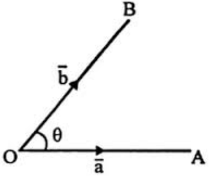
Inner Product and Orthogonality

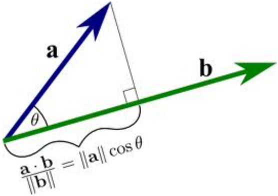
P. Sam Johnson

May 26, 2017

In the Euclidean space $\mathbb{R}^2$ and $\mathbb{R}^3$ there are two concepts, viz., length (or distance) and angle which have no analogues over a general field.

Fortunately there is a single concept usually known as scalar product or dot product which covers both the concepts of length and angle.





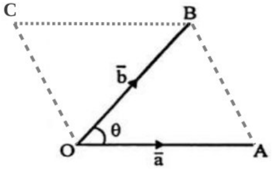
Let $a = (a_1, a_2)$ and $b = (b_1, b_2)$ be vectors in $\mathbb{R}^2$ represented by the points A and B as in figure. Then the scalar product of a and b is defined to be

$$\langle a, b\rangle = \ell_1 \ell_2 \cos\theta$$

where ℓ_1 is the length of OA, ℓ_2 is the length of OB and θ is the angle between OA and OB.

It can be shown using trigonometry that $\ell_1\ell_2\cos\theta = a_1b_1 + a_2b_2$, so $\langle a, b\rangle = a_1b_1 + a_2b_2$. **In Linear Algebra, scalar product is called inner product.**

Length, distance, angle : in terms of the inner product



1. **Length of a vector:** The length OA can be defined in terms of the inner product since
$$OA^2 = \langle a, a \rangle.$$
2. **Distance between vectors:** If $OABC$ is a parallelogram, the distance $AB = OC = \sqrt{\langle b-a, b-a \rangle}$ since $C = b - a$.
3. The angle θ can be obtained as
$$\theta = \cos^{-1}\left(\frac{\langle a, b \rangle}{\sqrt{\langle a, a \rangle . \langle b, b \rangle}}\right).$$

The above concepts and results have obvious analogues in $\mathbb{R}^3$.

We shall extend them to arbitrary (finite-dimensional) **vector spaces over** $\mathbb{R}$ **or** $\mathbb{C}$.

One does not extend inner product to vector spaces over a general field mainly because $\langle x, x\rangle \geq 0$ has no meaning in a general field.

Problems

1. Let z be a fixed nonnull vector in the plane. What is the locus of the point x such that $\langle x, z \rangle = 0$? What happens if 0 is replaced by a non-zero scalars?
2. If x_1, x_2, y_1, y_2 are real numbers, show that

$$(x_1 x_2 + y_1 y_2)^2 \leq (x_1^2 + y_1^2)(x_2^2 + y_2^2).$$

Hence deduce that $PQ + QR \geq PR$ for any three points P, Q and R in the plane.

Motivated by the usual inner product (dot product) on $\mathbb{R}^2$ we now give the axiomatic definition of inner product on a vector space over F, where F is either $\mathbb{R}$ or $\mathbb{C}$.

Definition

An inner product on a vector space V over F is a map $(x,y) \mapsto \langle x,y\rangle$ from $V \times V$ to F satisfying the following three conditions:

1. $\langle x,y\rangle = \overline{\langle y,x\rangle}$.
2. $\langle \alpha x + \beta y, z\rangle = \alpha\langle x,z\rangle + \beta\langle y,z\rangle$
3. $\langle x,x\rangle \geq 0$; $\langle x,x\rangle = 0 \iff x = 0$.

a vector space with an inner product	an **inner product space**
a real inner product space	a **Euclidean space**
a complex inner product space	a **unitary space**

Properties of an inner product

1. The restriction of an inner product to a subspace is an inner product.
2. In any inner product space, we have
   - $\langle x, \alpha y + \beta z \rangle = \overline{\alpha} \langle x, y \rangle + \overline{\beta} \langle x, z \rangle$.
   - $\langle 0, y \rangle = \langle x, 0 \rangle = 0$.
3. When the second argument is held fixed, **inner product is linear in the first argument**. Similarly, when the first argument is held fixed, **inner product is conjugate-linear in the second argument**.

