M = \begin{bmatrix} | & & | \\ [T(b_1)]_B & ... & [T(b_n)]_B \\ | & & | \end{bmatrix}$$

We could turn $[v]_{B'}$ to $[v]_B$ and then convert $[v]_B$ to $[T(v)]_B$ to $T[(v)]_{B'}$, giving us $M' = P_{B' \leftarrow B} M P_{B \leftarrow B'}$, or as $M' = QMQ^{-1}$, where $Q = P_{B' \leftarrow B}$, where M and M' are similar, representing the same linear transformation with respect to different coordinates.

Whenever $D = P^{-1}AP$ and $A = PDP^{-1}$, where A is diagonalizable, where we can find a basis where all vectors, when multiplied by A is just a scalar times that vector.

Eigenvalues and Eigenvectors

Let $L \to W$ is a linear transformation, were A is the square standard matrix.
$L(v) = A_L \cdot [v]_B$, where B is a basis

If $L(v) = \lambda v$, where $\lambda \in F$ is a scalar(non-zero), then

- v is an eigenvector of L
- λ is an eigenvalue of L
- $A - \lambda I$ is not invertible

For an $n \times n$ matrix A, $det(A - \lambda I_n) = p(\lambda)$ is a degree n polynomial, as $a_n \lambda_n + ... + a_1 \lambda + a_0$
Theorem. λ is an eigenvalue of A if and only if $det(A - \lambda I_n) = 0$. To find all eigenvectors / eigenvalues of a given matrix A, we find the characteristic polynomial where $det(A - \lambda I_n) = 0$, the roots of this polynomial are the eigenvalues. To find the eigenvector for eigenvalue λ, find the vector v such that $(A - \lambda I_n)v = 0$(the characteristic polynomial) through converting to RREF form.

If A has complex entries:

$P_a = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} ... (\lambda - \lambda_k)^{n_k}$

- $n_1 + n_2 + ... + n_k = n$

- $P(z) = 0 \to P(\bar{z}) = 0$, where $\bar{z}$ is the complex conjugate of z, so if z is a root, then so is $\bar{z}$

Definition Algebraic Multiplicity describes the power of a certain root in the characteristic polynomial. For example for $(1-\lambda)^2(-2-\lambda) = 0$, $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = -2$, where $\lambda = 1$ has algebraic multiplicity of 2. It is important to know that that characteristic polynomial has **3** roots of $1, 1, 2$, not just 2 roots.

Definition Geometric Multiplicity describes the dimension of the eigenspace of a given λ. It is essentially the number of vectors in the span produced by the solution to $(A - \lambda I_n)v = 0$. More formally, the geometric multiplicity is $dim\{v|Av = \lambda v\}$

Definition Geometric Multiplicity $\leq$ Algebraic Multiplicity

Corollary If A is $n \times n$, then the sum of the geometric multiplicities of its eigenvalues must be $\leq$n, and if the sum $= n$, then A has an eigenbasis and is diagonalizable.

Proposition: For any matrix A, the set $\{v|Av = \lambda v\}$ is a vector subspace for any given eigenvalue λ

Proposition: Eigenvalues of any triangular matrix are precisely its diagonal entries

Definition: The eigenspace of λ is defined to be $E_\lambda = Span\{v|Av = \lambda v\}$, where if v is eigenvector, v is colinear with Av

Theorem Calay-Hamilton Theorem - Any matrix satisfies its own characteristic equation.

Diagonalization

Definition An $n \times n$ matrix A is similar to an $n \times n$ matrix B if there is some invertible matrix P such that $A = P^{-1}BP$. If A is similar to a diagonal matrix, then A is called diagonalizable.

Proposition Similar matrices have the same characteristic polynomials and thus the same eigenvalues. If two matrices have the same eigenvalues, might not be similar.

Proposition If 0 is an eigenvalue of A, then A is not diagonalizable.

Theorem A is diagonalizable if any one of the following conditions hold:

- There exists n linearly independent eigenvectors, $v_1, v_2, ...v_n$
- For each eigenvalue λ_i, its algebraic multiplicity is equal to the geometric multiplicity

Proposition The sum of the algebraic multiplicity of all eigenvectors if n

Proposition If A is $n \times n$, its characteristic polynomial has degree n

Proposition An $n \times n$ matrix has at most n different eigenvalues

Corollary If an $n \times n$ matrix has n distinct eigenvalues, it has an eigenbasis and is therefore diagonalizable

Corollary If an $n \times n$ matrix has n different eigenvalues, then A is diagonalizable.

To diagonalize a matrix A, with eigenvalues $\lambda_1, \lambda_2, ...\lambda_n$ and corresponding eigenvectors $v_1, v_2, ...v_n$, we can describe the diagonalized form of A as $D = P^{-1}AP$, where $P = \begin{bmatrix} v_1 & ... & v_n \end{bmatrix}$ and $D = \begin{bmatrix} \lambda_1 & & \\ & ... & \\ & & \lambda_n \end{bmatrix}$

Proposition If the image has a smaller than the domain, then A cannot be diagonalizable

Proposition If A is a symmetric $n \times n$ real matrix, then all eigenvalues are real numbers

Theorem Fundamental Theorem of Algebra – for polynomial $a_n x^n + a_{n-1} x^{n-1} + ... + a_0$ is a function of real number coefficients, then it has n complex roots counting multiplicity. Any non-real root comes in pairs, $a + bi, a - bi$ for $b \neq 0$

Proposition $\lambda_1, \lambda_2, ...\lambda_k$ are all distinct eigenvalues of A. For each eigenspace $E_{\lambda i}$ has a set of basis $\{V_{\lambda i_1}, V_{\lambda i_1} ... V_{\lambda i_n}\}$, where $\{V_{\lambda i_k} | \forall i, \forall k\}$

Orthogonality

Definition Inner Product - Let $v = \begin{bmatrix} v_1 \\ ... \\ v_n \end{bmatrix}$ and $w = \begin{bmatrix} w_1 \\ ... \\ w_n \end{bmatrix}$ be vectors in R^n. Then $v \cdot w = v_1 w_1 + v_2 w_2 + ... + v_n w_n = v^T w =$ a scalar.

