Lecture 21

Last time we proved that if A is a symmetric matrix then $\max_{x \neq 0} \frac{x^t A x}{x^t x}$ & $\min_{x \neq 0} \frac{x^t A x}{x^t x}$ are eigenvalues of A and any other eigenvalue of A is in between these two eigenvalues.

We also proved that a symmetric matrix A is positive definite iff all its eigenvalues are positive.

Last time we did Power method to determine dominant eigenvalue of a matrix A (i.e., an eigenvalue λ such that $|\lambda| > |\mu|$ for any other eigenvalue μ.

In this method we do assume that A has a complete set of eigenvectors i.e., a basis of $\mathbb{R}^n$ consisting of eigenvectors of A.

We consider two cases

1. A is a general matrix
2. A is a symmetric matrix.

The general case is little bit complicated However when A is symmetric the algorithim is as follows

choose x_0 with $\|x_0\|_2 = 1$

for $n = 1, 2, \ldots$

$$y_n = A x_{n-1}$$

$$\mu_n = y_n^t x_{n-1}$$

$$x_n = \frac{y_n}{\|y_n\|_2}$$

Then $\mu_n \longmapsto$ dominant eigenvalue

$x_n \longmapsto$ corresponding eigenvector.

Inverse power method

It is used to find an eigenvector of A that is close to a number q.

In this case we apply the power method to the matrix

$$(A - qI)^{-1}$$

while solving $y^{(m)} = (A - qI)^{-1} x^{(m-1)}$

we infact solve

$$(*) \quad (A - qI)\, y^{(m)} = x^{(m-1)}$$

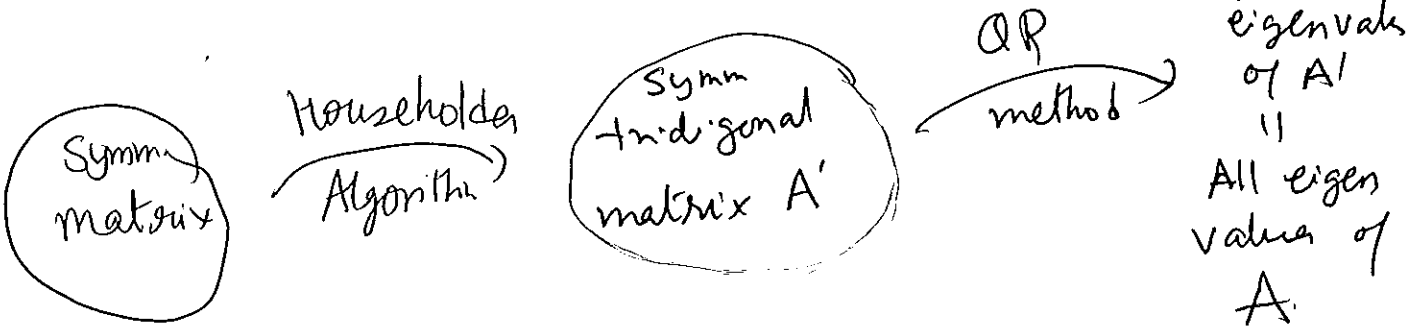
We do a PLU factorization of $A - qI$ and use it to solve (*).

Lecture 21

Householder method

In the next lecture we will use the QR method to reduce a symmetric tridigonal matrix to a similar matrix that is nearly diagonal. The diagonal entries of the reduced matrix are approximations to the eigenvalue of the given matrix

Householder method is used to reduce an arbitrary symmetric matrix to a similar symmetric tridigonal matrix.



Definition.

Let $w \in \mathbb{R}^n$, with $\|w\|_2 = 1$, i.e, $w^t w = 1$

The $n \times n$ matrix

$$P = I - 2ww^t$$

is called a House-holder transformation

House-holder transformations are used to selectively zero out blocks of entries in vectors or columns of matrices in a manner that is extremely stable with respect to round-off error.

Properties of House holder transformation are given in the following theorem

Theorem A Householder transformation $P = I - 2ww^t$ is symmetric and orthogonal (i.e., $P^{-1} = P^t = P$)

Proof Since

$$(ww^t)^t = (w^t)^t w^t = ww^t$$

it follows that

$$P^t = (I - 2ww^t)^t = I - 2ww^t = P$$

$$\begin{aligned} PP^t &= PP \\ &= (I - 2ww^t)(I - 2ww^t) \\ &= I - 4ww^t + 4(ww^t)(ww^t) \\ &= I - 4ww^t + 4w(w^tw)w^t \\ &= I - 4ww^t + 4ww^t \quad (\because w^tw = 1) \\ &= I. \end{aligned}$$

So $P^{-1} = P = P^t$.

