SC 639 (Autumn 2020) - Mathematical Structures for Control

Ravi N. Banavar [1]
Harsh Oza [1]

[1]Systems and Control Engineering,
IIT Bombay, India

June 28, 2021

Outline

1. Isomorphism
2. Rank and Nullity
3. Inner Product and Norm
4. Approximation

Outline

1. Isomorphism
2. Rank and Nullity
3. Inner Product and Norm
4. Approximation

Isomorphism

Isomorphism

A bijective linear transformation between two finite dimensional subspaces is called an **isomorphism**.

- Let $A : V \to W$ be a linear mapping between two subspaces.
- Subspaces V and W are called isomorphic to each other if the A is surjective and injective.
- If A is an isomorphism, then $dim(V) = dim(W)$.

Outline

1. Isomorphism
2. Rank and Nullity
3. Inner Product and Norm
4. Approximation

Range space and Null space

Range Space

Let $A : V \to W$ be a linear transformation. The **range space (R(A))** is given by following.

$$R(A) = \{w \in W | w = Av, \forall v \in V\} \tag{1}$$

Null space

Let $A : V \to W$ be a linear transformation. The **null space (N(A))** is given by following.

$$N(A) = \{v \in V | Av = 0, \forall v \in V\} \tag{2}$$

- R(A) is a subspace of W.
- N(A) is a subspace of V.
- A is one-one if $N(A) = \{0\}$

Rank and Nullity

- The dimension of Range space of A is called **Rank of A ($\rho(A)$)**.
- The dimension of Null space of A is called **Nullity of A ($\nu(A)$)**.

Rank-Nullity Theorem

Let $A : V \rightarrow W$ be linear with $dim(V) = k$. Then,

$$\rho(A) + \nu(A) = k \tag{3}$$

Proof

- Let $\nu(A) = m$ and basis for null space of A be $\{u_1, u_2, \dots, u_m\}$.
- Extend the set as $\{u_1, u_2, \dots, u_n, w_1, w_2, \dots, w_n\}$ to be a basis of the vector space V. Thus $dim(V) = m + n$
- We need to prove that $R(A)$ is finite and $dim(R(A)) = n$
- Let $v \in V$

$$v = a_1 u_1 + a_2 u_2 + \ldots + a_n u_n + b_1 w_1 + b_2 w_2 + \ldots + b_n w_n$$

applying transformation A on both sides,

$$Av = b_1 Aw_1 + b_2 Aw_2 + \ldots + b_n Aw_n$$

Here, Au_i disappears because $u_i \in N(A)$

- This implies $\{Aw_1, Aw_2, \dots, Aw_n\}$ spans the entire range of A.

Proof

- Now, we show $\{Aw_1, Aw_2, \ldots, Aw_n\}$ is linearly independent.

$$c_1Aw_1 + c_2Aw_2 + \ldots + c_nAw_n = 0$$

Then,

$$A(c_1w_1 + c_2w_2 + \ldots + c_nw_n) = 0$$

hence, $(c_1w_1 + c_2w_2 + \ldots + c_nw_n) \in N(A)$

$$(c_1w_1 + c_2w_2 + \ldots + c_nw_n) = (d_1u_1 + d_2u_2 + \ldots + d_nu_n)$$

- This implies all c and d are zero because $\{u_1, u_2, \ldots, u_n, w_1, w_2, \ldots, w_n\}$ is linearly independent.
- Thus,$\{Aw_1, Aw_2, \ldots, Aw_n\}$ is linearly independent and spans the range of A.

Outline

1 Isomorphism

2 Rank and Nullity

3 Inner Product and Norm

4 Approximation

Norm

Norm

Given a vector space V over a field $\mathcal{F}$, a norm on V is a non-negative valued function $||\cdot|| : V \rightarrow$ with following properties.

Let, $v, w \in V$ and if a is a scalar then,

1. $||av|| = |a| \cdot ||v||$,
2. $||v|| \geq 0$, with equality if and only if $v = 0_V$,
3. $||v + w|| \leq ||v|| + ||w||$,

Examples of Norm

1. Let $v \in \mathbb{R}^n$. The p-norm is given by,

$$||v||_p = \sqrt[p]{\sum_{i=1}^{n} |v_i|^p}$$

2. Let $A \in \mathbb{R}^{m \times n}$. The Frobenius norm is given by, $||A|| = \sqrt{Tr(A^T A)}$
3. Let V be the vector space consisting of all continuous functions on $[a, b]$ and $f \in V$. Then,$||f||_1 = \int_a^b |f(x)|dx$
4. The 2-norm of a signal $f(t)$ is $||f||_2 = \left(\int_{-\infty}^{\infty} f(t)^2 dt \right)^{1/2}$

