Orthonormal basis

Definition 1

A **unit vector** is a vector of length 1.

If $\vec{u}$ is a unit vector, then the orthogonal projection on its line can be written as

$$\vec{u}(\vec{u}^T\vec{u})^{-1}\vec{u}^T = \vec{u}\vec{u}^T.$$

Orthonormal basis

Definition 2

An **orthonormal basis** of a space is a basis $\vec{v}_1, ..., \vec{v}_n$ such that

$$\vec{v}_i^T \vec{v}_j = \begin{cases} 0 & i \neq j \text{ (\textbf{ortho})} \\ 1 & i = j \text{ (\textbf{normal})} \end{cases}$$

Letting $\Omega = \begin{pmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{pmatrix}$ be the matrix of collections of these vectors. Then $\vec{v}_1, ..., \vec{v}_n$ is an orthonormal basis if and only if $\Omega^T \Omega = I$

Orthonormal basis

!!

Note that columns of Ω being a basis means it is a squre matrix, and thus $\Omega^T\Omega = I$ is equivalent to $\Omega\Omega^T = I$. Therefore, **columns being orthonormal basis is equivalent as rows being orthogormal.**

Suppose $\Omega = \begin{pmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{pmatrix}$, then

$$\Omega^T\Omega = \begin{pmatrix} \vec{v}_1^T\vec{v}_1 & \vec{v}_1^T\vec{v}_2 & \cdots & \vec{v}_1^T\vec{v}_n \\ \vec{v}_2^T\vec{v}_1 & \vec{v}_2^T\vec{v}_2 & \cdots & \vec{v}_2^T\vec{v}_n \\ \dots & \dots & \ddots & \dots \\ \vec{v}_n^T\vec{v}_1 & \vec{v}_n^T\vec{v}_2 & \cdots & \vec{v}_n^T\vec{v}_n \end{pmatrix}$$

Orthogonal (Symmetric) Cross-Filling

Let A be a symmetric matrix, if we choose the center of the cross-filling to be on diagonal, then the cross-filling summands $A = A_1 + A_2$ are all symmetric matrix!

$$X = (Ae_i)(e_i^T Ae_i)^{-1}(e_i^T A) = (Ae_i)(e_i^T Ae_i)^{-1}(Ae_i)^T$$

Symmetric!

Orthogonal (Symmetric) Cross-Filling

Excercise. The following matrix

$$\begin{pmatrix} 0.5 & 0 & 0 & -0.5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -0.5 & 0 & 0 & 0.5 \end{pmatrix}$$

is a projection to the space

$$W : \left\{ \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} : x + y - z + w = 0 \right\}$$

Find an orthonormal basis $\vec{w}_1, \vec{w}_2, \vec{w}_3$ of W.

Orthogonal (Symmetric) Cross-Filling

Solution. We wanna keep the matrix symmetric while cross-filling, the method is to choose crosses in a symmetric way - with centers on diagonal

$$\begin{pmatrix} 0.5 & 0 & 0 & -0.5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -0.5 & 0 & 0 & 0.5 \end{pmatrix}$$

$$= \begin{pmatrix} 0.5 & 0 & 0 & -0.5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -0.5 & 0 & 0 & 0.5 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ \vec{0} & 1 & \vec{0} & \vec{0} \\ 0 & \vec{0} & 0 & 0 \\ 0 & \vec{0} & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & \vec{0} & 0 \\ 0 & 0 & \vec{0} & 0 \\ \vec{0} & \vec{0} & 1 & \vec{0} \\ 0 & 0 & \vec{0} & 0 \end{pmatrix}$$

Orthogonal (Symmetric) Cross-Filling

We may decompose this into

$$\begin{pmatrix} 0.5 \\ 0 \\ 0 \\ -0.5 \end{pmatrix} \frac{1}{0.5} \begin{pmatrix} 0.5 & 0 & 0 & -0.5 \end{pmatrix} + \begin{pmatrix} \vec{0} \\ 1 \\ \vec{0} \\ \vec{0} \end{pmatrix} \frac{1}{1} \begin{pmatrix} \vec{0} & 1 & \vec{0} & \vec{0} \end{pmatrix} + \begin{pmatrix} \vec{0} \\ \vec{0} \\ 1 \\ \vec{0} \end{pmatrix} \frac{1}{1} \begin{pmatrix} \vec{0} & \vec{0} & 1 & \vec{0} \end{pmatrix}$$

