Propositions and Theorems

Gauss-Jordan elimination Algorithm

Proposition 1

> For any mxn matrix A, there exist an RREF matrix which is a row equivalent to A

Note

If A is an nxn matrix, and B is an echelon form matrix obtained from A by gaussian reduction without applying any scaling operations,
Then, det(A) = $(-1)^k$ det(B)

where k = num of interchange ops applied

Equivalence relation

Let T be an equivalence relation on a set A, then

- **Reflexive:** aTa exists, for every a $\in$ A
- **Symmetric:** if aTb exists then bTa also exists, for every a, b $\in$ A
- **Transitive:** if aTb and bTc exists, then aTc also exists, for every a, b, c $\in$ A

Proposition 2

> Row equivalence is an equivalence relation of the set R^{mxn} of mxn matrices with entries from the set *R*.

Proof

Remark 1

Every equivalence relation includes a partition of the underlying set. The parts of the partition being the equivalence classes, i.e., the equivalence classes are pair-wise disjoint subsets whose union is the whole set.
Conversely, a corresponding equivalent relation exists given any partition of a set.

Proposition 2A

> A matrix can't be row equivalent to 2 RREF matrices
> And 2 RREF matrix can't be row equivalent

Proposition 2B

> Let A and B be two mxn matrices. Then, A is row-equivalent to B **if and only if** the two homogeneous systems A**x**=**0** and B**x**=**0** have the same solution set

Conclusive Remark

Each equivalence class for this equivalence relation contains a distinctive member, i.e., the unique RREF matrix. This fact can be used to determine whether two matrices are row-equivalent to each other

Solving linear system

Obs 1

The two important matrices related to the system A**x**=**b** are the coefficient matrix A and the *augmented matrix* of the system [A:**b**]

Obs 2

If we obtain a row equivalent matrix to either the coefficient matrix or augmented matrix, then the solution set of the system is the same

Homogeneous system

Obs 1

Solution of R**x**=**0** is the same as A**x**=**0**

Obs 2

If the number of non-zero rows r of R is less than the number of variables n, then the system has a non-trivial solution

Obs 3

If the number of non zero rows of R is equal to the number of variables, then there are no free variables, and the system has a unique solution (only the trivial solution of a zero vector)

Proposition 3

> If A is a square matrix, then A is row equivalent to the identity matrix if and only if the homogeneous system A**x**=**0** has only the trivial solution.

Proof

Non-Homogenous system

Proposition 4

> The system is consistent if and only if the rightmost column of R is not a pivot column

Proof

Obs 5

The non-homogeneous system A**x**=**b** can be inconsistent in either of the two cases of the associated homogenous system A**x**=**0** having a unique solution or infinite solution

Vector implementation of solutions

Obs 6

A vector is a solution of the system if and only if it is of the form **u**+x**v**, where **v** is a solution of the associated homogeneous system

> **Note**
>
> In case the homogeneous system has only the trivial solution, then **v**=**0** and the system has only a unique solution **u**; otherwise, we have infinitely many solutions

Geometric interpretation of solution

Obs 7

The solution of a homogeneous system is either the origin only or all the points on a line or plane through the origin

Obs 8

If a non-homogeneous system has even a single solution point, then its entire solution set consists of only that point or the line or the plane through that point which is parallel to the solution of the associated homogeneous system

Invertible matrices

Obs 1

The inverse of matrix A, if it exists, is unique

Obs 2

If A is invertible, then so is A^{-1} and $(A^{-1})^{-1}=A$

Obs 3

If A & B are invertible, then so is AB and $(AB)^{-1}=B^{-1}A^{-1}$

Obs 4

The product of invertible matrices is invertible, and the inverse is the product of the inverses taken in reverse order

Elementary Matrices

Proposition 5

> If e is an elementary row operation and E is the mxm elementary matrix $e(I_m)$, then for every mxn matrix A, e(A)=EA

Proof

Proposition 6

> Every elementary matrix is invertible

Proof

VIT

Theorem 1

The following are equivalent[1] for an mxm square matrix A:

1. A is invertible
2. A is row equivalent to the identity matrix
    - All the column vectors of the matrix are linearly independent
3. The homogeneous system A**x**=**0** has only the trivial solution
4. The system of equations A**x**=**b** has at least one solution for every **b** in R^m
5. det A ≠ 0
6. Col A = R^m
7. rank A = m
8. 0 is not an eigenvalue for A

Proof

Calculation of the inverse of Matrix-I

Corollary T1.1

An invertible sq matrix A is a product of an elementary matrix. Any sequence of row operations that reduce A to I also transform I to A^{-1}

Proof

Note

Form the augmented matrix [A: I] (sometimes known as the enlarged matrix of A) and carry out e-ops until the A part becomes I. The final result is of the form [I: A^{-1}]

Corollary T1.2

If A has a left inverse or a right inverse, then A has an inverse

Proof

Corollary T1.3

Suppose an sq matrix A is factored as a product of sq matrices,
i.e., $A = A_1A_2...A_n$ (**all sq matrices**) with $n \geq 2$,
Then A is invertible if and only if each A_i is invertible.

