Orthogonal Projector

P. Sam Johnson

May 26, 2017

We study the concept of orthogonal projection and give an explicit expression for the orthogonal projector into the column space of a matrix.

Definition

For any set A of vectors in an inner product space V,
$A^\perp := \{y \in V : y \perp \text{ for every } x \in A\}$.

- $A^\perp$ is a subspace of V for any set $A \subseteq V$.
- **Constructing an ONB for $S^\perp$ from an ONB of S.** Let $\{x_1, x_2, \ldots, x_k\}$ be any ONB of a subspace S and let $\{x_1, x_2, \ldots, x_k, x_{k+1}, \ldots, x_n\}$ be any extension of B to an ONB of V. Then $S^\perp$ is the space of $\{x_{k+1}, \ldots, x_n\}$.
- If S is a subspace of V, then $S^\perp$ **is a complement of** S;

  $$\boxed{d(S^\perp) = d(V) = d(S) \text{ and } (S^\perp)^\perp = S.}$$

  Becasue of the reasons ($S^\perp$ is a complement of S and is orthogonal to S), we call $S^\perp$, the **orthogonal complement** of S.

Properties of Orthogonal Complements

- If W is a complement of S and is orthogonal to S, then $W = S^{\perp}$.
- The union of an ONB of S and an ONB of W is an ONB of V.
- Suppose $S_1, S_2, \ldots, S_k$ are subspaces which are orthogonal to one another and $S_1 + S_2 + \cdots + S_k = V$. Then $S_1 \oplus S_2 \oplus \cdots \oplus S_k = V$. Now for any fixed i $(1 \leq i \leq k)$, $\sum_{j \neq i} S_j$ is a complement of S_i and is orthogonal to S_i, so it is the orthogonal complement of S_i.
- If $S \subseteq T$, then $T^{\perp} \subseteq S^{T}$.
- If S and T are subspaces, then $(S + T)^{\perp} = S^{\perp} \cap T^{\perp}$ and $(S \cap T)^{\perp} = S^{\perp} + T^{\perp}$.
- The result $S = (S^{\perp})^{\perp}$ is **quite powerful** and is closed related to a result known as Farkas lemma which is equivalent to the duality theorem of linear programming.

Orthogonal projection into a subspace

Definition

If S is a subspace of V and $x \in V$, the projection of x into S **along** *$S^{\perp}$ is called the* **orthogonal projection** *of x into S.*

- Geometrically, the orthogonal projection of x into S is the foot of the perpendicular drawn from x to S.
- Since $(S^{\perp})^{\perp} = S$, it follows that if y is the orthogonal projection of x into S then $x - y$ is the orthogonal projection of x into $S^{\perp}$.
- Let $\{x_1, x_2, \ldots, x_k\}$ be an ONB of S. Then for any $x \in V$, y defined by $y = \sum_{i=1}^{k} \langle x, x_i \rangle x_i$, $(y \in S)$ is the orthogonal projection of x into S ; the residual $x - y$ is the orthogonal to each of $x_1, x_2, \ldots, x_k$, $(x - y \in S^{\perp})$; $x - y$ is the orthogonal projection of x into $S^{\perp}$.
- The residual of x with respect to an ONB of S does not depend on the choice of the basis. Residual, is really with respect to the subspace S.

Orthogonal projection into a flat

Let W be a flat. Then there is a (unique) subspace S and a vector u (not unique) in V such that $W = u + S$. Any $x \in V$, $x - u$ has unique expression $x - u = s + t, s \in S, t \in S^{\perp}$. Hence $x = (u + s) + t$, $u + s \in W, t \in S^{\perp}$. Thus any vector $x \in V$ can be written unqiuely as $w + t$, where $w \in W$ and $t \in S^{\perp}$.

The vector w is called the orthogonal projection of x into W. Geometrically, it is the foot of the perpendicular from x to W.

If P is the orthogonal projector into S, then $w = u + P(x - u)$, where $w \in W = u + S$, and S is a subspace.

Theorem

Let w be the orthogonal projection of x into a flat W. Then $d(x, W) = \min_{z \in W} \|x - z\|$ is attained at w and only at w.

Consider $\mathbb{R}^n$ and $\mathbb{C}^n$ equipped with the canonical inner product. For a real matrix A, $N(A) = R(A)^\perp = C(A^\perp)^\perp$. For a complex matrix A, $N(A) = C(A^*)^\perp$.

Proof.

$$\begin{aligned} x \perp C(A^*) &\iff \langle x, A^*z \rangle \text{ for all } z \\ &\iff (A^*z)x = z^*Ax = 0 \text{ for all } z \\ &\iff Ax = 0 \iff x \in N(A). \end{aligned}$$

Definition

Let S be a subspace of F^n. The **orthogonal projector into** S *is the $n \times n$ matrix P such that for every $x \in F^n$, P_x is the orthogonal projection of x into S.*

- An $n \times n$ matrix Q is the orthogonal projector if it is the orthogonal projector into some subspace S of F^n.
- $S = C(Q)$.
- Q is the projector into S along $S^\perp$.
- $S^\perp = C(I - Q)$.
- $I - Q$ is the projector into $S^\perp$.