Now we decompose it into

$$\begin{pmatrix} 0.5 \\ 0 \\ 0 \\ -0.5 \end{pmatrix} \frac{1}{\sqrt{0.5}} \frac{1}{\sqrt{0.5}} \begin{pmatrix} 0.5 & 0 & 0 & -0.5 \end{pmatrix} + \begin{pmatrix} \vec{0} \\ 1 \\ \vec{0} \\ \vec{0} \end{pmatrix} \frac{1}{\sqrt{1}} \frac{1}{\sqrt{1}} \begin{pmatrix} \vec{0} & 1 & \vec{0} & \vec{0} \end{pmatrix}$$

$$+ \begin{pmatrix} \vec{0} \\ \vec{0} \\ 1 \\ \vec{0} \end{pmatrix} \frac{1}{\sqrt{1}} \frac{1}{\sqrt{1}} \begin{pmatrix} \vec{0} & \vec{0} & 1 & \vec{0} \end{pmatrix}$$

Orthogonal (Symmetric) Cross-Filling

Therefore, with

$$\vec{v}_1 = \frac{1}{\sqrt{0.5}} \begin{pmatrix} 0.5 & 0 & 0 & -0.5 \end{pmatrix}, \qquad \vec{v}_2 = \begin{pmatrix} \vec{0} \\ 1 \\ \vec{0} \\ \vec{0} \end{pmatrix} \; \vec{v}_3 = \begin{pmatrix} \vec{0} \\ \vec{0} \\ 1 \\ \vec{0} \end{pmatrix}$$

We have already decomposed

$$P = \vec{v}_1 \vec{v}_1^T + \vec{v}_2 \vec{v}_2^T + \vec{v}_3 \vec{v}_3^T$$

Note that P is projection, so the number 3 has been reflected in $\mathrm{tr}(P)$.

Orthogonal (Symmetric) Cross-Filling

Since P is a projection, so is $I - P$. I is symmetric, so is $I - P$.

$$I - P = \begin{pmatrix} 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0.5 \end{pmatrix}$$

We can write it as

$$\begin{pmatrix} 0.5 \\ \vec{0} \\ \vec{0} \\ \vec{0}.5 \end{pmatrix} \frac{1}{\sqrt{0.5}} \frac{1}{\sqrt{0.5}} \begin{pmatrix} 0.5 & \vec{0} & \vec{0} & \vec{0}.5 \end{pmatrix} = \vec{v}_4 \vec{v}_4^T$$

Orthogonal (Symmetric) Cross-Filling

Therefore, we have collected vectors

$$P = \vec{v}_1 \vec{v}_1^T + \vec{v}_2 \vec{v}_2^T + \vec{v}_3 \vec{v}_3^T \qquad I - P = \vec{v}_4 \vec{v}_4^T$$

Therefore

$$I_4 = \underbrace{\begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 \end{pmatrix}}_{\Omega} \underbrace{\begin{pmatrix} \vec{v}_1^T \\ \vec{v}_2^T \\ \vec{v}_3^T \\ \vec{v}_4^T \end{pmatrix}}_{\Omega^T}$$

This means columns of Ω is an orthonormal basis.

Orthogonal (Symmetric) Cross-Filling

Excercise. Some students make this argument for a 4×4 symmetric matrix A

$$A = \vec{v}_1 \vec{v}_1^T + \vec{v}_2 \vec{v}_2^T + \vec{v}_3 \vec{v}_3^T$$

$$I_4 - A = \vec{v}_4 \vec{v}_4^T + \vec{v}_5 \vec{v}_5^T$$

So he obtained that

$$I_4 = \underbrace{\begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 & \vec{v}_5 \end{pmatrix}}_{\Omega} \underbrace{\begin{pmatrix} \vec{v}_1^T \\ \vec{v}_2^T \\ \vec{v}_3^T \\ \vec{v}_4^T \\ \vec{v}_5^T \end{pmatrix}}_{\Omega^T}$$

Is columsn of Ω an orthonormal basis? why?

Orthogonal (Symmetric) Cross-Filling

Solution. In order for columns to be a basis, one has to be a square matrix. So in $\mathbb{R}^4$ the columns of Ω can only be basis when it consists of 4 columns.

Orthogonal (Symmetric) Cross-Filling

Theorem 1

Suppose $P = P^T = P^2$ orthogonal projection matrix. Diagonal centered Cross-filling splits P in to mutually orthogonal projections

$$P = \underbrace{v_1 v_1^T}_{P_1} + \cdots + \underbrace{v_r v_r^T}_{P_r}$$

and $v_1, \cdots, v_r$ is an orthonormal basis of $\mathrm{Col}(P)$.

Proof: Suppose P is of size $n \times n$, $\mathrm{trace}(P) = \mathrm{rank}(P) = r$. So $\mathrm{trace}(I_n - P) = \mathrm{rank}(I_n - P) = n - r$. Since P is symmetric, so $I_n - P$ is symmetric, using diagonal cross-filling we may decompose

$$I_n - P = \underbrace{v_{r+1} v_{r+1}^T}_{P_{r+1}} + \cdots + \underbrace{v_n v_n^T}_{P_n}$$

Since $\vec{v}_1 \vec{v}_1^T + ... + \vec{v}_n \vec{v}_n^T = I_n$, then $\vec{v}_1, ..., \vec{v}_n$ is an orthonormal basis!

