MA1101R Year 19/20 Sem 1 Summary Notes

Made by Ivan Tan (sarzorwyn) on github

Elementary Row Operations

Multiply by scalar: cR_i

$$\det(B) = c\,det(A)$$

Add a Row multiple: $R_i + cR_j$

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

Swap 2 rows: $R_i \Leftrightarrow R_j$

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

Matrices

$A^2B^2 \neq (AB)^2$

$(A+B)^T = A^T + B^T$ (same size)

$(AB)^T = B^TA^T$

$A \cdot B = A\,B^T = A^TB$

A^{-1} is a unique inverse of A

If B is an inverse of A, $A^{-1} = B$i

If AB is invertible, $(AB)^{-1} = B^{-1}A^{-1}$

If A is singular, AB and BA are singular. (same size)

If A and B are row equivalent,

$$E_1E_2 \dots E_nA = B$$

Where $E_1E_2 \dots E_n$ are elementary row matrices.

Likewise, $A = E_1^{-1}E_2^{-1} \dots E_n^{-1}B$

$(B|d) = E(A|b) = (EA|Eb)$

Determinant

$\det(A) = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}$ (cofactor expansion along the ith row)

$$= a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}$$

(cofactor expansion along the jth row)

$$\det(A) = a_{11}\det(M_{11})$$

$$\det(A) = a_{11}A_{ij} + a_{12}A_{12} + \cdots + a_{1n}A_{1n}$$

$$\det(EA) = \det(E)\det(A)$$

$\det(B) = (a_{ji} + ka_{ij})A_{j1} + (a_{j2} + ka_{i2})A_{j2} + \cdots + (a_{jn} + ka_{in})A_{jn}$

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

$\det(cA) = c^n\det(A)$

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

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

Cramer's Rule:

$$x = \frac{1}{det(A)}\begin{pmatrix} det(A_1) \\ det(A_2) \\ \vdots \\ det(A_n) \end{pmatrix}$$

$$x = A^{-1}b = \frac{1}{det(A)}[adj(A)]b$$

$$x_i = \frac{b_1A_1i + b_2A_2i + \cdots + b_nA_ni}{det(A)} = \frac{det(A_i)}{det(A)}$$

Vectors

n-vector: $\boldsymbol{u} = (u_1, u_2, \dots, u_n)$

Addition:

$\boldsymbol{u} + \boldsymbol{v} = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n)$

Subtraction:

$\boldsymbol{u} - \boldsymbol{v} = (u_1 - v_1, u_2 - v_2, \dots, u_n - v_n)$

Scalar multiple: $c \in \mathbb{R}$

$c\boldsymbol{n} = (cu_1, cu_2, \dots, cu_n)$

Zero vector: $\boldsymbol{0} = (0,0,\dots,0)$

Solution space in $\mathbb{R}^2$: (lines)

$\{(x,y) | ax + by = c\}$ (explicit)

$\left\{\left(\frac{c-bt}{a}, t \middle| t \in \mathbb{R}\right)\right\}$ if $a \neq 0$ or

$\left\{\left(t, \frac{c-at}{b} \middle| t \in \mathbb{R}\right)\right\}$ if $b \neq 0$

Solution space in $\mathbb{R}^3$: (planes)

$\{(x,y,z) | ax + by + cz = d\}$ (explicit)

$\left\{\left(\frac{d-bs-ct}{a}, s, t \middle| s,t \in \mathbb{R}\right)\right\}$ if $a \neq 0$ or

$\left\{\left(s, \frac{d-as-ct}{b}, t \middle| s,t \in \mathbb{R}\right)\right\}$ if $b \neq 0$ or

$\left\{\left(s, t, \frac{d-as-bt}{c} \middle| s,t \in \mathbb{R}\right)\right\}$ if $c \neq 0$

Line in $\mathbb{R}^3$:

$\{(a_0 + at, b_0 + bt, c_0 + ct \,|t \in \mathbb{R}\}$ =

$\{(a_0, b_0, c_0) + t(a,b,c) | t \in \mathbb{R}\}$

Linear combinations

Let $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}$ be vectors in $\mathbb{R}^n$

$$c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_n\boldsymbol{u_k}$$

is a linear combination of $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}$

Let $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}\}$ be a set of vectors in $\mathbb{R}^n$.