Householder method

$$A = \begin{bmatrix} * & * & * & * & \cdots & * \\ * & * & * & * & & * \\ * & * & * & * & & \\ & & & & & \\ * & * & * & * & & * \end{bmatrix}$$

$$A = A^{(1)}$$

$$A^{(2)} = P^{(1)} A P^{(1)}$$

($A^{(2)}$ is similar to $A^{(1)}$)

$$A^{(2)} = \begin{bmatrix} * & * & 0 & 0 & \cdots & 0 \\ * & * & * & * & & * \\ 0 & * & * & * & & * \\ 0 & * & * & * & & \\ \vdots & & & & & \\ 0 & * & * & & & * \end{bmatrix}$$

$$A^{(3)} = P^{(2)} A^{(2)} P^{(2)}$$

$A^{(3)}$ is similar to $A^{(2)}$ and hence is similar to $A^{(1)}$

$$A^{(3)} = \begin{bmatrix} * & * & 0 & 0 & \cdots & 0 \\ * & * & * & 0 & & 0 \\ 0 & * & * & * & & * \\ 0 & * & * & * & & * \\ \vdots & 0 & * & * & & \\ & \vdots & & & & \\ 0 & 0 & * & & & * \end{bmatrix}$$

Keep on doing this

we get

$$A^{(n)} = \begin{bmatrix} * & * & 0 & \cdots & 0 \\ * & * & * & 0 \cdots & 0 \\ 0 & * & * & & \\ \vdots & 0 & & & \\ & \vdots & & & * \\ 0 & 0 & 0 \cdots & 0 & * \; * \end{bmatrix}$$

which is a tri-diagonal matrix.

Householder method

We determine a Householder transformation $P^{(1)}$ with the property that

$$A^{(2)} = P^{(1)} A P^{(1)} \quad \text{has}$$

$$* \qquad a_{j1}^{(2)} = 0 \quad \text{for } j = 3, 4, \ldots, n$$

and as $A^{(2)}$ is symmetric we also have

$$a_{1j}^{(2)} = 0 \quad \text{for each } j = 3, 4, \ldots, n$$

The vector $w = (w_1, w_2, \ldots, w_n)^t$ is chosen so that $w^t w = 1$, $*$ holds and in the matrix

$$A^{(2)} = P^{(1)} A P^{(1)} = (I - 2ww^t) A (I - 2ww^t)$$

we have $a_{11}^{(2)} = a_{11}$ and

$$a_{j1}^{(2)} = 0 \quad \text{for } j = 3, 4, \ldots, n$$

This condition imposes n-condition on the unknowns $w_1, w_2, \ldots, w_n$.

Setting $w_1 = 0$ ensures that $a_{11}^{(2)} = a_{11}$

This is so since $P^{(1)} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{bmatrix}$

We want $P^{(1)} = I - 2ww^t$

$$(**) \quad \begin{aligned} P^{(1)} A P^{(1)}(e_1) &= P^{(1)} A(e_1) \\ &= P^{(1)} (a_{11}, a_{21}, \ldots, a_{n1})^t \\ &= (a_{11}, \alpha, 0, 0, \ldots 0) \end{aligned}$$

where α will be chosen later.

To simplify notation

$$\hat{w} = (w_2, \ldots, w_n)^t \in \mathbb{R}^{n-1}$$

$$\hat{y} = (a_{21}, \ldots, a_{n1})^t \in \mathbb{R}^{n-1}$$

and let $\hat{P}$ be the $(n-1) \times (n-1)$ Householder transformation

$$\hat{P} = I_{n-1} - 2\hat{w}\hat{w}^t$$

$(**)$ then becomes

$$P^{(1)} \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \hat{P} & \\ 0 & & & \end{array}\right] \begin{bmatrix} a_{11} \\ \hline \\ \hat{y} \\ \\ \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} \\ \hline \hat{P}y \end{bmatrix} = \begin{bmatrix} a_{11} \\ \hline \alpha \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