Gram–Schmidt orthogonalization

Excercise. Let $W \subset \mathbb{R}^3$ be subspace spaned by following vectors

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

Determine an orthogonal basis for W.

Solution.. The standard way is to write

$$A = \begin{pmatrix}1 & 1\\0 & 1\\1 & 1\end{pmatrix}$$

and therefore $W = \mathrm{Col}(A)$ We may use the formula of orthogonal projection to write

$$P = A(A^T A)^{-1} A^T$$

then P is an orthogonal projection and we may use diagonal corss-filling to find an orthogonal basis. However, the **Gram–Schmidt** method is a simpler algorithm of doing so.

Description of the algorithm

Since A is linearly independent, $A^T A$ is positive definite

$$0 = x^T A^T A x = (Ax)^T (Ax) \implies Ax = 0 \implies x = 0.$$

During the homework 1, we have showed that positive definite matrix always LU-decomposible. Therefore

$$A^T A = LDL^T$$

for some lower triangular matrix L and all digaonals of D is **positive**. .

Therefore

$$(A^T A)^{-1} = L^{T-1} D^{-1} L^{-1}$$

This decomposes

$$P = \underbrace{AL^{T-1}\sqrt{D^{-1}}}_{\Omega}\underbrace{\sqrt{D^{-1}}L^{-1}A^T}_{\Omega^T}$$

Description of the algorithm

$$P = \underbrace{AL^{T-1}\sqrt{D^{-1}}}_{\Omega}\underbrace{\sqrt{D^{-1}}L^{-1}A^{T}}_{\Omega^{T}}$$

Because both L and D invertible, A having left inverse implies that Ω having left inverse. So Ω^T have right inverse. Since $P = \Omega\Omega^T$ is a projection matrix, we have

$$\Omega^T\Omega = I_m$$

this implies columns of Ω is orthonormal basis of the column space of A.

Description of the algorithm

: Wait a moment, why $P = \Omega\Omega^T$ implies $\Omega^T\Omega = I_m$????

: Is that because when $P = P^2$, then $P = AB$ implies $BA = I$?

: You have to assume A has left inverse and B has right inverse for the statement

 I am sure that we have learned it befroe. But it was not on the midterm so I did not review. How to prove it?

 Do you know the proof?

Just in case if you forgot

 Because $P^2 = P$, we have $ABAB = AB = AIB$

Sine A has left inverse, so $BAB = IB$

Since B has right inverse, so $BA = I$.

Gram-Schmidt orgthogonalization

We describe the step of the algorithm for finding orthogonal basis of Column space of A.

- Using LU-decomposition, write $A^TA = LDL^T$
- Columns of $AL^{T-1}\sqrt{D^{-1}}$ is an orthonormal basis of $\mathrm{Col}(A)$.

Gram-Schmidt orgthogonalization

Let′s finish the problem!

Excercise. Let $W \subset \mathbb{R}^3$ be subspace spaned by following vectors

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

Determine an orthogonal basis for W.

Solution.. The standard way is to write

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$$

Calculate

$$A^T A = \begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix}$$

Gram-Schmidt orgthogonalization

Using cross-filling to find LU decomposition

$$\begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & frm{-}e \\ frm{-}e & 2 \end{pmatrix} + \begin{pmatrix} 0 & \vec{0} \\ \vec{0} & 1 \end{pmatrix}$$

So we got

$$A^T A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

So

$$(A^T A)^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

An orthonormal basis is given by

$$\underbrace{\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}}_{L^{T-1}} \underbrace{\begin{pmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{pmatrix}^{-1}}_{\sqrt{D^{-1}}} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \\ \frac{1}{\sqrt{2}} & 0 \end{pmatrix}$$

Gram-Schmidt orgthogonalization

Method to find **orthonormal basis** of a subspace

Summary

- For a subspace given by image of projection P: Do diagonal cross-filling of P
- For a subspace with given basis $W = \mathrm{Col}(A)$, A having left inverse : Using Gram-Schmidt method to do LU-decomposition of $A^T A = LDL^T$ and use the formula $AL^{T-1}\sqrt{D^{-1}}$.

QR decomposition

When W is the whole space, then columns of A being basis means that it is an invertible square matrix, the formula

$$AL^{T-1}\sqrt{D^{-1}}$$

gives a square matrix with columns orthornormal basis for the whole space. Therefore, we write

$$Q = AL^{T-1}\sqrt{D^{-1}}, \qquad R = \sqrt{D}L^{T}$$

and we call

$$A = QR$$

the QR-decomposition of matrix A.

