Linear Algebra Side Notes 2024

timetraveler314
2024 Spring
timetraveler314@outlook.com

Contents

1. Matrices

1.1. Blocked Matrices

1.1.1. Sylvester's Theorem

Sylvester's Theorem is a fantastic result using the technique of blocked matrices. The idea can be tweaked and applied to many other problems. Here are some examples.

Theorem 1.1.1.1 The First Rank Reduction Theorem

Let $\boldsymbol{A}$ be invertible, then

$$\operatorname{rank}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} = \operatorname{rank} \boldsymbol{A} + \operatorname{rank}(\boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}).$$

Proof: Apply elementary blocked matrix operations to the matrix $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}$:

$$\begin{pmatrix} \boldsymbol{I} & 0 \\ -\boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I} \end{pmatrix}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}\begin{pmatrix} \boldsymbol{I} & -\boldsymbol{A}^{-1}\boldsymbol{B} \\ 0 & \boldsymbol{I} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} & 0 \\ 0 & \boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B} \end{pmatrix}.$$

The result follows. □

Theorem 1.1.1.2 The Second Rank Reduction Theorem

Let $\boldsymbol{A}, \boldsymbol{D}$ be invertible, then

$$\operatorname{rank} \boldsymbol{A} + \operatorname{rank}(\boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}) = \operatorname{rank} \boldsymbol{D} + \operatorname{rank}(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}).$$

Proof: Similar to the first theorem. □

Theorem 1.1.1.3 Sylvester's Inequality

$$\operatorname{rank} \boldsymbol{AB} \geq \operatorname{rank} \boldsymbol{A} + \operatorname{rank} \boldsymbol{B} - n.$$

Proof: Apply the first rank reduction theorem, or use the trick on

$$\begin{pmatrix} \boldsymbol{A} & 0 \\ \boldsymbol{I} & \boldsymbol{B} \end{pmatrix} \to \begin{pmatrix} 0 & \boldsymbol{AB} \\ \boldsymbol{I} & 0 \end{pmatrix}.$$

The result follows from

$$\operatorname{rank} \begin{pmatrix} \boldsymbol{A} & 0 \\ \boldsymbol{I} & \boldsymbol{B} \end{pmatrix} \geq \operatorname{rank} \boldsymbol{A} + \operatorname{rank} \boldsymbol{B}.$$

Note 1.1.1.1

This is a direct result of the Frobenuis Inequality.

□

Theorem 1.1.1.4 Frobenuis Inequality

$$\operatorname{rank} \boldsymbol{ABC} \geq \operatorname{rank} \boldsymbol{AB} + \operatorname{rank} \boldsymbol{BC} - \operatorname{rank} \boldsymbol{B}.$$

Proof:

$$\begin{pmatrix} \boldsymbol{ABC} & 0 \\ 0 & \boldsymbol{B} \end{pmatrix} \to \begin{pmatrix} \boldsymbol{ABC} & \boldsymbol{AB} \\ 0 & \boldsymbol{B} \end{pmatrix} \to \begin{pmatrix} 0 & \boldsymbol{AB} \\ -\boldsymbol{BC} & \boldsymbol{B} \end{pmatrix}.$$

□

As for eigenvalues and eigenvectors, we want to obtain the form $\lambda \boldsymbol{I} - \boldsymbol{AB}$, and this can be done similarly.

Theorem 1.1.1.5 Sylvester's Theorem for Eigenvalues

Let $\boldsymbol{A}$ be an $m \times n$ matrix, $\boldsymbol{B}$ be an $n \times m$ matrix, and $m \geq n$. We have

$$|\lambda \boldsymbol{I}_m - \boldsymbol{AB}| = \lambda^{m-n} |\lambda \boldsymbol{I}_n - \boldsymbol{BA}|.$$

Proof: A direct result of the first theorem when $\lambda \neq 0$.