Proof

Corollary T1.4

The matrix A is invertible if and only if the system of equations A**x**=**b** has a unique solution for each and every vector **b** in R^m

LU Factorization of a matrix

Example

Formal Definition of Vector Space

Axioms

For **u**, **v** ∈ V and c, d ∈ F, the following axioms hold true:

1. Closure under the operations: **u**+**v** ∈ V & c**u** ∈ V
2. The following axioms hold for addition
    1. Associative property
    2. Identity
        - **0**+**u**=**u**+**0**=**u**
    3. Every vector **u** ∈ V has an additive inverse **v** ∈ V such that **u**+**v** = 0
    4. Commutative property
3. The following multiplication property are satisfied
    1. c(**u**+**v**) = c**u**+c**v**

2. (c+d)**u**=c**u**+d**u**
3. c(d**u**)=(cd)**u**
4. 1**u**=**u** (where 1 is the identity unit element of F)

Example of Vector Space

1. The space R^n of n-tuple of size n. These are often referred to as (column) vectors. The base field is R

2. ***The space R^{mxn} of mxn matrix with real entries. Again with base field R***

    Remark
    The fact that matrices form a vector space is of fundamental importance in both the theory and application of linear algebra

3. ***The space C[0, 1] of continuous function from the closed interval [0, 1] on the real line of R, i.e.,***

    C[0, 1]= {f: f is a continuous function, f:[0,1] -> R}

4. ***The space R^∞ of real sequences is a vector space over R, i.e.,***

    R^∞ = {$\langle a_n \rangle$:$\langle a_n \rangle$ is a sequence with real numbers}

    Of more interest than R^∞ itself is c, the subset of convergent sequences. It is also a vector space.

5. The space $R_n[t]$ of the polynomials of degree ≤ n with real coefficients.
6. The space R[t] of all polynomial with all real coefficients.

Consequences of Vector Space Axioms

Proposition 7

> Let V be a vector space. Then:
>
> 1. The zero vector is unique
> 2. The additive inverse of any vector **u** is unique; We use the notation -**u** for the inverse vector
> 3. 0**u** = **0** for every vector **u**
> 4. c**0** = **0** for every scalar c

> 5. -**u** = (-1)**u** for every vector **u**

Proof

Subspace

Let V be a vector space over the field F. A (vector) subspace of V is a non-empty subset W of V, which is a vector space over F with the operations of vector additions and scalar multiplication taken from V.

Test for subspaces

Proposition 8

> A subset W of V is a subset if and only if it satisfies the following three properties
>
> 1. The zero vector **0** is in W
> 2. W is closed under addition. That is, for each **u** & **v** ∊ W, the sum **u**+**v** ∊ W
> 3. W is closed under scalar multiplication. That is, for each **u** ∊ W, and each scalar c, the scalar product c**u** ∊ W

Proposition 9

> A **non-empty** subset W of V is a subspace if and only if for each **u** and **v** ∊ W, and each scalar c, the sum c**u**+**v** ∊ W.

Span of a set of vector

Definition 1

A linear combination of a finitely many given vectors in any sum of a scalar multiple of the vectors.

Definition 2

Let S = {$\mathbf{v_1}, \ldots, \mathbf{v_p}$} be a finite set of vectors in a vector space V. Then the span of S is the set of all the vectors that can be written as linear combinations of the vectors $\mathbf{v_1}, \ldots, \mathbf{v_p}$

Proposition 10

> If S = {$\mathbf{v_1}, \ldots, \mathbf{v_p}$} is a set of vectors in a vector space V, then Span S = Span{$\mathbf{v_1}, \ldots, \mathbf{v_p}$} is a subspace of V.

Proof

Corollary 10.1

Let V be a vector space

1. If U and W are two subspaces of V, then U∩W is also a subspace of V.
2. If S = {$\mathbf{v_1}, \ldots, \mathbf{v_p}$} is a set of vectors in a vector space V, then Span S is the smallest subspace which contains S, i.e., if W is a subspace such that S⊆W, then Span S⊆W

Remark

In terms of this, Span S is sometimes described as the intersection of all subspaces of V containing S.

Linear Dependence

Let $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_p}$ be a finite list of vectors in a vector space V. Then the vectors are said to be (linearly) dependent if there exist scalars $c_1, c_2, \ldots, c_p$, not all zeros, such that

$c_1\mathbf{v_1} + \ldots + c_p\mathbf{v_p} = \mathbf{0}$

Consequences of the Definition

- **Remark 1:** Any list which contains the **0** has to be linearly dependent
- **Remark 2:** A single non-zero vector is linearly independent
- **Remark 3:** A list of two non-zero vectors is linearly dependent if and only if one vector is a scalar multiple of the other.