Examples of an inner product

1. The inner product
$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i = y^T x$$
on $\mathbb{R}^2$ is called the **canonical inner product** on $\mathbb{R}^n$.
2. On $\mathbb{C}^n$, the **canonical inner product** is defined by
$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i} = y^* x.$$
y^*, the **adjoint** of y, to denote $\overline{y}^T$.
3. Fix any finite subset A of $\mathbb{R}$ with size $\geq n$. Let $V = \mathcal{P}_n$ over $\mathbb{R}$.
$$\langle p, q\rangle := \sum_{a \in A} p(a)q(a)$$
is an inner product on V.
4. $\langle A, B\rangle = tr(B^* A)$ is an inner product on $\mathbb{C}^{m \times n}$.

Examples of an inner product

1. On the vector space V of all real-valued continuous functions on an interval $[a, b]$,

$$\langle f, g\rangle = \int_a^b f(t)\, g(t)\, dt$$

defines an inner product.

2. If $h \in V$ is such that $h(t) > 0$ for all $t \in [a, b]$,

$$\langle f, g\rangle = \int_a^b h(t)f(t)\, g(t)\, dt$$

is also an inner product.

3. Let V be the vector space of all real-valued random variables with mean 0 and finite variance, defined on a fixed probability space. Let $F = \mathbb{R}$ and define $\langle x, y\rangle$ to be the covariance between x and y.

Examples of non-inner product spaces

1. $\langle x, y \rangle := y^T x$ and $\langle x, y \rangle := x^* y$ are **not** inner products on $\mathbb{C}^n$.
2. $\langle A, B \rangle = \sum_{i=1}^{n} a_{ii} \overline{b}_{ii}$ is **not** an inner product on $\mathbb{C}^{m \times n}$.
   What are all the axioms which are violated?

Inner product associated with a matrix

Let V be an inner product space and $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$ a basis of V. Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)^T$ and $\beta = (\beta_1, \beta_2, \dots, \beta_n)^T$ be the coordinate vectors of x and y respectively with respect to $\mathcal{B}$ and let $A = (a_{ij})$, where $a_{ij} = \langle u_j, u_i \rangle$. Then

$$\langle x, y \rangle := \langle \sum \alpha_i u_i, \sum \beta_j u_j \rangle = \sum \sum \overline{\beta_j} a_{ji} \alpha_i = \beta^* A \alpha. \tag{1}$$

The matrix A will satisfy the following conditions:

1. $A = A^*$
2. $\alpha^* A \alpha \geq 0$ for all $\alpha \in F^n$,
3. if $\alpha^* A \alpha = 0$ then $\alpha = 0$.

Matrix associated with an inner product

Conversely, if A is a matrix satisfying the above three conditions, then $\langle ., . \rangle$ defined by (1) is an inner product on V.

Suppose $A = B^*B$, where B is a matrix with n columns and rank n. Then

$$\langle x, y \rangle = y^*B^*Bx$$

is an inner product because

- $\overline{\langle y, x \rangle} = \overline{x^*B^*By} = (x^*B^*By)^* = y^*B^*Bx = \langle x, y \rangle.$
- $\langle x, x \rangle = (Bx)^*(Bx) \geq 0.$
- If $\langle x, x \rangle = 0$ then $Bx = 0$ and so $x = 0$.

We shall later show that any matrix A satisfying the above three conditions, can be written as B^*B for some non-singular B.

Concept of length : Norm

Inner combines the concepts of length and angle. We shall discuss the first concept, length.

Definition

A **norm** *on a (real or complex) vector space V is a map $x \mapsto \|x\|$ from V to $\mathbb{R}$ satisfying the following three conditions:*

1. $\|x\| \geq 0$; $x = 0$ *if* $\|x\| = 0$
2. $\|\alpha x\| = |\alpha|.\|x\|$
3. $\|x + y\| \leq \|x\| + \|y\|$.

A vector space together with a norm on it is called a **normed vector space** or **normed linear space**.