So Note. $\hat{P}\hat{y} = (I_{n-1} - 2\hat{\omega}\hat{\omega}^t)\hat{y}$

$= \hat{y} - 2\hat{\omega}\hat{\omega}^t\hat{y}$

Note $\hat{\omega}\hat{\omega}^t\hat{y} = \hat{\omega}^t\hat{y}\,\hat{\omega}$ ($\because \hat{\omega}^t y$ is a constant)

So $\hat{P}\hat{y} = \hat{y} - 2(\hat{\omega}^t\hat{y})\hat{\omega}$

$= (\alpha, 0, 0, \ldots, 0)^t$

$r = \hat{\omega}^t\hat{y}$. Then

$(\alpha, 0, 0, \ldots 0) = (a_{21} - 2r\omega_1, a_{31} - 2r\omega_2, \ldots, a_{n1} - 2r\omega_n)$

Equating components gives

$\alpha = a_{21} - 2r w_1$ and $0 = a_{j1} - 2r w_j$ for $j = 3, 4, \ldots, n$.

Thus $2r w_2 = a_{21} - \alpha$

and $2r w_j = a_{j1}$ for $j = 3, \ldots, n$

Squaring both sides of each equations and adding gives

$$4r^2 \sum_{j=2}^{n} w_j^2 = (a_{21} - \alpha)^2 + \sum_{j=3}^{n} a_{j1}^2$$

Since $w^t w = 1$ and $w_1 = 0$ we have that

$$\sum_{j=2}^{n} w_j^2 = 1$$

$$4r^2 = \sum_{j=2}^{n} a_{j1}^2 - 2\alpha a_{21} + \alpha^2$$

$$\alpha^2 = (\alpha, 0, \ldots, 0)(\alpha, 0, \ldots 0)^t$$

$$= (\hat{P}\hat{y})^t \hat{P}\hat{y} \qquad (\because \hat{P}\hat{y} = (\alpha, 0, \ldots, 0)^t)$$

$$= \hat{y}^t \hat{P}^t \hat{P} \hat{y}$$

$$= \hat{y}^t \hat{y} \qquad \because \hat{P} \text{ is orthogonal.}$$

Thus $\alpha^2 = \sum_{j=2}^{n} a_{j1}^2$

$$4r^2 = \sum_{j=2}^{n} a_{j1}^2 - 2\alpha a_{21} + \alpha^2$$

Substituting α^2 we get

$$2r^2 = \sum_{j=2}^{n} a_{j1}^2 - \alpha a_{21}$$

To ensure $2r^2 = 0$ if and only if

$a_{21} = a_{31} = \ldots = a_{n1} = 0$ we choose

$$\alpha = -\operatorname{sgn}(a_{21}) \left(\sum_{j=2}^{n} a_{j1}^2 \right)^{1/2}$$

which implies that-

$$2r^2 = \sum_{j=2}^{n} a_{j1}^2 + |a_{21}| \left(\sum_{j=2}^{n} a_{j1}^2 \right)^{1/2}$$

With this choice of α and $2r^2$ we get

$$w_2 = \frac{a_{21} - \alpha}{2r} \quad \text{and} \quad w_j = \frac{a_{j1}}{2r} \quad j = 3, \ldots n$$

To summarize the choice of $p^{(1)}$ we have

$$\alpha = -\operatorname{sgn}(a_{21}) \left(\sum_{j=2}^{n} a_{j1}^2 \right)^{1/2}$$

$$r = \left(\frac{1}{2}\alpha^2 - \frac{1}{2} a_{21} \alpha \right)^{1/2}$$

$$w_1 = 0$$

$$w_2 = \frac{a_{21} - \alpha}{2r}$$

$$w_j = \frac{a_{j1}}{2r} \quad \text{for } j = 3, \ldots, n$$

With this choice

$$A^{(2)} = P^{(1)} A P^{(1)} = \begin{bmatrix} a_{11}^{(2)} & a_{12}^{(2)} & 0 & \cdots & 0 \\ a_{21}^{(2)} & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & a_{32}^{(2)} & \cdots & \cdots & a_{3n}^{(2)} \\ \vdots & \vdots & & & \\ 0 & a_{n2}^{(2)} & \cdots & \cdots & a_{nn}^{(2)} \end{bmatrix}$$