- **Remark 4:** A list of non-zero vectors is linearly dependent if and only if at least one of the vectors is a linear combination of the others
- **Remark 5:** Consequently, any list which contains a repeated vector must be linearly dependent. A list that is linearly independent corresponds to a set
- **Remark 6:** Any list which contains a linearly dependent list is linearly dependent
- **Remark 7:** Any subset of a linearly independent set is linearly independent.

Sequences

Proposition S1

Suppose $<a_n>$ -> L_1 and $<b_n>$ -> L_2 and $c \in R$, then:

1. $<a_n>+<b_n>$ -> L_1+L_2
2. $c<a_n>$ -> cL_1
3. $<a_n><b_n>$ -> L_1L_2
4. If $a_n \neq 0$ for all n and $L_1 \neq 0$ then
   - $<1/a_n>$ -> $1/L_1$

Determinants

Definition

If $A \in F^{2x2}$ where $A = [a_{ij}]$, then det A is defined to be the scalar $a_{11}a_{22} - a_{12}a_{21}$.

If $A \in F^{nxn}$, let $A_{i,j}$ donate the (n-1)x(n-1) matrix obtained from A by omission of the i^{th} and j^{th} column.

Expansion Formulae

Column

$$detA = \sum_{i=1}^{n} a_{ij} detA_{i,j}$$

Row

$$det A = \sum_{j=1}^{n} a_{ij} det A_{i,j}$$

Proposition D1

> The following hold for the determinant of a sq. matrix A
>
> 1. If the matrix A' is obtained from A by interchanging two rows. then det A' = det A
> 2. If the matrix A' is obtained from A by multiplying some row by $\lambda \in F$, then det A' = λ det A
> 3. If the matrix A' is obtained from A by adding a multiple of one row to another, then det A' = det A

Procedure for computing the determinant

Proposition D2

> If a nxn matrix A is upper triangular, then det A = $a_{11}a_{22}...a_{nn}$

Corollary D2.1

In order to determine the determinant of a nxn matrix, use elementary row ops of interchange and replacement only to reduce A to an upper triangular matrix A'. If r is the number of row interchange carried out, then det A = $(-1)^r$ det A'

Further Properties of the determinant

Proposition D3

> A nxn matrix A is invertible if and only if det A ≠ 0
> => VIT

Proposition D4

> For all A, B $\in F^{nxn}$, det (AB) = (det A)(det B)

Corollary D4.1

If A is invertible, then det A^{-1} = $(\det A)^{-1}$

Proposition D5

> For all A $\in F^{nxn}$, det A^T = det A

Cramer's Rule

Definition

For any nxn matrix A and any vector **b** in R^n, define A_i(**b**) to be the matrix obtained by replacing the i^{th} column of A by **b**.

Proposition D6

> Let A be any invertible nxn matrix. For any vector **b** in R^n, The unique solution **x** of A**x**=**b** has entries given by:
> $\mathbf{x_i}$ = (det A_i(**b**))/(det A) for i=1,2,...,n

Application of Cramer's law

Terminology

For any nxn matrix A, we define the cofactor $C_{ij} = (-1)^{i+j}$ det A_{ij}

Definition

The adjoint of A is the matrix whose entries are the cofactor of A transposed.

Proposition D7

> Let A be any invertible nxn matrix, then
> A^{-1} = 1/(det A) adj(A)

Application of determinant to area and volume

Proposition D8

> 1. If A is a 2x2 matrix, then the area of the parallelogram determined by the column of A is |det A|
> 2. If A is a 3x3 matrix, then the volume of the parallelepiped determined by the column of A is |det A|

Proposition D9

> 1. Let T:R^2->R^2 be the linear transformation determined by a 2x2 matrix A. If S is a parallelogram in R^2, then
>   {area of T(S)} = |det A| {area of S}
> 2. Let T:R^3->R^3 be the linear transformation determined by a 3x3 matrix A. If S is a parallelepiped in R^3, then
>   {volume of T(S)} = |det A| {volume of S}

Proposition D10

> The conclusion of proposition D9 holds whenever S is a region in R^2 with finite area or a region in R^3 with finite volume.

In other words
{area or volume of T(S)} = |det A|{area or volume of S}

Basis and Dimension

Definitions

- A basis for a vector space V is a linearly independent set S of vectors such that V = Span S
- A space V that has a (finite) basis is finite-dimensional.
- A space that does not have a finite basis said to be infinite dimensional[2]

Infinite-dimensional space

The space R[t] is infinite-dimensional

Proposition 11

> B = $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p\}$ is a basis of the vector space V, if and only if every vector $\mathbf{v} \in V$ is uniquely expressible as a linear combination of the elements of B

Proposition 12

Steinitz Exchange Lemma

> Suppose $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ are linearly independent vectors in a vector space V, and suppose V = Span $\{\mathbf{w}_1, ..., \mathbf{w}_m\}$, then:
>
> 1. $n \leq m$
> 2. V = Span $\{\mathbf{v}_1, ..., \mathbf{v}_n, \mathbf{w}_{n+1}, ..., \mathbf{w}_m\}$, after some reordering the w's if necessary.