When $\lambda = 0$, it can either be done by the Cauchy-Binet formula or by the perbutation method.
□

The theorem reveals that $\boldsymbol{AB}$ and $\boldsymbol{BA}$ have the same non-zero eigenvalues with the same algebraic multiplicities. Furthermore, we can actually show that their geometric multiplicities are also the same.

Corollary 1.1.1.1

Let $\boldsymbol{A}$ be an $m \times n$ matrix, $\boldsymbol{B}$ be an $n \times m$ matrix. Then $\boldsymbol{AB}$ and $\boldsymbol{BA}$ have the same non-zero eigenvalues with the same algebraic and geometric multiplicities.

Proof: First theorem with

$$\begin{pmatrix} \lambda \boldsymbol{I}_m & 0 \\ \boldsymbol{B} & \lambda I_n - \boldsymbol{BA} \end{pmatrix} \leftarrow \begin{pmatrix} \lambda \boldsymbol{I}_m & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_n \end{pmatrix} \to \begin{pmatrix} \lambda \boldsymbol{I}_m - \boldsymbol{AB} & \boldsymbol{A} \\ 0 & \boldsymbol{I}_n \end{pmatrix}.$$

$$n + (n - \dim \ker(\lambda \boldsymbol{I}_n - \boldsymbol{BA})) = n + (n - \dim \ker(\lambda \boldsymbol{I}_m - \boldsymbol{AB})).$$

□

2. Polynomial

2.1. Coprime Polynomials and Linear Transformations

Coprime factorization of polynomials elicits the direct sum decomposition of the vector space. From the geometric perspective, this is the starting point of the Jordan Canonical Form.

Here are some results, all of which use the Bezout's Theorem:

Lemma 2.1.1

Let $f(x), g(x)$ be coprime polynomials, $\boldsymbol{A}$ be a matrix such that $f(\boldsymbol{A}) = 0$. Then $g(\boldsymbol{A})$ is invertible.

Proof: By the Bezout's Theorem, we have $u(x)f(x) + v(x)g(x) = 1$. Then

$$u(\boldsymbol{A})f(\boldsymbol{A}) + v(\boldsymbol{A})g(\boldsymbol{A}) = \boldsymbol{I}.$$

Given that $f(\boldsymbol{A}) = 0$, we have $g(\boldsymbol{A})v(\boldsymbol{A}) = \boldsymbol{I}$, which implies that $g(\boldsymbol{A})$ is invertible. □

Theorem 2.1.1

Let $f(x), g(x)$ be coprime polynomials, $\boldsymbol{A}$ be a matrix of order n. Then

$$f(\boldsymbol{A})g(\boldsymbol{A}) = 0 \Leftrightarrow \operatorname{rank} f(\boldsymbol{A}) + \operatorname{rank} g(\boldsymbol{A}) = n.$$

Proof:

$$\begin{pmatrix} f(\boldsymbol{A}) & 0 \\ 0 & g(\boldsymbol{A}) \end{pmatrix} \to \begin{pmatrix} f(\boldsymbol{A}) & f(\boldsymbol{A})u(\boldsymbol{A}) \\ 0 & g(\boldsymbol{A}) \end{pmatrix} \to \begin{pmatrix} f(\boldsymbol{A}) & \boldsymbol{I} \\ 0 & g(\boldsymbol{A}) \end{pmatrix} \to \begin{pmatrix} f(\boldsymbol{A}) & \boldsymbol{I} \\ -f(\boldsymbol{A})g(\boldsymbol{A}) & 0 \end{pmatrix}$$
$$\to \begin{pmatrix} 0 & \boldsymbol{I} \\ -f(\boldsymbol{A})g(\boldsymbol{A}) & 0 \end{pmatrix} \to \begin{pmatrix} f(\boldsymbol{A})g(\boldsymbol{A}) & 0 \\ 0 & \boldsymbol{I} \end{pmatrix}.$$

The result follows. □

This result is particularly useful when proving the diagonalizability of matrices that satisfy certain polynomial equations.