The set of all linear combinations of $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}$ is

$$\{c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_n\boldsymbol{u_k} | c_1, c_2, \dots, c_k \in \mathbb{R}\}$$

$$= span(S) = span\{\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}\}$$

which is also the linear span of S (or linear span of $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}$)

If k < n, then S cannot span $\mathbb{R}^n$.

$\boldsymbol{0} \in span(S)$

For any $\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_k} \in span(S)$ and $c_1, c_2, \dots, c_r \in \mathbb{R}$,

$c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \cdots + c_n\boldsymbol{v_k} \in span(S)$

Let $S_1 = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}\}$ and

$S_2 = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_m}\}$ be subsets of $\mathbb{R}^n$

$$span(S_1) \subseteq span(S_2)$$

if and only if each $\boldsymbol{u_i}$ is a linear combination of $\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_m}$.

To prove 2 sets A and B are equal, show that $A \subseteq B$ and $B \subseteq A$

i.e. show that $span\{\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}\} \subseteq span\{\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_m}\}$

If $\boldsymbol{u_k}$ is a linear combination of $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_{k-1}}$, then

$span\{\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_{k-1}}, \boldsymbol{u_k}\} = span\{\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_{k-1}}\}$

$\boldsymbol{u_k}$ is the 'redundant' vector.

Subspaces

Let V be a subset of $\mathbb{R}^n$. V is called a *subspace* of $\mathbb{R}^n$ if $V = span(S)$ where $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}\}$ for some vectors $(\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k})$. S spans (or $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}$ *span*) the subspace V.

V is a subspace of $\mathbb{R}^n$ <u>if and only if</u> for all $\boldsymbol{u}, \boldsymbol{v} \in V$ and $c, d \in \mathbb{R}$, $c\boldsymbol{u} + d\boldsymbol{v} \in V$

Zero space: $set\{\boldsymbol{0}\} = span\{\boldsymbol{0}\}$

$\boldsymbol{e_1} = (1,0,\dots,0)$, $\boldsymbol{e_2} = (0,1,0,\dots,0)$, …,

$\boldsymbol{e_n} = (0,\dots,0,1)$

Any vector $\boldsymbol{u} = (u_1, u_2, \dots, u_n) \in \mathbb{R}^n$ can be expressed as:

$$\boldsymbol{u} = u_1\boldsymbol{e_1} + u_2\boldsymbol{e_2} + \cdots + u_n\boldsymbol{e_n}$$

Example of subspaces of $\mathbb{R}^2$:

a) $\{\boldsymbol{0}\}$ given by span{$\boldsymbol{0}$})
b) Lines through the origin (given by span{**u**} for nonzero $\boldsymbol{u} \in \mathbb{R}^2$)
c) $\mathbb{R}^2$(given by span{(1,0),(0,1)})

Example of subspaces of $\mathbb{R}^3$:

a) $\{\boldsymbol{0}\}$ given by span{$\boldsymbol{0}$})
b) Lines through the origin (given by span{**u**} for nonzero $\boldsymbol{u} \in \mathbb{R}^3$)
c) Planes containing the origin (given by $span\{\boldsymbol{u}, \boldsymbol{v}\}$ for nonzero $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^3$ which are not parallel to each other)
d) $\mathbb{R}^3$(given by span{(1,0,0),(0,1,0),(0,0,1)})

The solution set of a homogeneous system (Ax = b) of linear equation in *n* variables is a subspace of $\mathbb{R}^n$

Linear Independence

Let $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}\}$ consider the equation:

$$c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k} = \boldsymbol{0}$$

The solution $c_1 = 0, c_2 = 0, \dots, c_k = 0$ is called the trivial solution.

S is a <u>*linearly independent*</u> *set* and $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}$ are *linearly independent* if there only is the trivial solution. No vector in S can be written as a linear combination of other vectors in S.

S is a <u>*linearly dependent*</u> *set* and $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}$ are *linearly dependent* if there is non-trivial solutions i.e. there exist real numbers $a_1, a_2, \dots, a_k$ not all them are zero, such that

$$a_1\boldsymbol{u_1} + a_2\boldsymbol{u_2} + \cdots + a_k\boldsymbol{u_k} = \boldsymbol{0}$$

At least one 'redundant' vector $\boldsymbol{u_i}$ can be written as a linear combination of other vectors in S.

$$\boldsymbol{u_i} = a_1\boldsymbol{u_1} + a_2\boldsymbol{u_2} + \cdots + a_k\boldsymbol{u_k}$$

If number of vectors $k > n$ in $\mathbb{R}^n$ space, then V is linearly dependent.