Proof

Proposition 13

> If V is a finite-dimensional vector space, then any two bases of V have the same number of elements.

Proof

Dimension

The dimension of a finite-dimensional space is the number of elements in a basis for V. This is written as dim V.

How to create Basis

Proposition 14

> Suppose S = {$\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$} is a linearly independent set in a vector space V. Suppose **v** is a vector which is not in Span S, then the set obtained by adjoining **v** to S is linearly independent

Proof

Proposition 15

> Any linearly independent set S of a finite-dimensional vector space can, if required be expanded to a basis by adjoining vectors.

Proposition 16

> Any finite spanning set S in a non-zero vector space can, if required, be contracted to a basis by deleting (some) vector(s)

Proposition 17

> Let V be a non-zero finite-dimensional vector span with dimension n. Then:
>
> 1. Any linearly independent set of vectors must have ≤ n vectors. If a linearly independent set has n vectors, then it must be a basis, i.e., it must also be spanning set of V.
> 2. Any spanning set of V that have ≥ n vectors must be linearly dependent. If a spanning set has n vectors, it must be a basis, i.e., it must also be linearly independent.

Proof

Dimension of Subspace

Proposition 18

> If W is a proper subspace of a finite dimensional space V, then W is also finite dimensional
> $0 < \dim W < \dim V$

Proof

Sum of Subspaces

Definition

Let U and W be subspaces of the vector space V. Then the sum of U and W,
$U+W = \{\mathbf{u} + \mathbf{w}: \mathbf{u} \in U, \mathbf{w} \in W\}$

U + W is the smallest subspace of V containing both U and W.

Proposition 19

> If U and W are finite dimensional subspace of the vector space V, then
> $\dim (U+W) = \dim U + \dim W - \dim (U \cap W)$

Proof

Direct Sum

Definition

V is said to be the direct sum of subspaces U and W if every vector $\mathbf{v} \in V$ is uniquely expressible in the form $\mathbf{v} = \mathbf{u} + \mathbf{w}$, where $\mathbf{u} \in U$ and $\mathbf{w} \in W$. We shall use the notation $V=U\oplus W$ to indicate that V is the direct sum of U and W.

Proposition 20

> If U and W are subspaces of the vector space of the vector V, then V = U ⊕ W if and only if V=U+W and U ∩ W = {**0**}

Proof

Corollary 20.1

If V is the direct sum of the finite dimensional subspaces U and W, then dim V = dim (U⊕W) = dim U + dim W

Three important subspaces

Null Space

The null space of an mxn matrix A, written Nul A, is the set of all solutions of the homogeneous system A**x**=**0**

Proposition 21

> The null space of an mxn matrix A is a subspace in R^n. Or equivalently, the set of all the solutions of a homogeneous system of m equation variables is a subspace in R^n
> Nul A = {**x** ∈ R^n: A**x** = **0**}

Column space

The column space of a mxn matrix A, written Col A, is the set of all linear combinations of the columns of A, i.e., the span of the column vectors obtained from A.

If A = [$\mathbf{a_1}, ..., \mathbf{a_n}$] then Col A = Span {$\mathbf{a_1}, ..., \mathbf{a_n}$}

Proposition 22

> Col A is a subspace in R^m
> Col A = {**b** ∈ R^m: **b**=A**x** for some **x** in R^n}

Proposition 23

> The pivot columns of a matrix A form the basis of Col A

Note: We take the columns of A for the basis and not the RREF

Proof

Row space

The row space of an mxn matrix, written Row A, is the set of all linear combinations of the rows of A

Proposition 24

> Row A is a subspace in R^n

Proposition 25

> Row equivalent matrices have the same row space

Difference between Nul A and Col A

Rank Theorem

If A is an mxn matrix, then the column rank of A is defines as dim(Col A). Similarly, for row rank and nullity, dim (Row A) and dim (Nul A) resp.

Theorem 2

1. The row rank and column rank of a matrix A are equal. This number is called rank of A
2. The rank of A is equal to the number of pivot positions in the RREF matrix obtained by A.
3. rank(A) + nullity(A) = n = number of columns of A

Proof

Corollary T2.1

A square mxm matrix is invertible if and only if rank(A) = m

Linear transformations

A map or function T: V -> W from a vector space V to a vector space W is said to be linear transformation if:

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V$
2. $T(c\mathbf{u}) = cT(\mathbf{u})$ for all $\mathbf{u} \in V$ and all scalar $c \in F$

Proposition 26(a)

> A linear transformation T:V -> W is completely determined by its action on a basis of V; in other words, by its values on the vectors in the basis of V.