Having found $P^{(1)}$ and computed $A^{(2)}$ the procen is repeated for $k = 2, 3, \ldots, n-2$ as follows

$$\alpha = -\operatorname{sgn}(a_{k+1,k}^{(k)}) \left(\sum_{j=k+1}^{n} (a_{jk}^{(k)})^2 \right)^{1/2}$$

$$r = \left(\frac{1}{2}\alpha^2 - \frac{1}{2}\alpha\, a_{k+1,k}^{(k)} \right)^{1/2}$$

$$w_1^{(k)} = w_2^{(k)} = \cdots = w_k^{(k)} = 0$$

$$w_{k+1}^{(k)} = \frac{a_{k+1,k}^{(k)} - \alpha}{2r}$$

$$w_j^{(k)} = \frac{a_{jk}^{(k)}}{2r} \quad \text{for } j = k+2, \ldots, n$$

$$P^{(k)} = I - 2\omega^{(k)}(\omega^{(k)})^t$$

and $A^{(k+1)} = P^{(k)} A^{(k)} P^{(k)}$

where

$$A^{(k+1)} = \begin{bmatrix} a_{11}^{(k+1)} & a_{12}^{(k+1)} & 0 & \cdots & 0 \\ a_{21}^{(k+1)} & \ddots & \ddots & & \vdots \\ 0 & \ddots & a_{k+1,k}^{(k+1)} & a_{k+1,k+1}^{(k+1)} & a_{k+1,k+2}^{(k+1)} \\ \vdots & & 0 & \ddots & \ddots \\ 0 & & 0 & a_{n,k+1}^{(k+1)} & \cdots \; a_{nn}^{(k+1)} \end{bmatrix}$$

keep on doing this

Then $A^{(n-1)}$ will be symmetric and ~~diag~~ tridigonal

$$A^{(n-1)} = P^{(n-2)} P^{(n-3)} \cdots P^{(1)} A P^{(1)} \cdots P^{(n-3)} P^{(n-2)}$$

So $A^{(n-1)}$ is similar to A.

Example

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

A is symmetric

We use Householder transformation to reduce A to a symmetric tridiagonal matrix.

determining $P^{(1)}$

$$\alpha = -(1)\left(\sum_{j=2}^{4} a_{j1}^2\right)^{1/2} = -3$$

$$r = \left(\frac{1}{2}\,9 + \frac{1}{2}\cdot 3\right)^{1/2} = \sqrt{6}$$

$$w_1 = 0$$

$$w_2 = \frac{1+3}{2\sqrt{6}} = \frac{2}{\sqrt{6}} = \frac{\sqrt{6}}{3}$$

$$w_3 = \frac{-2}{2\sqrt{6}} = \frac{-1}{\sqrt{6}} = -\frac{\sqrt{6}}{6}$$

$$\omega_4 = \frac{2}{2\sqrt{6}} = \frac{1}{\sqrt{6}} = \frac{\sqrt{6}}{6}$$

$$\omega = \left(0, \frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{6}\right)^t$$

$$= \frac{\sqrt{6}}{6}\,(0, 2, -1, 1)^t$$

$$P = I - 2\omega\omega^t$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - 2\left(\frac{\sqrt{6}}{6}\right)^2 \begin{bmatrix} 0 \\ 2 \\ -1 \\ 1 \end{bmatrix} [0, 2, -1, 1]$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1/3 & 2/3 & -2/3 \\ 0 & 2/3 & 2/3 & 1/3 \\ 0 & -2/3 & 1/3 & 2/3 \end{bmatrix}$$

$$A^{(2)} = \begin{bmatrix} 4 & -3 & 0 & 0 \\ -3 & 10/3 & 1 & 4/3 \\ 0 & 1 & 5/3 & -4/3 \\ 0 & 4/3 & -4/3 & 1 \end{bmatrix}$$

determining $P^{(2)}$

$$\alpha = -\operatorname{sgn}\left(a_{3,2}^{(2)}\right)\left(\sum_{j=3}^{4}\left(a_{j,2}^{(2)}\right)^2\right)^{1/2}$$

$$= -\left(1+\frac{16}{9}\right)^{1/2}$$

$$= -5/3$$

$$r = \left(\frac{1}{2}\left(-5/3\right)^2 - \frac{1}{2}\left(-\frac{5}{3}\right)^1\right)^{1/2}$$

$$= \left(\frac{1}{2}\cdot\frac{25}{9}+\frac{5}{6}\right) = \left(\frac{40}{2\cdot 9}\right)^{1/2}$$

$$= \left(\frac{20}{9}\right)^{1/2} = \frac{2\sqrt{5}}{3}.$$

$$w = \begin{pmatrix} 0 \\ 0 \\ 2\sqrt{5} \\ \sqrt{5}/5 \end{pmatrix}^t$$

$$P^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -3/5 & -4/5 \\ 0 & 0 & -4/5 & 3/5 \end{bmatrix}$$