Question 2.1.1

Let $\boldsymbol{A}$ be an idempotent matrix, prove that A is diagonalizable.

Proof: Let $f(x) = x^2 - x$. Then $f(\boldsymbol{A}) = 0$. By the previous theorem, we have $\operatorname{rank} \boldsymbol{A} + \operatorname{rank}(\boldsymbol{A} - \boldsymbol{I}) = n$. This implies that $\boldsymbol{A}$ is diagonalizable. □

Question 2.1.2

Let T be a linear transformation. Prove that

$$\ker \gcd(f, g)(T) = \ker f(T) \cap \ker g(T),$$
$$\ker \operatorname{lcm}(f, g)(T) = \ker f(T) + \ker g(T).$$

Proof: Let $\beta \in \ker f(T) \cap \ker g(T)$, then by the Bezout's Theorem, we have

$$u(x)f(x) + v(x)g(x) = \gcd(f, g)(x).$$

Then

$$u(T)f(T)\beta + v(T)g(T)\beta = \gcd(f, g)(T)\beta = 0,$$

which implies that $\beta \in \ker \gcd(f, g)(T)$.

Let $\beta \in \ker \gcd(f, g)(T)$, then because $f = \tilde{f} \gcd(f, g)$,

$$f(T)\beta = \tilde{f}(T)d(T)\beta = 0.$$

The result follows. □

Proof: If $v \in \ker\left(d\tilde{f}\tilde{g}\right)(T)$, then

$$\begin{aligned} v &= \left(a\tilde{f} + b\tilde{g}\right)(T)v \text{ (Bezout)} \\ &= \underbrace{u(T)f(T)v}_{\in \ker g(T)} + \underbrace{v(T)g(T)v}_{\in \ker f(T)}. \end{aligned}$$

□

Example 2.1.1

If $P \in \operatorname{Hom}(V, V)$ is a projection s.t. $P^2 = P$. Then $V = \ker P \oplus \operatorname{range} P$.

Proof: From the perspective of space decomposition:

Lemma 2.1.2

($I - P$ is also a projection)

$$\ker(I - P) = \operatorname{im} P.$$

Because $\gcd(x, 1-x) = 1, \operatorname{lcm}(x, 1-x) = x - x^2$, we have $V = \ker P \oplus \ker(I - P) = \ker P \oplus \operatorname{im} P$. □

Generally, the similar decomposition can be applied like this:

Theorem 2.1.2

Let $f(x), g(x)$ be coprime polynomials, φ be a linear transformation such that $f(\varphi)g(\varphi) = 0$. Prove that $V = V_1 \oplus V_2$ where $V_1 = \ker f(\varphi)$ and $V_2 = \ker g(\varphi)$.

Proof: By assumption, there exists $u(x), v(x)$ such that $u(x)f(x) + v(x)g(x) = 1$. Then substituting φ into the equation gives

$$u(\varphi)f(\varphi) + v(\varphi)g(\varphi) = \boldsymbol{I}.$$

Then for any $\alpha \in V$, we have $\alpha = u(\varphi)f(\varphi)\alpha + v(\varphi)g(\varphi)\alpha$. Note that

$$f(\varphi)u(\varphi) \in \ker g(\varphi), g(\varphi)v(\varphi) \in \ker f(\varphi),$$

we have $V = V_1 + V_2$.

By the last question, we have $V_1 \cap V_2 = 0$. Therefore $V = V_1 \oplus V_2$. □

3. Space Decomposition

In this section, we will discuss the decomposition of vector spaces. The main idea is to find a basis that can be used to represent the vector space in a more concise way. To begin with, we will consider the invariant subspaces of a linear transformation.

Before we start, let's recall our tools from the polynomial section.

Theorem 3.1

$f = f_1, \dots, f_k$ is coprime factorization, then

$$\ker f(T) = \bigoplus_{j=1}^{k} \ker f_{j(T)}.$$

3.1. Invariant Subspaces

Invariance $T(W) \subset W$ induces the restriction of the linear transformation to the subspace

$$T|_W : W \mapsto W, v \mapsto Tv.$$

Invariant subspace direct-sum decomposes the vector space in the following way.