Proposition 26(b)

> Conversely, given a basis $B = \{\mathbf{v_1}, ..., \mathbf{v_n}\}$ of V, and a list of n vectors $\mathbf{w_1}, ..., \mathbf{w_n}$ (not necessarily distinct) in the co-domain space W, there is a unique linear transformation T such that $T(\mathbf{v_1})=\mathbf{w_1}$, ..., $T(\mathbf{v_n})=\mathbf{w_n}$

Fundamental spaces for a linear transformation

There are two important subspaces associated with any linear transformation:

1. The kernel of T, Ker T = $\{\mathbf{v} \in V: T\mathbf{v}=\mathbf{0} \in W\}$ is a subspace of V
2. The range of T, Range T = $\{\mathbf{w} \in W: \mathbf{w}=T\mathbf{v}$ for some $\mathbf{v} \in V\}$

Remark: T is injective if and only if Ker T = $\{\mathbf{0}\}$

Rank of a linear transformation

Let T:V->W be a linear transformation, then rank of T is the dimension of range of T

Let T:V->W be a linear transformation, then nullity of T is the dimension of Ker T(if it is finite dimensional)

Rank theorem for linear transformation

Theorem 3

Suppose that T:V->W is a linear transformation and V is finite dimensional, then
rank(T) + nullity(T) = dim V

Proof

Isomorphism of Vector Space

A linear transformation T:V->W is said to be an isomorphism if it is injective (one - one) and surjective (onto).

Two isomorphic vectors are shown as V $\cong$ W

Proposition 27

> Two finite dimensional vector spaces V and W are isomorphic if and only if dim V = dim W

A very important linear transformation

Left Multiplication by a matrix: Let A be a fixed mxn matrix then function T_A: R^n->R^n defined by T(**x**) = A**x** is a linear transformation

Nul A = Ker T_A
Col A = Range T_A

Coordinate System

Observation 1

Given a basis for a finite dimensional vector space V, we recall that a vector can be expressed in one and only one way as a linear combination of the basis vectors

Definition

An ordered basis for a finite-dimensional space V is a finite sequence of vector which is linearly independent and spans V

Coordinate mapping

Observation 2

Given a fixed ordered basis B for a finite-dimensional vector space V, we can set up a correspondence between the vectors of V and n-tuple F^n

We see that coordinate mapping is an isomorphism from an n-dimensional vector space V over a field F to F^n

The matrix of a linear transformation

Let B = $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ be an ordered basis for V (dim V = n)
and C = $\{\mathbf{w_1}, \ldots, \mathbf{w_m}\}$ be an ordered basis for W (dim W = m)
and T:V->W be a linear transformation

Express each $T\mathbf{v_i} = A_{1i}\mathbf{w_1} + \ldots + A_{mi}\mathbf{w_m}$

We now form the mxn matrix A with the coefficients as columns
Example

The matrix A is called the matrix of T wrt the basis B and C.

$A = [T]_{B\to C}$

$T(\mathbf{v}) = A\mathbf{v}$

Change of Basis

Proposition 28

> Let $B = \{\mathbf{u_1, ..., u_n}\}$ and $C = \{\mathbf{v_1, ..., v_n}\}$ be two ordered basis of a vector space V. Then there is an invertible nxn matrix P such that $[\mathbf{x}]_C = P[\mathbf{x}]_B$

Proof

Finding change of base matrix

Example

Similarity of matrices

An nxn matrix B is said to be similar to an nxn matrix A if and only if there exist an invertible matrix P such that $B = PAP^{-1}$

Proposition 29

> Similarity of matrices is an equivalent relation on F^{nxn}, i.e., the set of nxn matrices with entries from the field F

Proof

Effect of change of Basis

Proposition 30

> Suppose A and B are the matrices of the linear operator T relative to the ordered bases α and β respectively.
> Then A and B are similar matrices, in fact
> $B = PAP^{-1}$, where $P = P_{\alpha\text{->}\beta}$ is the change of basis matrix

Proof

Algebra of linear transformations

Proposition 31

> 1. The set W^V is a set of all functions from V to W is a vector space over F
> 2. The set of all linear transformations(i.e. L(V, W)) from V to W is a subspace of W^V

Proof

Proposition 32

> Let V, W and Z be vector spaces over a field F.
> Let T be a linear transformation from V into W
> And U be a linear transformation from W into Z
> Then the composition function UT from V into Z defined by (UT)(**v**) = U(T(**v**)) for all **v** in V, is a linear transformation from V into Z

> Product (composition) of two linear operator is also a linear operator

Proof

Linear operator

Linear transformation of a vector space V into itself, i.e., space L(V, V) is a linear operator

Multiplication of linear operators satisfies:

1. IU = UI = U
2. $(T_1T_2)T_3 = T_1(T_2T_3)$

3. $U(T_1 + T_2) = UT_1 + UT_2$
4. $(T_1 + T_2)U = T_1U + T_2U$
5. $c(UT) = (cU)T = U(cT)$
6. however this multiplication is not commutative

Another fundamental isomorphism

Proposition 33

> Let V be a n-dimensional vector space over F and
> Let W be an m-dimensional vector space over F.
> Then there is an isomorphism between L(V,W) and F^{mxn}

Proof

Proposition 34

> If dim V = n and dim W = m, then dim (L(V,W)) = mn

Proof

Proposition 35

> Suppose T and U are linear operators on a finite dimensional vector space V and β to a fixed basis for V.
> Then $[UT]_\beta = [U]_\beta\ [T]_\beta$

Proposition 35a

> The mapping ϕ: L(V,V) -> F^{nxn} given by $\phi(T)=[T]_\beta$ is a vector space isomorphism which also preserves product, i.e.,
> $\phi(UT)= \phi(U)\phi(T)$

Proposition 35b

> Suppose that dim V = n, dim W = m, dim Z = k
> UT:V->Z would be a linear transformation from a space of dimension n to a space of dimension k, i.e., its matrix would be a kxn matrix

Proof

Eigenvectors and Eigenvalues

An eigenvector of an nxn matrix A is a non-zero vector **x** such that A**x** = λ**x**, for some scalar λ.
The scalar λ is called an eigenvalue of A if there is a non-trivial solution of A**x** = λ**x** such a vector **x** is called an eigenvector corresponding to λ

An eigenvector is not unique, since all scalar multiples of an eigenvector are also eigenvectors

Set of all eigenvectors corresponding to a fixed eigenvalue λ of the nxn matrix A together with the zero vector forms a subspace of $V=F^n$
X= {**v**: A**v**=λ**v**}
=> X is a eigenspace of A corresponding to λ

Fundamental results for Eigenvector

Proposition 36

> If $\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_p}$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, ..., \lambda_p$ of the matrix A, then the set $\{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_p}\}$ are linearly independent

Proof

Corollary 36.1:

An nxn matrix A can have at most n distinct eigenvalues

Proposition 37

> A scalar λ is an eigenvalue of an nxn matrix if and only if λ satisfies the characteristics equation det (A- λI) = 0

Proof

Eigenvalues of similar matrices

Proposition 38

> If the nxn matrices A and B are similar, then they have the same characteristic polynomial, and hence the same eigenvalue with the same multiplicities

However, the eigenvectors are not necessarily the same

Proof

Cayley-Hamilton theorem

If p is polynomial given by $p(t) = a_0 + a_1 t + ... + a_m t^m$ and A is a square matrix,

then $p(A) = a_0 I + a_1 A + ... + a_m A^m$

The square matrix A is said to satisfy the polynomial p(t) if $p(A) = \mathbf{0}$, i.e., the zero matrix

Theorem 4

Let q donate the characteristic polynomial of nxn square matrix A. Then $q(A) = \mathbf{0}$

Diagonalization of matrices

If A is a diagonal matrix, then its diagonal elements are its eigenvalues and the standard basis vectors are its eigenvectors

Definition

An nxn matrix A is said to be diagonalizable if A is similar to a diagonal matrix D,

in other words $A=PDP^{-1}$

Diagonalization Theorem

Theorem 5

1. An nxn matrix A is diagonalizable if and only if A has n linearly independent eigenvectors
2. In this case, $A=PDP^{-1}$, where the columns of P are linearly independent eigenvectors of A and the diagonal entries of D are eigenvalues corresponding to these eigenvector

Case 1

An nxn matrix A has n distinct (real) eigenvalues Then we get:

Proposition 39

> An nxn matrix A with n distinct eigenvalues is diagonalizable

Example of case 1

Two preliminary definition

Given an eigenvalues λ_1 for a matrix A we define:

- The algebraic multiplicity of λ_1 is the power of the factor (λ-λ_1) in the characteristic polynomial of A
- The geometric multiplicity of λ_1 is the dimension of the eigenspace corresponding to λ_1

Case 2

An nxn matrix A has p < n distinct eigenvalues, but counting the (algebraic) multiplicity, there are n real eigenvalues (not distinct)

Proposition 40

> Let A be an nxn matrix with n(real) eigenvalues (counting algebraic multiplicity) of which λ_1, ..., λ_p are distinct (p<n). Then the following holds true:

> 1. For 1≤k≤p, the geometric multiplicity of λ_k is less than or equal to the algebraic multiplicity of λ_k
> 2. A is diagonalizable, if and only if the geometric multiplicity for each λ_k equal to its algebraic multiplicity
> 3. If A is diagonalizable and B_k is a basis for the eigenspace corresponding to λ_k for each k, then the total collection of vectors in B_1, ..., B_p forms an eigenvectors basis for R^n