Let $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}$ be a set of linearly independent vectors in $\mathbb{R}^n$. If $\boldsymbol{u_{k+1}}$is a vector in $\mathbb{R}^n$ and not a linear combination of $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}$, then $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}, \boldsymbol{u_{k+1}}$ are linearly independent.

Bases

V is called a vector space if $V = \mathbb{R}^n$ or V is a subspace of $\mathbb{R}^n$. Let W be a vector space. V is a subspace of W if V is a vector space contained in W.

Let S be a subset of vector space V. S is called a *basis* (plural *bases*) for V if

1. S is linearly independent
2. S spans V

A basis for vector space V contains the smallest possible number of vectors that can span V.

The empty set ∅ is the basis for the zero space

Except for the zero base, any vector space and infinitely many different bases.

If S = $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}\}$ is a basis for vector space V, every vector $\boldsymbol{v}$ in V can be expressed in the form

$$\boldsymbol{v} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k}$$

in exactly 1 way where $c_1, c_2, \dots, c_k \in \mathbb{R}$

The coefficients $c_1, c_2, \dots, c_k$ are called the <u>*coordinates*</u> of $\boldsymbol{v}$ relative to the basis S.

$$(\boldsymbol{v})_S = (c_1, c_2, \dots, c_k) \in \mathbb{R}^k$$

$(\boldsymbol{v})_S$ is called the <u>*coordinate vector*</u> of $\boldsymbol{v}$ relative to basis *S*.

For any $\boldsymbol{u}, \boldsymbol{v} \in V$,

$\boldsymbol{u} = \boldsymbol{v}$ if and only if $(\boldsymbol{u})_S = (\boldsymbol{v})_S$

For any $\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_r} \in V$ and $c_1, \dots, c_r \in \mathbb{R}$,

$$\boxed{\begin{aligned}&(c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \cdots + c_r\boldsymbol{v_r})_S = \\ &c_1(\boldsymbol{v_1})_S + c_2(\boldsymbol{v_2})_S + \cdots + c_r(\boldsymbol{v_r})_S\end{aligned}}$$

Let $|S| = k$, $\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_k}$ be vectors in V.

1. $\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_k}$ are linearly independent vectors in V if and only if $(\boldsymbol{v_1})_S, (\boldsymbol{v_2})_S, \dots, (\boldsymbol{v_r})_S$ are linearly independent vectors in $\mathbb{R}^k$
2. $span\{\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_r}\} = V$ if and only if $span\{(\boldsymbol{v_1})_S, (\boldsymbol{v_2})_S, \dots, (\boldsymbol{v_r})_S\} = \mathbb{R}^k$

Dimensions

Let V be a vector space with basis of k vectors

1. Any subset of V with more than k vectors is always linearly dependent
2. Any subset of V with less than k vectors cannot span V

All bases for a vector space have the same number of vectors.

The *dimension* of a vector space V, dim(V), is the number of vectors in a basis for V. The dimension of zero space is defined to be 0.

$$\dim(\mathbb{R}^n) = n$$

Let V be a vector space, dim(V) = k, $S \subseteq V$

The following are equivalent:

1. S is a basis for V
2. S is linearly independent and $|S| = k$
3. S spans V and $|S| = k$

Let U be a subspace of a vector space V.

$\dim(U) \leq \dim(V)$ and if $U \neq V, \dim(U) < \dim(V)$

Transition Matrices

For $\boldsymbol{v} \in V$ and $\boldsymbol{v} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k}$

Coordinate vectors:

1. Row form: $(\boldsymbol{v})_S = (c_1, c_2, \ldots, c_k)$
2. Column form: $[v]_S = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$

Let S = $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$ and T be two bases for a vector space.

$$\boldsymbol{P} = ([\boldsymbol{u_1}]_T[\boldsymbol{u_2}]_T \ldots [\boldsymbol{u_k}]_T)$$

The square matrix $\boldsymbol{P}$ is the transition matrix from S to T.