Inner Product

Inner Product

An inner product in a vector space is a numerically valued function of ordered pair of vectors v and w, such that

1. $\langle \alpha_1 v_1 + \alpha_2 v_2, w \rangle = \alpha_1 \langle v_1, w \rangle + \alpha_2 \langle v_2, w \rangle$
2. $\langle v, w \rangle = \overline{\langle w, v \rangle}$;
3. $\langle v, v \rangle \geq 0; \qquad \langle v, v \rangle = 0$ if and only if $v = 0$

Examples:

1. Let $u, v \in V = \mathbb{R}^n$. Then, $\langle u, v \rangle = u^T v$ is an inner product.
2. Let V be the vector space consisting of all continuous functions and $f, g \in V$. Then, $\langle f, g \rangle = \int_0^1 f(t)g(t)\, dt$ is an inner product.
3. Let $u, v \in^2$. Then $\langle u, v \rangle = u_1 v_1 + 10 u_2 v_2$ is an inner product.

Norm induced by inner product

Norm (induced by inner product)

Let V be an inner product space. The norm of a vector $v \in V$ is defined to be a scalar $||v|| = \sqrt{\langle v, v\rangle}$.

Let V be an inner product space. If $v, w \in V$ and if a is a scalar then, it satisfies following additional properties.

1. $||v+w||^2 + ||v-w||^2 = 2(||v||^2 + ||w||^2)$
2. $||v-w|| \geq |||v|| - ||w|||$

Bessel's Inequality

Theorem

Let V be an inner product space and $u = \{u_1, u_2, \ldots, u_n\}$ be an orthonormal set in V. Let $x \in V$ be arbitary. Define $u \in V$ by $u = \sum_{k=1}^{n} \langle x, u_k \rangle u_k$. Then,

$$||u||^2 = \sum_{k=1}^{n} |\langle x, u_k \rangle|^2 \leqslant ||x||^2 \tag{4}$$

Proof.

Let $g = \sum_{k=1}^{n} \langle x, u_k \rangle u_k$

$$||g||^2 = \langle g, g \rangle = \sum_{k=1}^{n} \langle x, u_k \rangle^2 ||u_k||^2 = \sum_{k=1}^{n} \langle x, u_k \rangle^2 \tag{5}$$

For each $x \in V$ we have that,

$$\begin{aligned} 0 \leq ||x - g||^2 &= \langle x - g, x - g \rangle \\ &= ||x||^2 - 2\langle x, g \rangle = ||g||^2 \\ &= ||x||^2 - 2\langle x, \sum_{k=1}^{n} \langle x, u_k \rangle u_k \rangle + ||g||^2 \\ &= ||x||^2 - 2\sum_{k=1}^{n} |\langle x, u_k \rangle|^2 + ||g||^2 \\ &= ||x||^2 - ||g||^2 \end{aligned} \tag{6}$$

Therefore, $||g||^2 \leq ||x||^2$, which implies that, $||g|| \leq ||x||$.

$$\sum_{k=1}^{n} |\langle x, u_k \rangle|^2 \leqslant ||x||^2 \tag{7}$$

Cauchy-Schwarz Inequality

Recall: We may write a vector u as a scalar multiple of a nonzero vector v, plus a vector orthogonal to the vector v.

$$u = \frac{\langle u, v\rangle}{||v||^2} v + \left(u - \frac{\langle u, v\rangle}{||v||^2} v\right) \tag{8}$$

Theorem

Let V be an inner product space and $u, v \in V$. Then,

$$|\langle u, v\rangle| \leqslant ||u||\,||v|| \tag{9}$$

Proof.

Let $u, v \in V$. If $v = 0$, then both sides of (9) equal 0 and the desired inequality holds. Thus we assume that $v \neq 0$. Consider orthogonal decomposition,

$$u = \frac{\langle u, v\rangle}{||v||^2} v + w$$

By Pythagorean theorem,

$$\begin{aligned} ||u||^2 &= ||\frac{\langle u, v\rangle}{||v||^2} v||^2 + ||w||^2 \\ &= \frac{|\langle u, v\rangle|^2}{||v||^2} + ||w||^2 \\ &\geq \frac{|\langle u, v\rangle|^2}{||v||^2} \end{aligned}$$

Multiplying both sides of this inequality by $||v||^2$ and taking square root gives the Cauchy-Schwarz inequality. □

Adjoint Transformation

- Let V be any vector space and V' be its dual space. A is a linear transformation V.
- For each fixed $y \in V'$, the function y' defined by $y'(x) = [Ax, y]$ is a linear functional on V.
- The adjoint $A' : V' \to V$ is defined by the following property.

$$[Ax, y] = [x, A'y] \tag{10}$$

Adjoint Transformation (based on inner product)

Let V and W be inner product spaces and $A : V \to W$ be a linear transformation. A linear transformation $A' : W \to V$ satisfying the condition $\langle A(v), w \rangle = \langle v, A'(w) \rangle$ for all $V \in V$ and $W \in W$ is called an adjoint transformation of A.