Example of case 2

The determinant of a linear operator

Let T:V->V be a linear operator where V is a vector space of finite dimension n. Let α be any basis for V, and let A be the matrix of T wrt the (ordered) basis α, then det T = det A

Eigenvalues of Linear operators

An eigenvector of a linear operator T:V->V is a non zero vector **v** such that T**v**=λ**v** for some scalar λ. Such a scalar is called the eigenvalue of the operator

Inner Product

An inner product on a (real) vector space V is a function, that to each pair of vector **u** and **v** in V associates a scalar (real number) <**u, v**> and satisfies the following axioms for all vectors **u, v, w** in V and all scalars c in *R*

1. <**u, v**> = <**v, u**>
2. <**u + v, w**> = <**u, w**> + <**v, w**>
3. <c**u, v**> = c<**u, v**>
4. <**u, u**> ≥ 0, <**u, u**> = 0 if and only if **u** = **0**

Example

Length and distance in Inner product spaces

The length or norm of any vector **u** in V is the non negative number $||\mathbf{u}|| = \sqrt{<\mathbf{u}, \mathbf{u}>}$

Orthogonality

Two vectors **u, v** in V are said to be orthogonal to each other if $<\mathbf{u}, \mathbf{v}> = 0$; written as $\mathbf{u}\perp\mathbf{v}$

A set of vector $\{\mathbf{u_1}, ..., \mathbf{u_p}\}$ is said to be orthogonal if any two distinct vectors in the set are orthogonal to each other

Proposition 41

> An orthogonal set of non zero vectors in V is linearly independent

Proof

Definition

If W is a subspace of V, then a vector **v** is said to be orthogonal to W if **v** is orthogonal to every vector in W.

The set of all vectors orthogonal to W is called the orthogonal complement of W, written $W^\perp$

$W^\perp = \{\mathbf{v}\in V: \mathbf{v}\perp\mathbf{w} \text{ for every } \mathbf{w} \in W\}$

Proposition 42

> 1. **v** belongs to $W^\perp$ if and only if **v** is orthogonal to every vector in a spanning set for W.
> 2. $W^\perp$ is a subspace of V and $W\cap W^\perp = \{\mathbf{0}\}$

Proof

Orthogonal Basis

An orthogonal basis for a subspace W is a basis which is also an orthogonal set

Proposition 43

> Let {$\mathbf{u_1}, ..., \mathbf{u_p}$} be an orthogonal basis for a subspace W. Then if $\mathbf{y} = c_1\mathbf{u_1} + ... + c_p\mathbf{u_p}$ is any vector in W, we have:
> $c_j = \langle \mathbf{y}, \mathbf{u_j} \rangle / \langle \mathbf{u_j}, \mathbf{u_j} \rangle$, for j = 1, ..., p

Proof

The gram-Schmidt Process

Theorem 6

Given a basis {$\mathbf{x_1}, ..., \mathbf{x_p}$} for a subspace W of V, we can generate an orthogonal basis {$\mathbf{v_1}, ..., \mathbf{v_p}$} for W such that Span {$\mathbf{v_1}, ..., \mathbf{v_k}$} = Span {$\mathbf{x_1}, ..., \mathbf{x_k}$} for k = 1, ..., p

$\mathbf{v_1} = \mathbf{x_1}$

$\mathbf{v_2} = \mathbf{x_2} - (\langle \mathbf{x_2}, \mathbf{v_1} \rangle / \langle \mathbf{v_1}, \mathbf{v_1} \rangle)\mathbf{v_1}$

.

.

.

$\mathbf{v_p} = \mathbf{x_p} - (\langle \mathbf{x_p}, \mathbf{v_1} \rangle / \langle \mathbf{v_1}, \mathbf{v_1} \rangle)\mathbf{v_1} - ... - (\langle \mathbf{x_p}, \mathbf{v_{p-1}} \rangle / \langle \mathbf{v_{p-1}}, \mathbf{v_{p-1}} \rangle)\mathbf{v_{p-1}}$

Proof

Orthogonal Decomposition

Theorem 7

> Let W be any finite dimensional subspace of V.
> Then each vector **y** in V can be written uniquely in the form $\mathbf{y} = \mathbf{y}^{\wedge} + \mathbf{z}$, where $\mathbf{y}^{\wedge} \in W$ and $\mathbf{z} \in W^{\perp}$
>
> Given any finite dimensional subspace W of V, then we can express $V = W+W^{\perp}$, with $W \cap W^{\perp} = \{\mathbf{0}\}$,
> i.e., $V=W \oplus W^{\perp}$

Proof
Example

Proposition 44

Pythagorean Theorem

> **u, v** are orthogonal to each other if and only if
> $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$