The transition matrix $\boldsymbol{P}$ is invertible and $\boldsymbol{P^{-1}}$ is the transition matrix from T to S.

$$[\boldsymbol{w}]_T = \boldsymbol{P}[\boldsymbol{w}]_S$$
$$\boldsymbol{P^{-1}}[\boldsymbol{w}]_T = [\boldsymbol{w}]_S$$

Row Spaces and Column Spaces

Let $\boldsymbol{A} = (a_{ij})$ be a *m* x *n* matrix.

The row space of $\boldsymbol{A}$ is the subspace of $\mathbb{R}^n$ spanned by the rows of $\boldsymbol{A}$. Let $\boldsymbol{r_1}, \boldsymbol{r_2}, \ldots, \boldsymbol{r_m}$ be m rows of $\boldsymbol{A}$.

$$\begin{aligned} r_1 &= (a_{11} \quad a_{12} \quad \cdots \quad a_{1n}) \\ r_2 &= (a_{21} \quad a_{22} \quad \cdots \quad a_{2n}) \\ &\vdots \\ r_m &= (a_{m1} \quad a_{m2} \quad \cdots \quad a_{mn}) \end{aligned}$$

The column space of $\boldsymbol{A}$ is the subspace of $\mathbb{R}^n$ spanned by the columns of $\boldsymbol{A}$. Let $\boldsymbol{c_1}, \boldsymbol{c_2}, \ldots, \boldsymbol{c_n}$ be n rows of $\boldsymbol{A}$.

$$\boldsymbol{c_1} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \boldsymbol{c_2} = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \ldots, \boldsymbol{c_n} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

column space of $\boldsymbol{A}$ = $\{\boldsymbol{Au} \mid \boldsymbol{u} \in \mathbb{R}^n\}$

row space of $\boldsymbol{A}$ = column space of $\boldsymbol{A}^T$

column space of $\boldsymbol{A}$ = row space of $\boldsymbol{A}^T$

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are row equivalent matrices,

row space of $\boldsymbol{A}$ = row space of $\boldsymbol{B}$

i.e. elementary row operations preserve the row space of a matrix

Let $\boldsymbol{R}$ be a row-echelon form of $\boldsymbol{A}$. The set of nonzero rows (pivots) in $\boldsymbol{R}$ is a basis for the row space of $\boldsymbol{A}$.

Elementary row operations may not preserve the column space of a matrix.

A basis for the column space of $\boldsymbol{A}$ can be obtained by taking the columns of $\boldsymbol{A}$ that correspond to the pivot columns in $\boldsymbol{R}$.

A given set of columns of $\boldsymbol{A}$ is linearly independent/form a basis for the column space of $\boldsymbol{A}$ if and only if the set of corresponding columns of $\boldsymbol{B}$ is linearly independent/form a basis for the column space of $\boldsymbol{B}$.

Ranks

The row space and column space of a matrix has the same dimension.

The *rank* of a matrix is the dimension of its row/column space. rank($\boldsymbol{A}$) = non-zero rows/pivot columns in rref($\boldsymbol{A}$).

For a *m* x *n* matrix $\boldsymbol{A}$, $rank(\boldsymbol{A}) \le \min\{m, n\}$

If $\operatorname{rank}(\boldsymbol{A}) = \min\{m, n\}$, $\boldsymbol{A}$ has full rank.

A square matrix has full rank $\Leftrightarrow \det(\boldsymbol{A}) \neq 0$

$rank(\boldsymbol{A}) = rank(\boldsymbol{A}^T)$

$\boldsymbol{Ax} = \boldsymbol{b}$ consistent $\Leftrightarrow rank(\boldsymbol{A}) = rank(\boldsymbol{A}|\boldsymbol{b})$

$rank(\boldsymbol{AB}) \le min\{rank(\boldsymbol{A}), rank(\boldsymbol{B})\}$

Nullspaces and Nullities

Solution space $\boldsymbol{Ax} = \boldsymbol{0}$ is the nullspace of $\boldsymbol{A}$.