Note 3.1.1 Invariant Subspace Decomposition

$$V = W_1 \oplus \dots \oplus W_k,$$

and $T|_{W_j}$ has matrix representation $[T]_j$ with respect to a basis of W_j. Then

$$T = \operatorname{diag}\left([T]_1, \dots, [T]_k\right).$$

3.1.1. Commutation and Invariance

Theorem 3.1.1.1

Let T, S be linear transformations that commute, then

$$\ker T, \operatorname{range} T$$

are invariant under S.

We've got an important special case of this theorem where $S = f(T)$ is a polynomial of T. This is closely related to the decomposition by characteristic polynomials or sth else.

3.2. Primary Decomposition

3.2.1. Null Space Stops Growing

Definition 3.2.1.1

$T \in \operatorname{Hom}(V, V)$ (finite dimensional). Define

$$\ker T^{\infty} = \lim_{k \to \infty} \ker T^k = \bigcup_{k=1}^{\infty} \ker T^k,$$
$$\operatorname{im} T^{\infty} = \lim_{k \to \infty} \operatorname{im} T^k = \bigcap_{k=1}^{\infty} \operatorname{im} T^k.$$

Theorem 3.2.1.1 Null Space Stops Growing

$$\ker T \subset \ker T^2 \subset \dots$$
$$\operatorname{im} T \supset \operatorname{im} T^2 \supset \dots$$

$$\ker T^{\infty} = \ker T^k, \operatorname{im} T^{\infty} = \operatorname{im} T^k$$

for some k.

Theorem 3.2.1.2

The following statements are equivalent:

(1) $V_1 = \ker T^{\infty}, V_2 = \operatorname{im} T^{\infty}$.

(2) $V = V_1 \oplus V_2$, where V_1, V_2 are T-invariant subspaces s.t. $T|_{V_1}$ is nilpotent and $T|_{V_2}$ is invertible.

Proof: (1) If $v \in V_1 \cap V_2$, then there exists w s.t. $v = T^k w$,

Then $w \in \ker T^{2k} = \ker T^{\infty} = \ker T^k$. $v = T^k w = 0$.

□

Note 3.2.1.1

(2) shows a nice decomposition of the vector space to study the structure of the linear transformation, while (2)⇒(1) means the only way to depict the vector space is to consider the null space and the range of the linear transformation.

3.2.2. Primary Decomposition

Theorem 3.2.2.1

If W is invariant under T, then $W = \bigoplus \ker\left(p_j(T)\right) \cap W$.

Note 3.2.2.1 Projection onto Root Spaces

Bezout's Theorem gives us a way to project onto the root spaces of the characteristic polynomial.

$$\ker f(T) = \ker p_j(T) \oplus \ker\left(f/p_j\right)(T).$$

By Bezout's Theorem, we have $u(x)\frac{f}{p_j}(x) + v(x)p_j(x) = 1$. Then

$$u(T)\frac{f}{p_j}(T)(x) + 0 = v \in \ker p_j(T)$$

Thus the projection is $u(T)\frac{f}{p_j}(T)$, which is a polynomial of T.

3.3. Rational Canonical Form

3.3.1. Cayley-Hamilton Theorem

3.3.2. Cyclic Spaces

Definition 3.3.2.1 Krylov Space, or Cyclic Space

$$K^m(v) = \operatorname{span}\left(\{v, Tv, ..., T^{m-1}v\}\right).$$

$$K^\infty(v) = \bigcup_{m=1}^{\infty} K^m(v).$$

They are also called cyclic spaces, generated by a single vector v.

By finiteness, the union must stop as some point. If m_0 is the smallest integer s.t. $T^{m_0}v$ is a linear combination of $v, ..., T^{m_0-1}v$, then $K^\infty(v) = K^{m_0}(v)$, whereafter the sequence becomes stable. In fact, the vectors $v, ..., T^{m_0-1}v$ form a basis of $K^{m_0}(v)$.