Proposition 45

Best approximation theorem

> Let W be any finite dimensional subspace of V, **y** be any vector in V, and $\mathbf{y}^{\wedge}$ be the orthogonal projection of **y** onto W. Then,
> $\|\mathbf{y}-\mathbf{y}^{\wedge}\| < \|\mathbf{y}-\mathbf{v}\|$ for all **v** in W distinct from $\mathbf{y}^{\wedge}$

Corollary 45.1

If **y** is any vector and W is a finite dimensional subspace then
$\|\mathrm{proj}_W\, \mathbf{y}\| \leq \|\mathbf{y}\|$

Proposition 46

The Cauchy Schwarz Inequality

> For all **u, v** in V,
> $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq ||\mathbf{u}|| \; ||\mathbf{v}||$

Proof

Proposition 47

Triangular Inequality

> For all **u, v** in V,
> $||\mathbf{u} + \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||$

Diagonalization of symmetric matrices

Definition

A matrix A is said to be symmetric if $A = A^T$.
A symmetric matrix is necessarily square

Proposition 48

> If A is symmetric then any two eigenvectors from different eigenspace are orthogonal

Proof

Orthogonal matrices

Definition

A square matrix P is said to be orthogonal if its columns are orthonormal

Proposition 49

> A orthogonal matrix is necessarily invertible and
> $P^{-1} = P^T$

Proof

Definition

A square matrix A is said to be orthogonally diagonalizable if there is an orthogonal matrix P and diagonal matrix D such that
$A = PDP^{-1} = PDP^T$

Proposition 50

> If an nxn matrix A is orthogonally diagonalizable, then A is symmetric

Proof

Definition

The set of eigenvalues of a matrix A is called the spectrum of A

Spectral theorem for symmetric matrices

Theorem 8

An nxn symmetric matrix A has the following properties:

1. The eigenspace are mutually orthogonal
2. A has n real eigenvalues, counting (algebraic) multiplicities
3. A is orthogonally diagonalizable
4. The dimension of the eigenspace for each eigenvalue of λ equals the multiplicity of λ, i.e., geometric multiplicity is equal to algebraic multiplicity

Corollary T8.1

A is orthogonally diagonalizable if and only if A is symmetric

Singular Value Decomposition

Observation 1

Let A be an mxn matrix. Then A^TA, being a symmetric nxn matrix, can be orthogonally diagonalizable.

Let $\{\mathbf{v_1}, ..., \mathbf{v_n}\}$ be an orthonormal basis for R^n consisting of eigenvectors of A^TA with corresponding eigenvalues λ_1, ..., λ_n. Then:
$||A\mathbf{v_i}||^2 = \lambda_i$

=> all eigenvalues of matrix A^TA are non-negative
=> Eigenvalues can be rearranged so that: $\lambda_1 \geq ... \geq \lambda_n \geq 0$

Proposition 51

> Suppose that $\{\mathbf{v_1}, ..., \mathbf{v_n}\}$ is an orthonormal basis for R^n consisting of eigenvectors of A^TA with corresponding eigenvalues arranged so that $\lambda_1 \geq ... \geq \lambda_n \geq 0$.
> Suppose that A has r nonzero singular values then $\{A\mathbf{v_1}, ..., A\mathbf{v_r}\}$ is an orthogonal basis for Col A and rank A = r

Proof

Definition

Let A be a mxn matrix.
The singular values of A are the square roots of the eigenvalues of A^TA, denoted by $\sigma_1, \sigma_2, ..., \sigma_n$ arranged in descending order, i.e., $\sigma_i = \sqrt{\lambda_i}$ for i = 1, ..., n

Note: The singular values are lengths of the vectors $A\mathbf{v_1}, A\mathbf{v_2}, ..., A\mathbf{v_n}$

Theorem 9

Let A be a mxn matrix with rank r. Then A can be factored as a product
$A_{mxn} = U_{mxm} \Sigma_{mxn} V_{nxn}^T$

Σ is an mxn matrix consisting an rxr diagonal matrix D with the r non-zero singular values of A, $\sigma_1 \geq ... \geq \sigma_r \geq 0$, along the main diagonal.
D is places in the upper left hand side corner of Σ. Remaining entries of Σ are zero.

U is an mxm orthogonal matrix and V is an nxn orthogonal matrix

1. The matrix V has as it columns the orthogonal basis {$\mathbf{v_1, ..., v_n}$} of the eigenvectors of A^TA
2. In order to obtain U, we take the r vectors A$\mathbf{v_i}$ corresponding to the non zero singular values, extend then to an orthogonal basis of R^m using the gram Schmidt Process and finally normalize the vectors to obtain an orthogonal basis {$\mathbf{u_1, ..., u_m}$}. U has the vectors $\mathbf{u_i}$ as the columns.

Example

1. If any condition is satisfied, then all conditions are satisfied ↩
2. Note that this is a negative definition; nothing is said about the existence of the basis ↩