Dimension of nullspace = nullity($\boldsymbol{A}$)

$\text{nullity}(\boldsymbol{A}) \le n$ since nullspace is subspace of $\mathbb{R}^n$

$$rank(\boldsymbol{A}) + nullity(\boldsymbol{A}) = n$$

If $\boldsymbol{Ax} = \boldsymbol{b}$ has a solution $\boldsymbol{v}$, solution set is:

$M = \{\boldsymbol{u} + \boldsymbol{v} | \boldsymbol{u}$ is a element of nullspace of $\boldsymbol{A}\}$

$\boldsymbol{x} =$ (a general solution for $\boldsymbol{Ax} = \boldsymbol{0}$) + (one particular solution to $\boldsymbol{Ax} = \boldsymbol{0}$)

A consistent linear system $\boldsymbol{Ax} = \boldsymbol{0}$ has one solution if and only if the nullspace of $\boldsymbol{A}$ is equal to $\{\boldsymbol{0}\}$

Dot Product

$$\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{uv}^T = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

norm: $||\boldsymbol{u}|| = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$

distance between $\boldsymbol{u}$ and $\boldsymbol{v}$: $d(\boldsymbol{u} \cdot \boldsymbol{v}) = ||\boldsymbol{u} - \boldsymbol{v}|| = \sqrt{(u_1 - v_1)^2 + \cdots + (u_n - v_n)^2}$

angle between $\boldsymbol{u}$ and $\boldsymbol{v}$: $\theta = \cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{||\boldsymbol{u}||||\boldsymbol{v}||}\right)$

$(c\boldsymbol{u}) \cdot \boldsymbol{v} = \boldsymbol{u} \cdot (c\boldsymbol{v}) = c(\boldsymbol{u} \cdot \boldsymbol{v})$

$(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w} \mid \boldsymbol{w} \cdot (\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{w} \cdot \boldsymbol{u} + \boldsymbol{w} \cdot \boldsymbol{v}$

$\boldsymbol{u} \cdot \boldsymbol{u} \ge 0$; $\boldsymbol{u} \cdot \boldsymbol{u} = 0$ if and only if $\boldsymbol{u} = \boldsymbol{0}$

Orthogonal and Orthonormal Bases

$\boldsymbol{u}$ and $\boldsymbol{v}$ are *orthogonal* if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$

$\cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{||\boldsymbol{u}||||\boldsymbol{v}||}\right) = \cos^{-1}(0) = \frac{\pi}{2}$

Set S of vectors are *orthogonal* if every pair of distinct vectors are orthogonal and *orthonormal* if all vectors are unit vectors.

S is linearly independent if it is orthogonal.

A basis S for a vector space is called *orthogonal*/*orthonormal* if S is orthogonal/orthonormal.

If S = $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$ is an orthogonal basis for a vector space V, then for any vector $\boldsymbol{w}$ in V:

$$\boldsymbol{w} = \frac{\boldsymbol{w} \cdot \boldsymbol{u_1}}{\boldsymbol{u_1} \cdot \boldsymbol{u_1}}\boldsymbol{u_1} + \frac{\boldsymbol{w} \cdot \boldsymbol{u_2}}{\boldsymbol{u_2} \cdot \boldsymbol{u_2}}\boldsymbol{u_2} + \cdots + \frac{\boldsymbol{w} \cdot \boldsymbol{u_n}}{\boldsymbol{u_n} \cdot \boldsymbol{u_n}}\boldsymbol{u_n}$$

$$(\boldsymbol{w})_S = \left(\frac{\boldsymbol{w} \cdot \boldsymbol{u_1}}{\boldsymbol{u_1} \cdot \boldsymbol{u_1}}, \frac{\boldsymbol{w} \cdot \boldsymbol{u_2}}{\boldsymbol{u_2} \cdot \boldsymbol{u_2}}, \ldots, \frac{\boldsymbol{w} \cdot \boldsymbol{u_n}}{\boldsymbol{u_n} \cdot \boldsymbol{u_n}}\right)$$

Where $\frac{\boldsymbol{w} \cdot \boldsymbol{u_i}}{\boldsymbol{u_i} \cdot \boldsymbol{u_i}}$ represents the coordinates in the base.