It's easy to see that $K^m(v)$ is T-invariant. Moreover, under the induced basis, the matrix representation of T is a companion matrix, whose minimal polynomial is the same as the characteristic polynomial.

4. Diagonalization

4.1. Simultaneous Diagonalization

Theorem 4.1.1 Simultaneous Diagonalization

Let $T_1, \ldots, T_k$ be commuting linear transformations. Then there exists a basis s.t. T_j is diagonal for all j.

Proof: By induction on $\dim V$. Restricting T to smaller spaces, we might find their common invariant subspaces.

Take the eigen decomposition

$$\bigoplus \ker(\lambda I - T_j) = W_1 \oplus \ldots \oplus W_k,$$

of T_1. By commutativity, W_j **is T_r-invariant for all** r.

Then $T_1|W, \ldots, T_k|W$ are diagonalizable (induction hypothesis). The result follows. □

4.2. Diagonalization of Tensor Product

Example 4.2.1

If $A, B \in \mathbb{F}^{n \times n}$ are diagonalizable, then

$$T \in \mathrm{Hom}(\mathbb{F}^{n\times n}, \mathbb{F}^{n\times n}) : X \mapsto AXB$$

is diagonalizable.

Proof:

$$[T] = B^T \otimes A$$
$$(P \otimes Q)^{-1}[T](P \otimes Q) = (P^{-1}B^T Q) \otimes (P^{-1}AQ) = D_1 \otimes D_2 = D.$$

Alternative: $T = L_A R_B$ where L_A, R_B are left and right multiplication operators. The two multiplication operators are commutative, so it suffices to diagonalize them separately (because they are simultaneously diagonalizable, the P^{-1} and P in the middle cancel out). □

5. Commutativity

5.1. Commutating Linear Transformations

Theorem 5.1.1

Let T, S be commuting linear transformations. Then

$$\ker{(T - \lambda I)^\infty} \text{ is invariant under } S.$$

Choosing a basis s.t. T is Jordan form $\mathrm{diag}(J_1, \ldots, J_n)$, then the matrix representation of S is block diagonal with respect to the Jordan blocks, and $S_i J_i = J_i S_i$.

6. Miscellaneous

6.1. Minimal Polynomial

Question 6.1.1

If the minimal polynomial of $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ is $m(\lambda) = \prod(\lambda - \lambda_i)^{k_i}$. Prove that the minimal polynomial of $\boldsymbol{B} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{I} \\ 0 & \boldsymbol{A} \end{pmatrix}$ is $\prod(\lambda - \lambda_i)^{k_i+1}$.

Proof: Note that $\boldsymbol{B}^t = \begin{pmatrix} \boldsymbol{A}^t & t\boldsymbol{A}^{t-1} \\ 0 & \boldsymbol{A}^t \end{pmatrix}$, we have

$$f(\boldsymbol{B}) = \begin{pmatrix} f(\boldsymbol{A}) & f'(\boldsymbol{A}) \\ 0 & f(\boldsymbol{A}) \end{pmatrix}, \text{where } f \text{ is a polynomial.}$$

Denote $g(\lambda)$ as the minimal polynomial of $\boldsymbol{B}$. Then $g(\boldsymbol{B}) = 0$ implies that $g(\boldsymbol{A}) = 0, g'(\boldsymbol{A}) = 0$. Then $m \mid g, m \mid g'$.

By $m \mid g$, we have $g = mp$, and $m \mid g' = mp' + m'p$, then $m \mid m'p$. Consider every factor $(\lambda - \lambda_i)^{k_i}$ in m, we know that

$$(\lambda - \lambda_i)^{k_i - 1} \parallel m'.$$

Therefore, the minimal $p = \prod(\lambda - \lambda_i)$, which implies that the minimal polynomial $g = \prod(\lambda - \lambda_i)^{k_i+1}$. □