We shall prove that every inner product induces a norm. We shall give a family of norms which are not induced by inner product. For this we need some famous inequalities.

Concept of angle

Inner combines the concepts of length and angle. We shall discuss an important special case of the second concept, viz., the angle between two vectors being 90°.

Let V be an inner product space, $x, y \in V$. Let A, B be subsets of V.

$\langle x, y\rangle = 0$ (we write $x \perp$ y)	x and y are **orthogonal** to each other
$x \perp y$ for every pair of distinct vectors x, y in A	A is **orthogonal**
A is orthogonal and every vector in A has norm 1	A is **orthonormal**
every vector in A is orthogonal to every vector in B	A is **orthogonal** to B

1. $x \perp y \iff y \perp x$.
2. $0 \perp x$ for all x.
3. $x \perp x \iff x = 0$.
4. if $x \perp y, y \perp z$, then $x \perp (\alpha y + \beta z)$.
5. A set of vectors is orthogonal iff its elements are pair-wise orthogonal. **Is the corresponding statement for linear independence true?**. Linear independence is a property of the entire set whereas orthogonality is a property of pairs.
6. The empty set is orthonormal (in a vacuous sense).

Pythagoras theorem

In a **real inner product space**, if $x \perp y$, then

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

The converse is true for real inner product space but not for complex inner product space.

More generally,

$$\|\sum_{i=1}^{k} x_i\|^2 = \sum_{i=1}^{k} \|x_i\|^2$$

if $\{x_1, x_2, \ldots, x_k\}$ is orthogonal. The converse is not true for both real and complex inner product spaces.

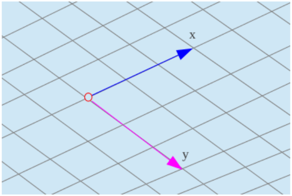
1. Any orthogonal set A not containing the null vector is linearly independent.
2. Any orthonormal set is linearly independent.
3. If the subspaces $S_1, S_2, \ldots, S_k$ are orthogonal to one another then $S_1 + S_2 + \cdots + S_k$ is direct.

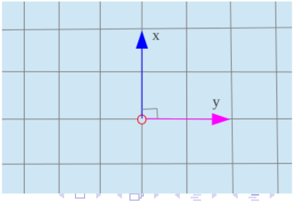
Definition

Let S be a subspace of an inner product space. We say that B is an **orthogonal basis** *(resp. an* **orthonormal basis***) of S if B is a basis of S and B is an orthonormal (resp. an orthonormal) set.*

We have seen that a basis corresponds to a **coordinate system**.

An orthonormal basis corresponds to a **system of rectangular coordinates** where the reference point on each axis is at unit distance from the origin.





For a given orthonormal basis, finding the coordinates with respect to such a coordinate system is easy as shown in the following.

Theorem

Let $B = \{x_1, x_2, \ldots, x_n\}$ be an orthonormal basis of an inner product space V. Then for any $x \in V$, we have

$$s = \sum_{j=1}^{n} \langle x, x_j \rangle x_j.$$

Suppose $A = \{x_1, x_2, \ldots, x_k\}$ be an orthogonal set (not a basis) of non-null vectors in V. Then for any $x \in V$, we call

$$z := x - \sum_{j=1}^{k} \frac{\langle x, x_j \rangle}{\langle x_j, x_j \rangle} x_j$$

the **residual of x with respect to A**. The residual z is orthogonal to each x_i.

Exercise.

Let $x_1, x_2, \ldots, x_k$ form an orthonormal set.

1. Show that $\| \sum_{i=1}^k \alpha_i x_i \|^2 = \sum_{i=1}^k \|\alpha_i\|^2$.
2. If z is the residual of x on $\{x_1, x_2, \ldots, x_k\}$, show that

$$\|z\|^2 = \|x\|^2 - \| \sum_{i=1}^{k} \langle x, x_i \rangle x_i \|^2 = \|x\|^2 - \sum_{i=1}^{k} |\langle x, x_i \rangle|^2.$$

3. **Bessel's inequality:**

$$\|x\|^2 \geq \sum_{i=1}^{k} |\langle x, x_i \rangle|^2$$

for any x. Show also that equality holds iff $x \in Sp(\{x_1, x_2, \ldots, x_k\})$.