If T = $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_k}\}$ is an orthonormal basis for a vector space V, then for any vector $\boldsymbol{w}$ in V:

$$\boldsymbol{w} = (\boldsymbol{w} \cdot \boldsymbol{u_1})\boldsymbol{u_1} + (\boldsymbol{w} \cdot \boldsymbol{u_2})\boldsymbol{u_2} + \cdots + (\boldsymbol{w} \cdot \boldsymbol{u_n})\boldsymbol{u_n}$$
$$(\boldsymbol{w})_T = (\boldsymbol{w} \cdot \boldsymbol{u_1}, \boldsymbol{w} \cdot \boldsymbol{u_2}, \ldots, \boldsymbol{w} \cdot \boldsymbol{u_n})$$

A vector $\boldsymbol{v} \in \mathbb{R}^n$ is *orthogonal* to V if $\boldsymbol{v}$ is orthogonal to all vectors in V. $\boldsymbol{v}$ is orthogonal to V if and only if $\boldsymbol{v} \cdot \boldsymbol{u_i} = 0$ for $i = 1, 2, \ldots, k$

Every vector $\boldsymbol{u} \in \mathbb{R}^n$ can be written uniquely as

$$\boldsymbol{u} = \boldsymbol{n} + \boldsymbol{p}$$

such that $\boldsymbol{n}$ is a vector orthogonal to V, $\boldsymbol{p}$ is a vector in V, also called the orthogonal projection of $\boldsymbol{u}$ onto V.

$$\boldsymbol{p} = \frac{\boldsymbol{w} \cdot \boldsymbol{u_1}}{\boldsymbol{u_1} \cdot \boldsymbol{u_1}}\boldsymbol{u_1} + \frac{\boldsymbol{w} \cdot \boldsymbol{u_2}}{\boldsymbol{u_2} \cdot \boldsymbol{u_2}}\boldsymbol{u_2} + \cdots + \frac{\boldsymbol{w} \cdot \boldsymbol{u_k}}{\boldsymbol{u_k} \cdot \boldsymbol{u_k}}\boldsymbol{u_k}$$

This generalizes $\boldsymbol{w}$ contained in V (i.e. $\boldsymbol{n}$ = 0,$\boldsymbol{p} = \boldsymbol{w}$)

Gram-Schmidt Process

Let $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$ be a basis for vector space V.

$$\boldsymbol{v_1} = \boldsymbol{u_1}$$
$$\boldsymbol{v_2} = \boldsymbol{u_2} - \frac{\boldsymbol{u_2} \cdot \boldsymbol{v_1}}{\boldsymbol{v_1} \cdot \boldsymbol{v_1}}\boldsymbol{v_1}$$
$$\vdots$$
$$\boldsymbol{v_k} = \boldsymbol{u_k} - \frac{\boldsymbol{u_k} \cdot \boldsymbol{v_1}}{\boldsymbol{v_1} \cdot \boldsymbol{v_1}}\boldsymbol{v_1} - \frac{\boldsymbol{u_k} \cdot \boldsymbol{v_2}}{\boldsymbol{v_2} \cdot \boldsymbol{v_2}}\boldsymbol{v_2} - \cdots - \frac{\boldsymbol{u_k} \cdot \boldsymbol{v_{k-1}}}{\boldsymbol{v_{k-1}} \cdot \boldsymbol{v_{k-1}}}\boldsymbol{v_{k-1}}$$

$\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_k}\}$ is an orthogonal basis for V.

$$\boldsymbol{w_1} = \frac{1}{||\boldsymbol{v_1}||}, \boldsymbol{w_2} = \frac{1}{||\boldsymbol{v_2}||}, \ldots, \boldsymbol{w_k} = \frac{1}{||\boldsymbol{v_k}||}$$

$\{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_k}\}$ is an orthogonal basis for V.

Best Approximations

Let V be a subspace in $\mathbb{R}^n$. If $\boldsymbol{u}$ is a vector in $\mathbb{R}^n$ and $\boldsymbol{p}$ is the projection of $\boldsymbol{u}$ onto V,

$$d(\boldsymbol{u}, \boldsymbol{p}) \le d(\boldsymbol{u}, \boldsymbol{v}) \text{ for all } \boldsymbol{v} \in V$$

i.e. $\boldsymbol{p}$ is the *best approximation* of $\boldsymbol{u}$ in V

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a linear system where $\boldsymbol{A}$ is a *m* x *n* matrix. $\boldsymbol{u} \in \mathbb{R}^n$ is the *least square solutions* if $||\boldsymbol{b} - \boldsymbol{Au}|| \le ||\boldsymbol{b} - \boldsymbol{Av}||$ for all $\boldsymbol{v} \in \mathbb{R}^n$