Let V, W and Y be inner product spaces. Let $A : v \to W$, $B : v \to W$ and $C : W \to Y$ be linear transformations having adjoints and k be a scalar. Then,

1. $(A + B)' = A' + B'$
2. $(kA)' = \bar{k}A'$
3. $(CA)' = A'C'$
4. $(A')' = A$

Riesz Representation Theorem

Riesz Representation Theorem

Let V be an inner product space. If $\delta \in V'$ then there exists a unique vector $y \in V$ satisfying $\delta(v) = \langle v, y \rangle$ for all $v \in V$.

Example: Let $n > 1$ be an integer and let V be the subspace of $\mathcal{P}_n$ consisting of all polynomial functions of degree at most n, on which we have an inner product defined by $\langle f, g \rangle = \int_{-1}^{1} f(t)g(t)\, dt$. Let $\delta \in V'$ be a linear functional defined by $\delta : f \mapsto f(0)$. Then there exists a polynomial function $p \in V$ satisfying the condition $f(0) = \int_{-1}^{1} f(t)p(t)\, dt$ for all $f \in V$.

Projection

Projection

If V is the direct sum of $\mathcal{M}$ and $\mathcal{N}$, so that every $z \in V$ maybe written, uniquely, in form of $z = x + y$, with $x \in \mathcal{M}$ and $y \in \mathcal{N}$, the projection on $\mathcal{M}$ along $\mathcal{N}$ is the transformation E defined by, $Ez = x$.

- A linear transformation E is a projection on some subspace if and only if it is idempotent, that is, $E^2 = E$.

Proof: If E is the projection on $\mathcal{M}$ along $\mathcal{N}$, and if $z = x + y$, then

$$E^2 z = EEz = Ex = x = Ez \tag{11}$$

Conversely, suppose $E^2 = E$. Let $\mathcal{N}$ be a set of all vectors $z \in V$ such that $Ez = 0$. Let $\mathcal{M}$ be a set of all vectors $z \in V$ such that $Ez = z$.
For an arbitary z we have,

$$z = Ez + (1 - E)z$$

If we write $Ez = x$ and $(1 - E)z = y$, then $Ex = E^2 z = Ez = x$ and $Ey = E(1 - E)z = Ez - E^2 z = 0$, so that $x \in \mathcal{M}$ and $y \in \mathcal{N}$. This proves $V = \mathcal{M} \bigoplus \mathcal{N}$ and that the projection on $\mathcal{M}$ along $\mathcal{N}$ is E.

Properties of Projections

- If E is a projection on $\mathcal{M}$ along $\mathcal{N}$, then $\mathcal{M}$ and $\mathcal{N}$ are, respectively, the sets of all solutions of equations $Ez = z$ and $Ez = 0$.
- A linear transformation E is a projection if and only if $1 - E$ is a projection. If E is a projection on $\mathcal{M}$ along $\mathcal{N}$ then $1 - E$ is the projection on $\mathcal{N}$ along $\mathcal{M}$.
- $E_1 + E_2$ is a projection $\iff$ $E_1E_2 = E_2E_1 = 0$; then $E = E_1 + E_2$ is a projection on $\mathcal{M}$ along $\mathcal{N}$, where $\mathcal{M} = \mathcal{M}_1 \bigoplus \mathcal{M}_2$ and $\mathcal{N} = \mathcal{N}_1 \cap \mathcal{N}_2$
- $E_1 - E_2$ is a projection $\iff$ $E_1E_2 = E_2E_1 = E_2$; then $E = E_1 - E_2$ is a projection on $\mathcal{M}$ along $\mathcal{N}$, where $\mathcal{M} = \mathcal{M}_1 \cap \mathcal{N}_2$ and $\mathcal{N} = \mathcal{N}_1 \bigoplus \mathcal{M}_2$
- If $E_1E_2 = E_2E_1 = E$, then E is a projection on $\mathcal{M}$ along $\mathcal{N}$, where $\mathcal{M} = \mathcal{M}_1 \cap \mathcal{M}_2$ and $\mathcal{N} = \mathcal{N}_1 + \mathcal{N}_2$
- If $\mathcal{M}$ is a subspace invariant under linear transformation A, then $EAE = AE$ for every projection E on $\mathcal{M}$. Conversely, if $EAE = AE$ for some projection E on $\mathcal{M}$, then $\mathcal{M}$ is invariant under linear transformation A.