Exercise

Let $B = \{x_1, x_2, \ldots, x_k\}$ be an orthonormal set in a finite-dimensional inner product space V. Show that the following statements are equivalent:

1. B is maximal,
2. $\langle x, x_i \rangle = 0$ for $i = 1, 2, \ldots, k \Rightarrow x = 0$,
3. B generates V,
4. if $x \in V$ then $x = \sum_{i=1}^{k} \langle x, x_i \rangle x_i$,
5. if $x, y \in V$ then $\langle x, y \rangle = \sum_{i=1}^{k} \langle x, x_i \rangle . \langle x_i, y \rangle$,
6. if $x \in V$ then $\|x\|^2 = \sum_{i=1}^{k} |\langle x, x_i \rangle|^2$.

Gram-Schmidt orthogonalization process

Theorem

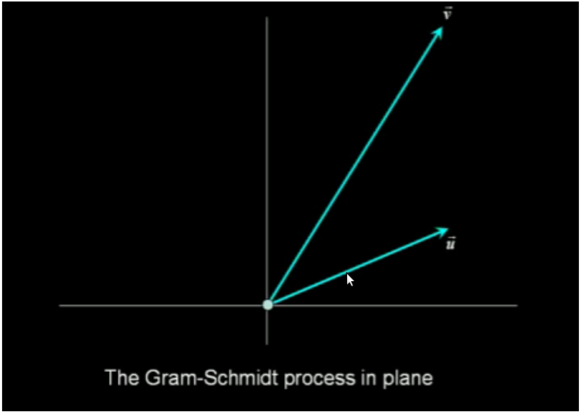
Let $\{x_1, x_2, \dots, x_k\}$ be a basis of a subspace S of an inner product space V. Define $z_1, z_2, \dots, z_k$ inductively by :

$$z_i = x_i - \sum_{j=1}^{i-1} \frac{\langle x_i, z_j \rangle}{\langle z_j, z_j \rangle} z_j \quad (i = 1, 2, \dots, k).$$

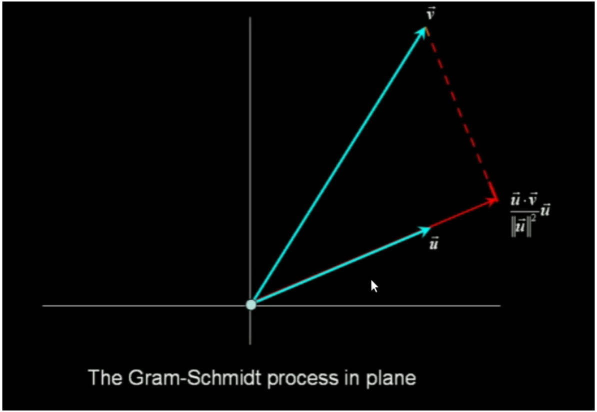
Then $z_1, z_2, \dots, z_k$ is an **orthogonal basis** *of S.*

- *An orthonormal basis of S can be obtained by normalizing the z_i's.*
- *Note that each z_i is the residual of x_i with respect to $z_1, z_2, \dots, z_{i-1}$; $z_1, z_2, \dots, z_k$ and $x_1, x_2, \dots, x_k$ have the same span.*
- *Let S be a subspace of a finite-dimensional inner product space V. Starting from any basis of S we can construct an orthonormal basis by the Gram-Schmidt process:* **Every subspace S of a finite-dimensional inner product space has an orthonormal basis.**