Let $\boldsymbol{p}$ be the projection of $\boldsymbol{b}$ onto the column space of $\boldsymbol{A}$

$$||\boldsymbol{b} - \boldsymbol{Ap}|| \le ||\boldsymbol{b} - \boldsymbol{Av}|| \text{ for all } \boldsymbol{v} \in \mathbb{R}^n$$

i.e. $\boldsymbol{u}$ is the least squares solution to

$$\boldsymbol{Ax} = \boldsymbol{b} \text{ if and only if } \boldsymbol{Au} = \boldsymbol{p}$$

Alternatively, $\boldsymbol{u}$ is the least squares solution to $\boldsymbol{Ax} = \boldsymbol{b}$ if and only if $\boldsymbol{u}$ is a solution to

$$\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b} \rightarrow \boldsymbol{x} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\,\boldsymbol{A}^T\boldsymbol{b}$$

Note: $\boldsymbol{A}^T\boldsymbol{A}$ is always a symmetric matrix.

Orthogonal Matrices

A square matrix $\boldsymbol{A}$ is called *orthogonal* if $\boldsymbol{A}^{-1} = \boldsymbol{A}^T$

A square matrix $\boldsymbol{A}$ will be orthogonal if and only if

$$\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I}$$

1) $\boldsymbol{A}$ is orthogonal
2) The rows/columns of $\boldsymbol{A}$ form a orthonormal basis for $\mathbb{R}^n$.

Let S and T are two orthonormal bases for a vector space, and $\boldsymbol{P}$ be the transition matrix from S to T. Then $\boldsymbol{P}$ is orthogonal and $\boldsymbol{P}^T$ will be the transition matrix from T to S.

Eigenvalues and Eigenvectors

Let $\boldsymbol{A}$ be a square matrix of order n. A nonzero column vector $\boldsymbol{u} \in \mathbb{R}^n$ is called an *eigenvector* if

$$\boldsymbol{Au} = \lambda\boldsymbol{u}$$

for some scalar λ, which is called a *eigenvalue* associated with the eigenvector $\boldsymbol{u}$.

The equation $\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$ is called the characteristic equation of $\boldsymbol{A}$ and $\det(\lambda\boldsymbol{I} - \boldsymbol{A})$ is the characteristic polynomial of $\boldsymbol{A}$.

The solution space $(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ is the *eigenspace* associated with eigenvalue λ denoted as E_λ.

If $\boldsymbol{A}$ is a triangular matrix (upper/lower), the eigenvalues of $\boldsymbol{A}$ are the diagonal entries of $\boldsymbol{A}$.

Diagonalization

A square matrix $\boldsymbol{A}$ is *diagonalizable* if there exists an invertible matrix $\boldsymbol{P}$ such that $\boldsymbol{P^{-1}AP}$ is a diagonal matrix. Matrix $\boldsymbol{P}$ is said to *diagonalize* $\boldsymbol{A}$.

$\boldsymbol{A}$ is diagonalizable if and only if $\boldsymbol{A}$ contains n linearly independent eigenvectors.

If $\boldsymbol{A}$ has *n* distinct eigenvalues, $\boldsymbol{A}$ is diagonalizable.

$$\boldsymbol{P^{-1}A^nP} = \boldsymbol{D^n} \rightarrow \boldsymbol{A^n} = \boldsymbol{PD^nP^{-1}}$$

To diagonalize a matrix, find all distinct eigenvalues. For each eigenvalue λ_i, find a basis S_{λ_i} for eigenspace E_{λ_i}. Let $S = S_{\lambda_1} \cup S_{\lambda_2} \cup \ldots S_{\lambda_k}$.

If $|S| < n$, then $\boldsymbol{A}$ is not diagonalizable.

If $|S| = n$, then say $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$, $\boldsymbol{P} = (\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k})$ is an invertible matrix that diagonalizes $\boldsymbol{A}$. S is always linearly independent.

$\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = (\lambda - \lambda_1)^{r1}(\lambda - \lambda_2)^{r2} \cdots (\lambda - \lambda_k)^{r_k}$

Where $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of $\boldsymbol{A}$.