QR decomposition

Excercise. Find QR decomposition of the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$$

We calculate

$$A^T A = \begin{pmatrix} 2 & 5 \\ 5 & 13 \end{pmatrix}$$

Use cross-filling we decompose

$$A^T A = \begin{pmatrix} 2 & 5 \\ 5 & 12.5 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 0.5 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 2.5 & 1 \end{pmatrix}}_{L} \underbrace{\begin{pmatrix} 2 & 0 \\ 0 & 0.5 \end{pmatrix}}_{D} \underbrace{\begin{pmatrix} 1 & 2.5 \\ 0 & 1 \end{pmatrix}}_{L^T}$$

So

$$R = \underbrace{\begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{0.5} \end{pmatrix}}_{\sqrt{D}} \underbrace{\begin{pmatrix} 1 & 2.5 \\ 0 & 1 \end{pmatrix}}_{L^T}, \qquad Q = AR^{-1} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

Gram-Schmidt for abstract functions

Note that writting $A = \begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{pmatrix}$, we have in fact

$$A^T A = \begin{pmatrix} \vec{v}_1^T \vec{v}_1 & \vec{v}_1^T \vec{v}_2 & \cdots & \vec{v}_1^T \vec{v}_n \\ \vec{v}_2^T \vec{v}_1 & \vec{v}_2^T \vec{v}_2 & \cdots & \vec{v}_2^T \vec{v}_n \\ \dots & \dots & \ddots & \dots \\ \vec{v}_n^T \vec{v}_1 & \vec{v}_n^T \vec{v}_2 & \cdots & \vec{v}_n^T \vec{v}_n \end{pmatrix}$$

Symbolically, we write the inner product as a pair

$$\vec{v}_1^T \vec{v}_2 =: \langle \vec{v}_1, \vec{v}_2 \rangle$$

it has property that $\langle \vec{v}_1, \vec{v}_2 \rangle = \langle \vec{v}_2, \vec{v}_1 \rangle$ and $\langle \vec{v}_1, \vec{v}_1 \rangle = 0 \iff \vec{v}_1 = 0$. Having this pair is the same as having notion of length and angles.

Gram-Schmidt for abstract functions

Generally, we may be able to define length and angles for functions for future applications. For example, we may define

$$\langle f(x), g(x) \rangle = \int_0^\pi f(x)g(x)dx$$

In the abstract word, if the vector is given $\vec{v} = f(x)$, then the meaning of transpose is in fact a linear transformation

$$\vec{v}^T = \langle f(x), - \rangle$$

and the projection can be written by

$$\frac{\vec{v}\vec{v}^T}{\vec{v}^T\vec{v}} = \frac{f(x)\langle f(x), - \rangle}{\langle f(x), f(x) \rangle}$$

Gram-Schmidt for abstract functions

This approach is extremely useful in Fourier analysis.

Excercise. Let W be vector space of polynomials with degree at most 2. Find an orthonormal basis for the inner product defined by

$$\langle f(x), g(x)\rangle = \int_0^1 f(x)g(x)dx.$$

Solution. Let $A = \begin{pmatrix} 1 & x & x^2 \end{pmatrix}$ We find the table of inner product

$$A^TA = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{3} & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{12} & 0 \\ 0 & 0 & \frac{1}{180} \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Gram-Schmidt for abstract functions

So an orthonormal basis is given by

$$\begin{pmatrix} 1 & x & x^2 \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{12} & 0 \\ 0 & 0 & \sqrt{180} \end{pmatrix}$$

Application of orthonormal basis

Let $v_1 = f_1(x)$ $v_2 = f_2(x)$ $v_3 = f_3(x)$ be orthonormal basis given as before. Note that this means

$$\underbrace{\begin{pmatrix} v_1^T \\ v_2^T \\ v_3^T \end{pmatrix}}_{A^T} \underbrace{\begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix}}_{A} = I.$$

This implies that

$$P := AA^T = v_1 v_1^T + v_2 v_2^T + v_3 v_3^T$$

is an orthogonal projection matrix projecting to space spanned by v_1, v_2, v_3. Note that for any vector w, we have

$$Pw = v_1 v_1^T w + v_2 v_2^T w + v_3 v_3^T w$$

$$= v_1 \langle v_1, w \rangle + v_2 \langle v_2, w \rangle + v_3 \langle v_3, w \rangle$$

Application of orthonormal basis

This means that for any function g, the orthogonal projection to the subspace is given by the formula

$$f_1(x) \cdot \int_0^1 f_1(x)g(x)dx + f_2(x) \cdot \int_0^1 f_2(x)g(x)dx + f_3(x) \cdot \int_0^1 f_3(x)g(x)dx$$

Therefore, finding an orthonormal basis can help writing down orthogonal projections clearly.