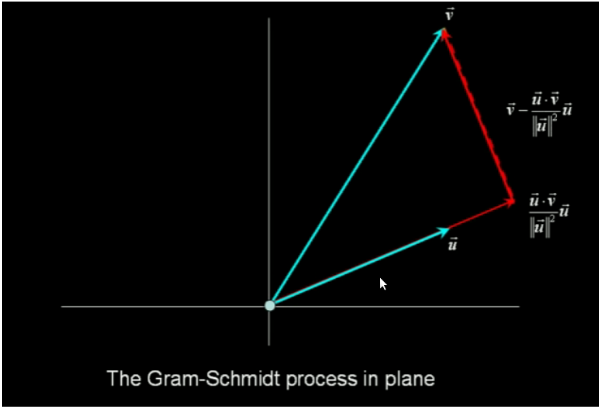
Gram-Schmidt process in plane



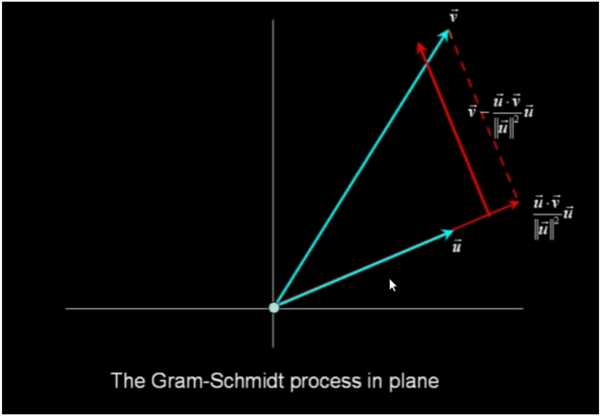
Gram-Schmidt process in plane



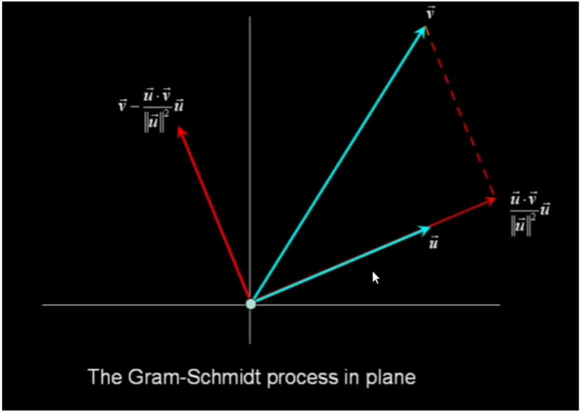
Gram-Schmidt process in plane



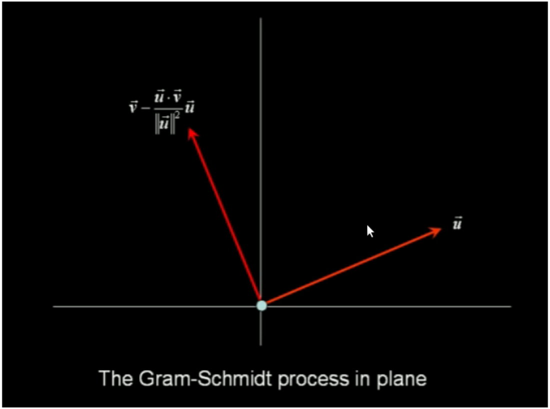
Gram-Schmidt process in plane



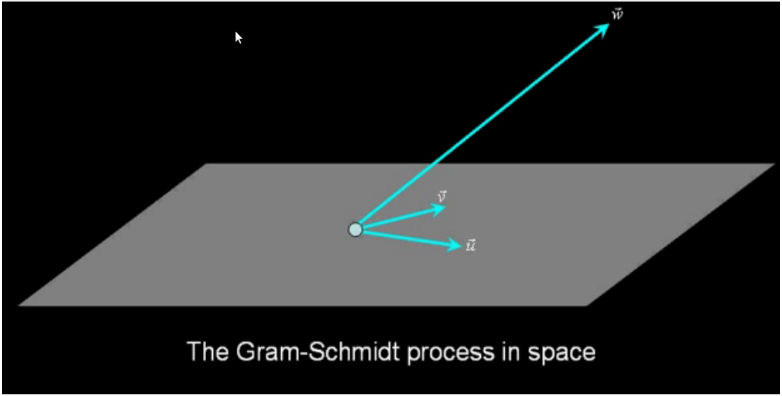
Gram-Schmidt process in plane



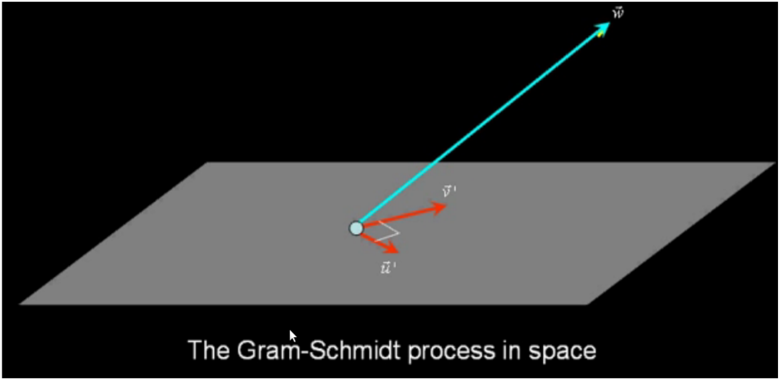
Gram-Schmidt process in plane



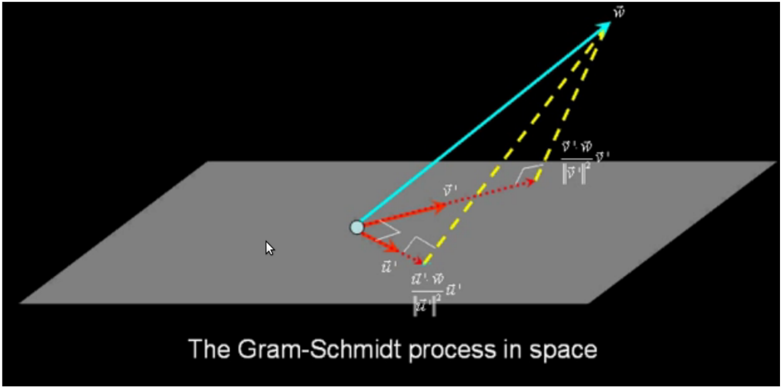
Gram-Schmidt process in space



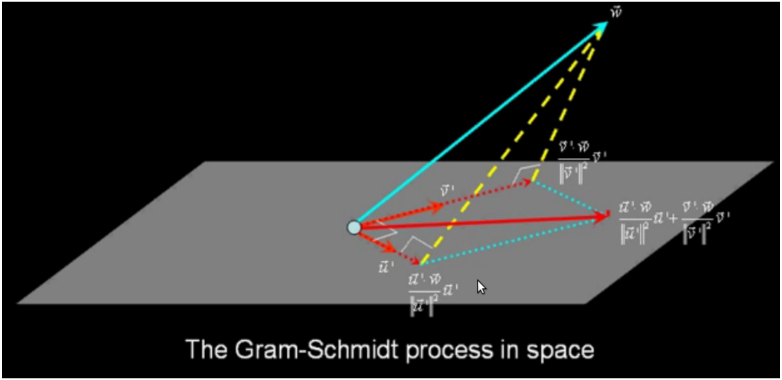
Gram-Schmidt process in space