TFAE

1. Q is an orthogonal projector.
2. $Q^* Q = Q$.
3. $Q^* = Q$ and $Q^2 = Q$.

Proof.

$$\begin{aligned} Q \text{ is an ort.proj.} &\iff Qx \text{ is the orth.proj. of } x \text{ into } C(Q), \forall\, x \in F^n. \\ &\iff x = Qx \perp C(Q) \forall\, x \in F^n. \\ &\iff \langle Qy, (I-Q)x \rangle = 0 \forall\, x\ y \in F^n. \\ &\iff (I-Q)^* = 0 \iff Q^*Q = Q. \end{aligned}$$

Hence $(a) \iff (c)$.

We next obtain an explicit formula for the orthogonal projector into the column space of an arbitrary matrix.

Theorem

The orthogonal projector P_A into $C(A)$ is given by $P_A = A(A^*A)^-A^*$, where B^- denotes a **generalized inverse** *of A.*

Proof. page 270

Remark

Page 270

Computation of P_A. Page 270

Transformation which preserves distances are called **isometries**. Linear transformations which preserve inner product (and so distances and angles) are isometries. We study their matrices.

We consider only the canonical inner product in $\mathbb{C}^n$ and $\mathbb{R}^n$.

Definition

A **unitary matrix** *is a complex square matrix A such that $A^*A = I$ (is equivalent to, $A^* = A^{-1}$). We know that if a square matrix A has a left inverse, then A has an inverse, so $A^*A = I$.*
An **orthogonal matrix** *is a real square matrix A such that $AA^T = I$ (is equivalent to, $A^T = A^{-1}$).*

$$(AA^*)_{ij} = \langle A_{i*}, A^*_{j*} \rangle = \langle A_{i*}, A_{j*} \rangle$$

and $(A^*A)_{ij} = \langle A_{*j}, A_{*i} \rangle$. Hence A is unitary iff the rows as well as the columns of A from orthonormal bases of F^n.

Unitary Matrices	Orthogonal Matrices
Examples : the identity matrix and all permutation matrices	Examples : the identity matrix and all permutation matrices
The unitary matrices of order 1 are $e^{i\theta}, 0 \leq \theta \leq 2\pi$.	The orthogonal matrices of order 1 are 1 and -1.
The determinant of an unitary matrix has modulus 1 because $I = AA^*$ and $1 = det(AA^*) = \|det(A)\|^2$.	The determinant of an orthogonal matrix is 1 or -1 because $I = AA^*$ and $1 = det(AA^*) = det(A)^2$, so $det(A) = 1$, or -1.

1. If A is unitary, the matrix obtained from A by any permutation of rows or columns is also unitary.
2. The matrix obtained by multiplying any row or column of a unitary matrix by a scalar of unit modulus is also unitary.
3. Any 2×2 orthogonal matrix A is

$$A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \text{ or } \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \text{ for some } \theta.$$

Proof. Let P and Q be the points in $\mathbb{R}^2$ corresponding to the two columns of A and Q be the angle between the x-axis and OP.

Then length $PO = 1$, so $P = (\cos\theta, \sin\theta)^T$.

Length OQ is also 1 and OQ is perpendicular to OP, so $Q = (\cos\phi, \sin\phi)^T$ where ϕ is $(\theta + \frac{\pi}{2})$ or $(\theta - \frac{\pi}{2})$. Hence A is A_θ or B_θ.

Definition

An orthogonal matrix is said to be **proper** *or* **improper** *according as its determinant is* 1 *or* -1.
Note that A_θ *is proper and* B_θ *is improper.*

Construction of Hermitian-unitary matrix from a vector $u \in \mathbb{C}^n$ with $\|u\| = 1$.

Let u be any vector in $\mathbb{C}^n$ with $\|u\| = 1$ and set $A = I - 2uu^*$.

Then A is Hermitian and $AA^* = A^2 = I - 4uu^* + 4uu^*uu^* = I$, since $u^*u = I$. Thus A is unitary.

If u is real and A is symmetric and orthogonal, then $A = I - 2uu^T$ is symmetric-orthogonal matrix.

Theorem

Let A be an $n \times n$ matrix. TFAE

1. *A is unitary.*
2. *$\langle Ax, Ay \rangle = \langle x, y \rangle$ for all $x, y \in \mathbb{C}^n$ (the map $x \mapsto Ax$ preserves angles).*
3. *$\|Ax\| = \|x\|$ for all $x \in \mathbb{C}^n$ (the map $x \mapsto Ax$ preserves length).*
4. *$\|Ax\| = 1$ whenever $\|x\| = 1$ and $x \in \mathbb{C}^n$ (the map $x \mapsto Ax$ leaves the surface of a sphere with centre at the origin, invariant.*
5. *$\|Ax - Ay\| = \|x - y\|$ for all $x, y \in \mathbb{C}^n$ (the map $x \mapsto Ax$ preserves distance).*
6. *$\{Ax_1, Ax_2, \ldots, Ax_n\}$ is an orthonormal basis of $\mathbb{C}^n$ whenever $\{x_1, x_2, \ldots, x_n\}$ is an orthonormal basis of $\mathbb{C}^n$.*

Further, if A is real, then 'unitary' can be replaced by 'orthogonal' in (i) and $\mathbb{C}^n$ by $\mathbb{R}^n$ in (ii) through (vi).