Outline

1. Isomorphism
2. Rank and Nullity
3. Inner Product and Norm
4. Approximation

Least Squares Method

- Let a linear system be given by $Ax = b$, where $A \in^{m\times n}, x \in^m$ and $b \in^n$ then, the aim is to find best approximation of b which is in the range space of A.
- The problem is converted to minimizing the error $e = ||Ax - b||^2$.
- The solution $\hat{x}$ is which minimizes e is the same as locating the point $p = A\hat{x}$ that is closer to b than any other point in the column space of A.
- The error vector e must be perpendicular to the column space.
- All vectors perpendicular to the column space lie in the left nullspace. Thus the error vector e must be in the nullspace of A^T:

$$\begin{aligned} A^T(b - A\hat{x}) &= 0 \\ \hat{x} &= (A^T A)^{-1} A^T b \\ p = A\hat{x} &= A\left((A^T A)^{-1} A^T b\right) \end{aligned} \tag{12}$$

Example

Let a linear system be $Ax = b$ where,

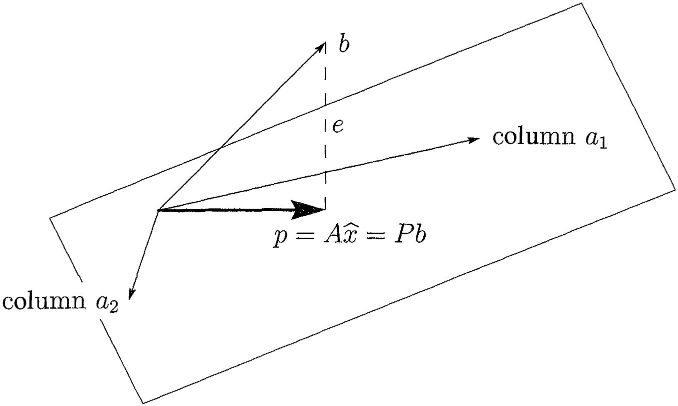
$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$

Here, b does not lie in the column space of A. So we have to find a best approximation which lies in column span of A.
By (12) we get,

$$p = \begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix}$$

Geometrically, p is an orthogonal projection of b on column space of A.

Geometric representation of least squares approximation



Example

Problem: Sheldon wants to buy an apartment. But he wants to make sure that he pays the right market value for the apartment. So he decides to do a survey and tabulates them as under.

Area (in hundred Sq. ft)	12	10	15	6
No. of Bedrooms	2	3	4	1
Price(in Lacs)	70	75	85	50

Now, Sheldon has finalized a 750 Sq. ft. and 3 bedroom apartment. What do you think he should pay? (For 0 Sq. ft. and 0 bedroom apartment the price should be zero.)

Example

$$
\begin{aligned}
12x + 2y &= 70 \\
10x + 3y &= 75 \\
15x + 4y &= 85 \\
60x + 1y &= 50
\end{aligned}
$$

We write it as

$$AX = b$$

where,

$$A = \begin{pmatrix} 12 & 2 \\ 10 & 3 \\ 15 & 4 \\ 6 & 1 \end{pmatrix} \quad X = \begin{pmatrix} x \\ y \end{pmatrix} \quad b = \begin{pmatrix} 70 \\ 75 \\ 85 \\ 50 \end{pmatrix}$$

$$
\begin{aligned}
X &= (A^T A)^{-1} A^T b \\
&= \begin{pmatrix} 5.800 \\ 1.967 \end{pmatrix}
\end{aligned}
$$

Example

Now we have to find the value of a 750 sq. ft. and 3 bedroom apartment.

$$
\begin{aligned}
Price &= 7.5x + 3y \\
&= 49.4
\end{aligned}
$$

Hence, Sheldon should pay 49.4 Lacs for the apartment.

References

1. Halmos, Paul R. Finite-dimensional vector spaces. Courier Dover Publications, 2017.
2. Golan, Jonathan S. The Linear Algebra. Springer, 2007.
3. Strang, Gilbert. Linear Algebra and its applications. Cengage Learning, 2005.

Thank You