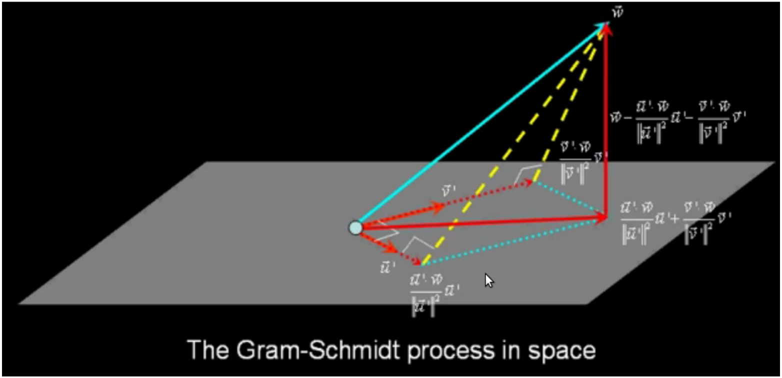
Gram-Schmidt process in space



Gram-Schmidt process in space

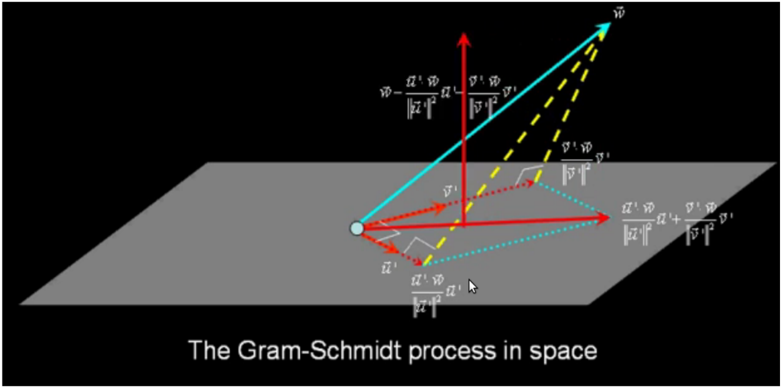


Gram-Schmidt process in space



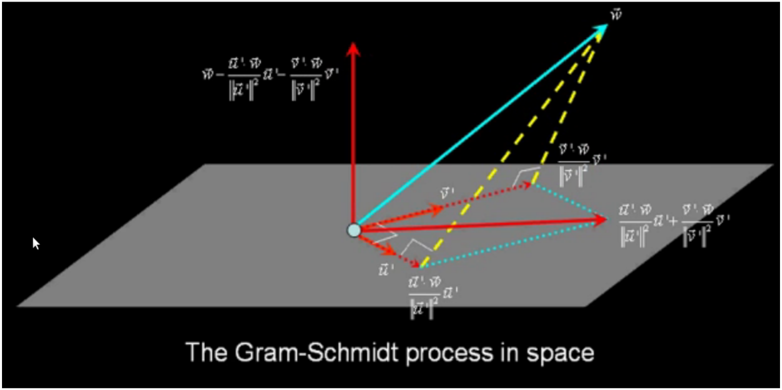
The Gram-Schmidt process in space

Gram-Schmidt process in space



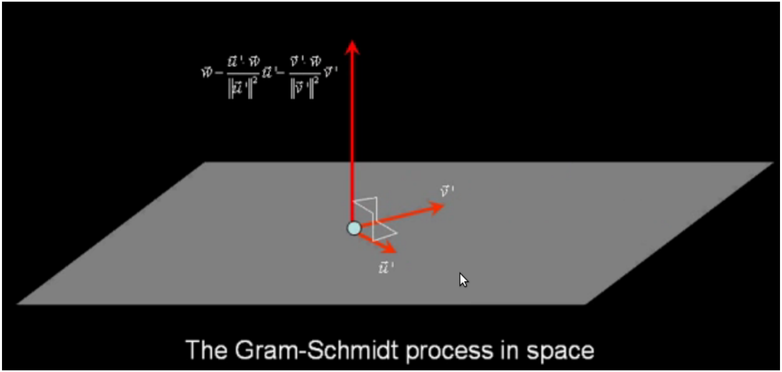
The Gram-Schmidt process in space

Gram-Schmidt process in space



The Gram-Schmidt process in space

Gram-Schmidt process in space



Generalized Gram-Schmidt Process

Let $x_1, x_2, \ldots, x_s$ be a given vectors in V, not necessarily basis.

1. **Step 1:** Set $k = 1$.
2. **Step 2:** Compute $z_k = x_k - \sum_{j=1}^{k-1} \frac{\langle x_k, y_j \rangle}{y} {}_j$.
3. **Step 3:** Compute $y_k := \frac{z_k}{\|z_k\|}$ or 0 according as $z_k \neq 0$ or $z_k = 0$.
4. **Step 4:** If $k < s$, increase k by 1 and go to Step 2. Otherwise go to Step 5.
5. **Step 5:** For $i = 1, 2, \ldots, s$, the set B_i of all non-null vectors **among** $y_1, y_2, \ldots, y_i$ is an orthonormal basis of the span S_i of $\{x_1, x_2, \cdots, x_i\}$.

If $x_1, x_2, \ldots, x_\ell$ form an orthonormal set then $y_j = x_j$ for $j = 1, 2, \ldots, \ell$.