Theorem For $u, v, w \in \mathbf{R}^n$:

- $u \cdot v = v \cdot u$
- $(u + v) \cdot w = u \cdot w + v \cdot w$
- $(cu) \cdot v = c(u \cdot v) = u \cdot (cv)$
- $u \cdot u \geq 0$ and $u \cdot u = 0 \iff u = 0$

In particular $T : R^n \to R^n, T(v) = v \cdot u$ is a linear transformation

Definition For $v \in R^n$, then the length of v is $||v|| = \sqrt{v \cdot v}$

Proposition Let $c \in R$, then $||c \cdot v|| = |c| \cdot ||v||$

Lemma Every non-zero vector v is proportional to a unit vector v' where $v' = \frac{v}{||v||}$

Definition The length of the vector is the distance between a vecotr and the origin, where $dist(u, v) = ||u - v||$

Definition We can also describe the angle between two vectors as $cos(\theta) = \frac{u \cdot v}{||u||||v||}$

Remark If $u \perp v$, $||u + v||^2 = ||u||^2 + ||v||^2$

Definition Orthogonal Complement Let $W \subseteq R^n$ vector subspace. Define a set $W^T = \{v \in \mathbf{R}^n | v$ is $\perp$ to every Vector in $W\}$

Proposition W^T is a vector space

Remark V^T is not always a line, if $u, v \in R^3$, V^T is a plane

Definition We can describe the projection of vector w onto v as $proj_v w = \frac{v \cdot w}{||u||^2} v$

Definition

- A set of vectors is called an orthogonal set if every two distinct vectors are $\perp$
- If this set is a set of basis, this set is the orthogonal basis

Proposition Any set of nonzero, orthogonal vectors are linearly independent

Definition Orthonormal basis is a set of length 1 vectors which are orthogonal to each other.

Theorem A set of n vectors $v_1, ...v_n \in B$ is orthogonal $\iff$ $[v_1...v_n]^T \cdot [v_1...v_n] =$ a diagonal matrix

Proposition If A is an orthonormal matrix, the eigenvalues of A can only be 1 or -1, and if A is orthogonal, then $det(A) = \pm 1$

Definition A unitary matrix is a matrix with complex numbers such that $U(n) = \{A \in M_n C | A^* A = I_n$, where $A^* = \bar{A}^T$

Definition Modulus: $||\lambda|| = \sqrt{a^2 + b^2} = \sqrt{\lambda \cdot \bar{\lambda}}$ for $\lambda = a + bi$

Theorem Eigenvalues of $A \in U(n)$ has a modulus equal to 1

Theorem Let $B = \{v_1, ...v_n\}$ be a set of basis for R^n, and $A = [A_1, A_2, ...A_n]$

- Then if B is orthogonal, then $A^T A$ is diagonal
- If B is orthonormal, then $A^T A = I_n$

Theorem Given $v \in V, W$ a subspace of V and $\{w_1, w_2, ...w_k\}$ an Orthonormal Basis of W, then defining $\hat{v} = \sum_{i=1}^{k} < v, w_i > w_i$, where $proj_w v = \hat{v}$, we have

- $\hat{v} \in W$
- $v - \hat{v}$ is $\perp$ to all vectors in W
- $\hat{v}$ is the only vector satisfying the above two corollaries
- $\hat{v}$ is the closest vector to v in W

Remark If $A^T = A_{-1}$ then A preserves innerproducts, i.e. $< v, w > = < Av, Aw >$ for all $v, w \in R^n$. Therefore A preserves norms $||Av|| = ||v||$ for all $v \in R^n$ **Definition** $B = \{v_1, ...v_n\}$, orthonormal of R^n, and $A = [v_1|v_2|...|v_n]$, where each column is a basis vector, then A is an orthogonal matrix. In addition, I_n is an orthogonal matrix. A matrix A is orthogonal if

- $A^t = A^{-1}$
- $< Av, Aw > = < v, w >$ for all $v, w \in R^n$
- $||Ax = ||x||$ for all $x \in R^n$

Theorem If A is an orthogonal matrix, then $A^T A = I_n$

Theorem If A is an orthogonal matrix, then the absolute value of each eigenvalue must be 1

Gram-Schmidt Process We can produce an orthogonal basis from $B = \{v_1, ..., v_n\}$ where

$$\begin{aligned} v_1' &= v_1 \\ v_2' &= v_2 - proj_{v_1} v_2 \\ v_3' &= v_3 - proj_{v_1} v_3 - proj_{v_2} v_3 \\ &\vdots \\ v_n' &= v_n - proj_{v_i} v_n - ... - proj_{v_{n-i}} v_n \end{aligned}$$

Spectral Theorem

Studying eigenvalues and diagonalization of real symmetric matrices.

Theorem 1 All eigenvalues of a real symmetric matrix are real numbers

Theorem 2 If A is a real symmetric matrix, then the eigenvalues corresponding to different eigenvalues are $\perp$

Theorem 3 Every symmetric real matrix A can be diagonalized by an orthogonal matrix

Proposition A has an orthonormal eigenbasis if and only if A is symmetric

Proposition $v_1, ...v_n$ are eigenvectors with eigenvalues $\lambda_1, ...\lambda_n$. For $\lambda_i \neq \lambda_j, i \neq j$, $v_1...v_n$ are linearly independent.