And the symmetric tridigonal matrix is

$$A^{(3)} = \begin{bmatrix} 4 & -3 & 0 & 0 \\ -3 & 10/3 & -5/3 & 0 \\ 0 & -5/3 & -\frac{33}{25} & 68/15 \\ 0 & 0 & \frac{68}{75} & \frac{149}{75} \end{bmatrix}$$

Remark

$$P = I - 2ww^t$$

On the k^{th} step

$$P^{(k)} = I - 2w^{(k)}(w^{(k)})^t$$

$$w^{(k)} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ * \\ * \\ x \\ \vdots \\ x \end{pmatrix} \leftarrow k^{th} \text{ row}$$

$$P^{(k)} = I - 2w^{(k)}(w^{(k)})^t$$

$$= \left[\begin{array}{c|c} I_k & 0 \\ \hline 0 & \begin{matrix} * & x \\ x & 0 \\ \vdots & \end{matrix} \end{array}\right]$$

$$A^{(k+1)} = P^{(k)} A^{(k)} P^{(k)}$$

$$= \begin{bmatrix} \text{same as } A^{(k)} & 0 \\ 0 & \begin{matrix} * & * & * \\ * & & * \\ * & & \end{matrix} \end{bmatrix}$$

determine w_k such that we get appropriate zeros

Theoretical background regarding QR-method

Inner product spaces

Let V be a vectorspace over $\mathbb{R}$. We say V is an inner product space over $\mathbb{R}$ if $\exists$ a function $\langle , \rangle : V \times V \to \mathbb{R}$ such that

1) $\langle a, a \rangle \geq 0 \quad \forall a \in V$
with $\langle a, a \rangle = 0 \iff a = 0$

2) $\langle v, u \rangle = \langle u, v \rangle$

3) $\langle \alpha_1 v_1 + \alpha_2 v_2, u \rangle = \alpha_1 \langle v_1, u \rangle + \alpha_2 \langle v_2, u \rangle$
$\forall \alpha_1, \alpha_2 \in \mathbb{R}, \; v_1, v_2, u \in V$

Examples

1) $V = \mathbb{R}^n \qquad \langle a, b \rangle = a^t . b$

2) $V = C[a,b] = \{ f : [a,b] \to \mathbb{R} \mid f \text{ is cts} \}$

$$\langle f, g \rangle = \int_a^b f(x)\, g(x)\, dx.$$

Then $\langle , \rangle$ is an inner product in $C[a,b]$

~~Orthogonal~~ · Orthogonal set

Let V be an inner product space.

A set $\{v_1, v_2, \ldots, \}$ is said to be orthogonal if $v_i \neq 0$ for all i and

$$\langle v_i, v_j \rangle = 0 \quad \text{for } i \neq j$$

Example

$\{(1,1), (1,-1)\}$ is an orthogonal set in $\mathbb{R}^2$.

Norm :- Let V be an inner product space

$$\|v\| \overset{\text{defn}}{:=} \sqrt{\langle v, v \rangle}$$

$\| \ \|$ is called norm of v

One can prove

1) $\|v\| \geq 0$ and $\|v\| = 0$ iff $v = 0$

2) $\|\alpha v\| = |\alpha| \|v\|$ for $\alpha \in \mathbb{R}$.

3) $\|v+w\| \leq \|v\| + \|w\|$ (triangle inequality)

Triangle inequality is usually proved by proving

Lemma (Cauchy - Schwarz inequality)

$$|\langle v, w\rangle| \leq \|v\| \, \|w\|$$

Proof Exercise

(we proved it for $\mathbb{R}^n$. The same proof generalizes)

Orthonormal set

A subset $\{v_1, v_2, \ldots\}$ of V is said to be orthonormal if

1) it is orthogonal, i.e., $v_i \neq 0 \ \forall i$ and $\langle v_i, v_j \rangle = 0$ for $i \neq j$

2) $\|v_i\| = 1$ for all $i \geq 1$.

Remark :- A orthogonal set is L.I.