Then $\dim(E_{\lambda_i}) \le r_i$ and $\boldsymbol{A}$ is diagonalizable if and only if $\dim(E_{\lambda_i}) = r_i$, i.e. $|S_{\lambda_i}| = r_i$ (r_i is multiplicity)

$$a_{n+1} = pa_{n-1} + qa_n \rightarrow \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ p & q \end{pmatrix}\begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix}$$

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \boldsymbol{P}\begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix}\boldsymbol{P^{-1}}\begin{pmatrix} a_0 \\ a_1 \end{pmatrix}, a_0, a_1 \textit{ base case}$$

Orthogonal Diagonalization

A square matrix $\boldsymbol{A}$ is *orthogonally diagonalizable* if there exists an invertible matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^T\boldsymbol{AP}$ is a diagonal matrix. $\boldsymbol{A}$ is *orthogonally diagonalizable* if and only if it is symmetric. $\boldsymbol{A}^T = \boldsymbol{A}$

To orthogonally diagonalize a symmetric matrix $\boldsymbol{A}$, perform Gram-Schmidt on eigenbases with same eigenvalue. Different eigenvalue bases are orthogonal.

Linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$

Linear transformation mapping from T: $\mathbb{R}^n \rightarrow \mathbb{R}^m$

$T((\boldsymbol{x_1}, \ldots, \boldsymbol{x_n})^T) = \boldsymbol{A}(\boldsymbol{x_1}, \ldots, \boldsymbol{x_n})^T$ $\boldsymbol{A}$:standard matrix

$$O(\boldsymbol{u}) = \boldsymbol{0} \qquad I(\boldsymbol{u}) = \boldsymbol{u} \qquad T(\boldsymbol{0}) = \boldsymbol{0}$$

$T(c_1\boldsymbol{u_1} + \cdots + c_k\boldsymbol{u_k}) = c_1T(\boldsymbol{u_1}) + \cdots + c_kT(\boldsymbol{u_k})$

If $\boldsymbol{v} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_n\boldsymbol{u_n}, c_1, c_2, \ldots, c_r \in \mathbb{R}$

Then $T(\boldsymbol{v}) = c_1T(\boldsymbol{u_1}) + c_2T(\boldsymbol{u_2}) + \cdots + c_nT(\boldsymbol{u_n})$

$T(\boldsymbol{e_i}) = \boldsymbol{Ae_i} =$ the *i*th column of $\boldsymbol{A}$

$$\boldsymbol{A} = (T(\boldsymbol{e_1}) \quad T(\boldsymbol{e_2}) \quad \ldots \quad T(\boldsymbol{e_n}))$$

Let S: $\mathbb{R}^m \rightarrow \mathbb{R}^k$, Composition of S with T: S($\boldsymbol{u}$)=$\boldsymbol{Bu}$

$$(S \circ T)(\boldsymbol{u}) = S(T(\boldsymbol{u})) = \boldsymbol{B}(\boldsymbol{A}(\boldsymbol{u})) = (\boldsymbol{BA})\boldsymbol{u}$$

Ranges and Kernels

Range: $R(T) = \{T(\boldsymbol{u}) | \boldsymbol{u} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$

$R(T) =$ column space of $\boldsymbol{A} = T(\boldsymbol{u_1}) + \cdots + T(\boldsymbol{u_n})$

$$\dim(R(T)) \rightarrow \operatorname{rank}(T) = \operatorname{rank}(\boldsymbol{A})$$

Kernel: $\operatorname{Ker}(T) = \{\boldsymbol{u} | T(\boldsymbol{u}) = \boldsymbol{0}\} \subseteq \mathbb{R}^n$

Ker(T) $= \{\boldsymbol{u} | \boldsymbol{Au} = \boldsymbol{0}\} =$ nullspace of $\boldsymbol{A}$

$$\text{nullity}(T) = nullity(\boldsymbol{A})$$

Let $\boldsymbol{A}$ be a square matrix of order n.

1) $\boldsymbol{A}$ is invertible
2) $\boldsymbol{Ax} = \boldsymbol{0}$ has only the trivial solution
3) RREF of $\boldsymbol{A}$ is an identity matrix $\boldsymbol{I_n}$
4) $\boldsymbol{A}$ can be expressed as product of elementary matrices ($\boldsymbol{E_1E_2} \ldots \boldsymbol{E_n} = \boldsymbol{A}$)
5) $\det(\boldsymbol{A}) \neq 0$
6) rows/columns of $\boldsymbol{A}$ form a basis for $\mathbb{R}^n$
7) $\operatorname{rank}(\boldsymbol{A}) = n$
8) 0 is not an eigenvalue of $\boldsymbol{A}$.