Written out if full, (ii) says: if we make a change of variables from $x_1, x_2, \ldots, x_n$ to $y_1, y_2, \ldots, y_n$ by $y = Ax$, where A is orthogonal, then

$$y_1^2 + y_2^2 + \cdots + y_n^2 = x_1^2 + x_2^2 + \cdots + x_n^2.$$

This is what makes orthogonal transformations useful in many subjects.

For example, this is used in Statistics to show that the sample mean and sample variance are independently distributed if the popoulation is normal.

Let A be a linear map. The map $K : x \mapsto Ax + c$ is known as an **affine transformation**.

If A is unitary, then K preserves distances. We prove a strong form of the converse.

Theorem

Let f be any map (not necessarily affine transformation) from $\mathbb{R}^n$ to itself such that $\|f(x) - f(y)\| = \|x - y\|$ for all $x, y \in \mathbb{R}^n$. Then there exist an orthogonal matrix A and a vector $c \in \mathbb{R}^n$ such that $f(x) = Ax + c$ for all $x \in \mathbb{R}^n$.

Proof. Page 276

We have seen that every 2×2 orthogonal matrix corresponds to either a rotation or a reflection of the plance depending upon whether it is proper or improper.

We will prove that every orthogonal matrix of order 3 corresponds to either a rotation of $\mathbb{R}^3$ about a line through the origin or such a rotation followed by a reflection in the origin, depending upon whether it is proper or improper.

Transition matrices and the effect of a change of bases on the matrix of a linear transformation. Page 277

Definition

Matrices A and B are **unitarily similar** *to each other if there exists an unitary matrix P such that $B = P^{-1}AP$.*

Constructing a large class of unitary matrices provided one can invert a matrix.

Definition

A **skew-hermitian matrix** *is a square matrix S such that $S^* = -S$. A real skew-hermitian matrix is said to be* **skew-symmetric**.

Let A be a square matrix such that $I + A$ is non-singular. Let $\tilde{A} = (I - A)(I + A)^{-1}$. Then $\tilde{A}$ is unitary iff A is skew-hermitian.

If A is real, $\tilde{A}$ is orthogonal iff A is skew-symmetric.

Let A be a square matrix such that $I + A$ is non-singular. Let $\tilde{A} = (I - A)(I + A)^{-1}$. Then $I + \tilde{A}$ is also singular and $\tilde{\tilde{A}} = A$. We will prove later that if S is skew-hermitian then $I + S$ is non-singular.

Thus $S \leftrightarrow \tilde{S}$ is a $1 - 1$ correspondence between skew-hermitian matrices and unitary matrices U such that $I + U$ is non-singular.

How to generate the skew-hermitian matrices? Put arbitrary purely imaginary numbers on the diagona, arbitrary complex numbers above the diagonal and then fill the cells below the diagonal by using $s_{ij} = -\overline{s}_{ji}$.

Now taking $\tilde{S}$ we get all unitary matrices U such that $I + U$ is non-singular.

Exercises

1. Let A be an $n \times n$ matrix. Show that the following statements are equivalent:
   (a) $Ax \perp Ay$ iff $x \perp y$,
   (b) A is non-zero scalar times a unitary martix,
   (c) the columns of A are orthogonal and have equal norms,
   (d) the rows of A are orthogonal and have equal norms.
2. Show that if A is unitary, then $C(I - A)$ and $N(I - A)$ are orthogonal complements.

Exercises

3. Show that the set of all $n \times n$ orthogonal matrices forms a group under multiplication. This group is denoted by O_n. Show that the set of all proper orthogonal matrices of order n forms a subgroup of O_n. This subgroup is denoted by SO_n.
4. Show that for an $m \times n$ matrix A, $\|Ax\| = \|x\|$ for all $x \in \mathbb{C}^n$ iff $A^*A = I$. (Such a rectangular matrix is called a **semi-unitary matrix**.
5. Given u and v in $\mathbb{R}^n$ with $\|u\| = \|v\|$, explain how an orthogonal matrix C can be obtained so that $Cu = v$.
6. Let u and x be fixed vectors in $\mathbb{R}^n$. Find the maximum and the minimum values of $(x^T Cu)^2$ as C varies over all $n \times n$ orthogonal matrices.

Reference

- A. Ramachandra Rao and P. Bhimasankaram, "*Linear Algebra*", Hindustan Book Agency, 2000.