Proof Suppose $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0$

$$\langle \alpha_1 v_1 + \cdots + \alpha_k v_k, v_j \rangle = \langle 0, v_i \rangle = 0$$

$$\text{L.H.S} = \langle \alpha_1 v_1 + \cdots + \alpha_k v_k, v_i \rangle$$

$$= \langle \alpha_1 v_1, v_i \rangle + \cdots + \langle \alpha_k v_k, v_i \rangle$$

$$= \alpha_1 \langle v_1, v_i \rangle + \alpha_2 \langle v_2, v_i \rangle + \cdots + \alpha_k \langle v_k, v_i \rangle$$

$$= \alpha_i \langle v_i, v_i \rangle$$

$$\alpha_j \langle v_i, v_i \rangle = 0$$

$v_i \neq 0 \Rightarrow \langle v_i, v_i \rangle \neq 0$. So $\alpha_i = 0$

Corollary :- Let V be a finite dimensional inner product space. Let A be an orthogonal set

Then $\#A \leq \dim V$.

Question:- Can we find a basis of V consisting of orthogonal elements ?

Ans Yes.

Gram-Schmidt orthonormalization

Let $\{v_1, v_2, \ldots\}$ be a linearly independent subset of an inner product space V.

Define $y_1 = v_1$ and $u_1 = \frac{y_1}{\|y_1\|}$

for $n = 2, 3, 4, \ldots$

$$y_n = v_n - \langle v_n, u_1 \rangle u_1 - \langle v_n, u_2 \rangle u_2 \cdots - \langle v_n, u_{n-1} \rangle u_{n-1}$$

i.e $$y_n = v_n - \sum_{j=1}^{n-1} \langle v_n, u_i \rangle u_i$$

$$u_n = \frac{y_n}{\|y_n\|}$$

Theorem $\{u_1, u_2, \ldots\}$ is an orthonormal set in V and for $n = 1, 2, \ldots$

$$\text{Span}\{u_1, \ldots u_n\} = \text{Span}\{v_1, v_2, \ldots, v_n\}$$

Corollary Let V be a f.d inner product space. Then there exists an orthonormal basis for V

Proof Let $\{v_1, \ldots v_n\}$ be a basis of V.

Let $\{u_1, \ldots u_n\}$ be the set obtained by Gram-Schmidt process

Then $\{u_1, \ldots u_n\}$ is orthonormal and hence linearly independent

Also $\text{span}\{u_1, \ldots u_n\} = \text{span}\{v_1, \ldots v_n\} = V$.

So $\{u_1, \ldots u_n\}$ is a basis of V.

Proof of Gram-Schmidt orthonormalization

Proof by induction on n

$n = 1$ $\{v_1\}$ is L.I

So $v_1 \neq 0$. Thus $y_1 = v_1 \neq 0$

$$u_1 = \frac{y_1}{\|y_1\|}.$$

$$\|u_1\| = 1.$$

$$\text{span}\{u_1\} = \text{span}\{y_1\} = \text{span}\{v_1\}.$$

for $n \geq 1$ assume that we have defined y_n and u_n as stated above and proved that $\{u_1, u_2, \ldots, u_n\}$ is an orthonormal set satisfying $\text{span}\{u_1, \ldots u_n\} = \text{span}\{v_1, \ldots v_n\}$.

Define

$$y_{n+1} = v_{n+1} - \langle v_{n+1}, u_1\rangle u_1 - \cdots - \langle v_{n+1}, u_n\rangle u_n$$

As $\{v_1, v_2, \ldots v_{n+1}\}$ is a L.I set

$v_{n+1} \notin \text{span}\{v_1, \ldots v_n\} = \text{span}\{u_1, \ldots u_n\}$

So $y_{n+1} \neq 0$

Let $u_{n+1} = \dfrac{y_{n+1}}{\|y_{n+1}\|}$

$$\|u_{n+1}\| = 1.$$

Also for $j \leq n$ we have

$$\langle y_{n+1}, u_j \rangle = \langle v_{n+1}, u_j \rangle - \sum_{k=1}^{n} \langle v_{n+1}, u_k \rangle \langle u_k, u_j \rangle$$

$$= \langle v_{n+1}, u_j \rangle - \langle v_{n+1}, u_j \rangle \langle u_j, u_j \rangle$$

$\because \langle u_k, u_j \rangle = 0$ for $k \neq j$

Also $\langle u_j, u_j \rangle = 1$

$$= \langle v_{n+1}, u_j \rangle - \langle v_{n+1}, u_j \rangle$$

$$= 0$$

$$\langle u_{n+1}, u_j \rangle = \frac{1}{\|y_{n+1}\|} \langle y_{n+1}, u_j \rangle = 0 \quad \text{for } j = 1, \ldots, n.$$

Thus $\{u_1, u_2, \ldots, u_{n+1}\}$ is an orthonormal set.