Theorem

Let S be a subspace of a finite-dimensional inner product space V. Any orthonormal subset of S can be extended to an orthonormal basis of S.

Proof. Let $A = \{x_1, x_2, \ldots, x_\ell\}$ be an orthonormal subset of S. Extend A to a spanning set $\{x_1, x_2, \ldots, x_\ell, x_{\ell+1}, \ldots, x_s\}$ of S by **appending a basis**. Applying the generallized Gram-Schmidt process to $\{x_1, x_2, \ldots, x_s\}$, get $\{y_1, y_2, \ldots, y_s\}$. Then the non-null vectors among $y_1, y_2, \ldots, y_s$ form an orthonormal basis of S which contains $A = \{x_1, x_2, \ldots, x_\ell\}$.

We note that the orthonormal basis obtained by the Gram-Schmidt process from $x_1, x_2, \ldots, x_\ell$ may be quite different from that obtained from generallized Gram-Schmidt process (a rearrangement of $x_1, x_2, \ldots, x_\ell$).

QR-decomposition.

Let A be an $n \times s$ matrix with rank p. Let $y_1, y_2, \dots, y_s$ be the vectors obtained when generalized Gram-Schmidt process is applied to the columns of A. For each $k = 1, 2, \dots, s$,

$$z_k = A_{*k} - \sum_{j=1}^{k-1} \langle A_{*k}, y_j \rangle y_j = \|z_k\| y_k$$

and, $y_k := \frac{z_k}{\|z_k\|}$ or 0 according as $z_k \neq 0$ or $z_k = 0$.

Hence k-th column of A is a linear combination of $y_1, y_2, \dots, y_s$. That is,

$$A_{*k} = \langle A_{*k}, y_1 \rangle y_1 + \langle A_{*k}, y_2 \rangle y_2 + \cdots + \langle A_{*k}, y_{k-1} \rangle y_{k-1} + \|z_k\| y_k.$$

QR-decomposition.

$$A = [y_1 \ y_2 \ \cdots y_s] \begin{bmatrix} \|z_1\| & \langle A_{*2}, y_1\rangle & \langle A_{*3}, y_1\rangle & \cdots & \langle A_{*s}, y_1\rangle \\ 0 & \|z_2\| & \langle A_{*3}, y_2\rangle & \cdots & \langle A_{*s}, y_2\rangle \\ 0 & 0 & \|z_3\| & \cdots & \langle A_{*s}, y_3\rangle \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \|z_3\| \end{bmatrix}.$$

Let U be the $s \times s$ upper triangular matrix (u_{ik}) where

$$u_{ik} = \begin{cases} \langle A_{*k}, y_i\rangle & \text{if } i < k \\ \|z_k\| & \text{if } i = k \\ 0 & \text{otherwise} \end{cases}$$

Then $A = PU$.

Also if <u>Q is the submatrix of P</u> formed by the non-null columns (the columns of Q form an orthonormal basis, $Q^*Q = I_p$) and <u>R the submatrix of U</u> formed by the corresponding rows, then (Q, R) is a rank-factorization of A and $Q^*Q = I_p$.

When A is of full column rank $(Q, R) = (P, U)$ is known as a QR-**decomposition of** A.

<u>**Uniqueness.**</u> QR-factorization is unique if **we insist that the diagonal elements of** R **are real and positive**, i.e., if A is of full column rank, then there exist unique matrices Q and R such that $A = QR, Q^*Q = I$, R is upper triangular and $r_{ii} > 0$ for all i.

Exercises

1. Let x, y, u and v belong to $\mathbb{R}^n$. Then show that

$$\langle x + iy, u + iv \rangle := u^T x + v^T y$$

   is an inner product on the vector space $\mathbb{C}^n$ over $\mathbb{R}$.
   What is its connection with the canonical inner product on $\mathbb{C}^n$?

2. Show that $\langle x, y \rangle = 0$ for all y iff $x = 0$.

References

- S. Kumaresan, "*Linear Algebra - A Geometric Approach*", PHI Learning Pvt. Ltd., 2011.
- A. Ramachandra Rao and P. Bhimasankaram, "*Linear Algebra*", Hindustan Book Agency, 2000.