Also $\text{span}\{u_1, \ldots u_{n+1}\} = \text{span}\{v_1, \ldots v_n, u_{n+1}\}$

$= \text{span}\{v_1, \ldots v_n, y_{n+1}\}$

$= \text{span}\{v_1, \ldots v_n, v_{n+1}\}$.

Thus by mathematical induction proof is complete

Examples

1. $V = \mathbb{R}^3$

$$\mathcal{B} = \left\{ v_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \; v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \; v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

is a basis of $\mathbb{R}^3$.

By Gram-Schmidt orthonormalization process

$$y_1 = v_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \|y_1\| = 1$$

$$u_1 = y_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$y_2 = v_2 - \langle v_2, u_1 \rangle u_1$$

$$= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$\|y_2\| = 1$

$$u_2 = y_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$y_3 = v_3 - \langle v_3, u_2 \rangle u_1 - \langle v_3, u_2 \rangle u_2$$

$$= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\|y_3\| = 1$$

$$u_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Example 2 $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$ is a basis of $\mathbb{R}^2$.

$$y_1 = v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\|y_1\| = \sqrt{1^2 + 1^2} = \sqrt{2}$$

$$u_1 = \frac{1}{\sqrt{2}} y_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

$$y_2 = v_2 - \langle v_2, u_1 \rangle u_1$$

$$= \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \left(\frac{1}{\sqrt{2}} + \frac{2}{\sqrt{2}} \right) \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \frac{3}{\sqrt{2}} \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 3/2 \\ 3/2 \end{pmatrix}$$

$$= \begin{pmatrix} -1/2 \\ 1/2 \end{pmatrix}$$

$$\|y_2\| = \sqrt{1/4 + 1/4} = \frac{1}{\sqrt{2}}$$

$$u_2 = \frac{y_2}{\|y_2\|} = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

$$Q = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

Note $\det A > 0$

$\det Q = 1$

One can show

$\text{sign}(\det[v_1, \ldots v_n]) = \text{sgn}(\det[u_1, \ldots u_n])$.

Example $C[-1, 1]$ $\langle f, g \rangle = \int_{-1}^{1} fg\, dx$.

$V = \{1, x, x^2\}$
$\quad\;\; \| \quad\; \| \quad\; \|$
$\quad\;\; v_1 \quad v_2 \quad v_3$

Do Gram-Schmidt orthonormalization on V.

$y_1 = v_1$ $\qquad \|y_1\| = \sqrt{\int_{-1}^{1} 1^2} = \sqrt{2}$

$u_1 = \frac{1}{\sqrt{2}}$.

$y_2 = v_2 - \langle v_2, u_1 \rangle u_1$

$= x - \left(\int_{-1}^{1} x \cdot \frac{1}{\sqrt{2}}\right) \frac{1}{\sqrt{2}}$

$= x$

$\|y_2\| = \sqrt{\int_{-1}^{1} x^2\, dx}$

$= \sqrt{x^3/3 \,\Big|_{-1}^{1}}$

$= \sqrt{2/3}$.

$u_2 = \sqrt{\frac{3}{2}}\, x$.

$y_3 = v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2$

$= x^2 - \left(\int_{-1}^{1} x^2 \frac{1}{\sqrt{2}}\right)\left(\frac{1}{\sqrt{2}}\right) - \left(\int_{-1}^{1} x^2 \sqrt{\frac{3}{2}}\, x\right)\sqrt{\frac{3}{2}}\, x$

$$y_3 = x^2 - \frac{1}{2}\int_{-1}^{1} x^2\,dx$$

$$= x^2 - \frac{1}{2}\cdot\frac{x^3}{3}\Big|_{-1}^{1}$$

$$= x^2 - \frac{1}{2}\cdot\frac{2}{3}$$

$$y_3 = x^2 - \frac{1}{3}$$

$$\|y_3\| = \sqrt{\int_{-1}^{1}\left(x^2 - \tfrac{1}{3}\right)^2 dx}$$

$$= \quad -$$

Compute