Linear Algebra and Matrices

C.T.A

This chapter deals with the study of linear algebra and matrices as they are used in this course.

1 Review of Matrix Algebra

We start this chapter by introducing matrices and their algebra. Notationally, matrices are a rectangular array of elements denoted by capital letters A, M, Γ. Those elements are referred to as scalars and denoted by lowercase letters, $a, b, \alpha, etc.$. Note that the scalars are not necessarily constants: they maybe real or complex numbers, polynomials, or general functions.

Example 1 Consider the matrix

$$A = \begin{pmatrix} -0.02567 & -36.617 & -18.897 & -32.09 & 3.2509 & -0.76257 \\ 0.257 \times 10^{-5} & -1.8997 & 0.98312 & -7.256 \times 10^{-4} & -0.1708 & -4.965 \times 10^{-3} \\ 0.012338 & 11.72 & -2.6316 & 8.758 \times 10^{-4} & -31.604 & 22.396 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -30 \end{pmatrix}$$

This is a 5x6 matrix with real entries. On the other hand,

$$P(s) = \begin{bmatrix} s+1 & s+3 \\ s^2+3s+2 & s^2+5s+4 \end{bmatrix}$$

is a 2×2 matrix of polynomials in the variable s.

$\triangle$

A matrix which has m rows and n columns, is said to be $m \times n$. A matrix may be denoted by $A = [a_{ij}]$ where $i = 1, \cdots m$ and $j = 1, \cdots n$.

Example 2 The matrix A in Example 1 has 5 rows and 6 columns. the element $a_{11} = -0.02567$ while $a_{43} = 1$.

$\triangle$

A scalar is a 1×1 matrix. If $m = n$ the matrix is said to be square. If $m = 1$ the matrix is a row matrix (or vector). If $n = 1$ the matrix is an m column matrix (or vector). If A is square then we define the $Trace$ of A by the sum of its diagonal elements, i.e.

$$Trace(A) = \sum_{i=1}^{n} a_{ii} \tag{1}$$

Example 3 Consider $P(s)$ of Example 1, then $Trace(P(s)) = s^2 + 6s + 5$.

△

1.1 Matrix Algebra

Two matrices A and B are equal, written $A = B$, if and only if A has the same number of rows and columns as B and if $a_{ij} = b_{ij}$ for all i, j. Two matrices A and B that have the same numbers of rows and columns may be added or subtracted element by element, i.e.

$$C = [c_{ij}] = A \pm B \Longleftrightarrow c_{ij} = a_{ij} \pm b_{ij} \ \forall i, j \tag{2}$$

the multiplication of 2 matrices A and B to obtain $C = A.B$ may be performed if and only if A has the same number of columns as B has rows. In fact,

$$C = AB \Longleftrightarrow c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \tag{3}$$

the matrix C is $m \times q$ if A is $m \times n$ and B is $n \times q$. Note that BA may not even be defined and it certainly is not equal to AB even when both products are defined. The matrix of all zeros is the Null matrix, and the square matrix A with $a_{ii} = 1$ and $a_{ij} = 0$ for $i \neq j$ is the identity matrix. The identity matrix is denoted by I. Note that $AI = IA = A$ assuming A is $n \times n$ as is I.

1.2 Properties of the Laws of Matrix Algebra

The following properties of matrix Algebra are easily verified

1. $A \pm B = B \pm A$
2. $A + (B + C) = (A + B) + C$
3. $\alpha(A + B) = \alpha A + \alpha B$ for all scalar α.
4. $\alpha A = A\alpha$
5. $A(BC) = (AB)C$
6. $A(B + C) = AB + AC$
7. $(B + C)A = BA + CA$

1.3 Transpose, Conjugate, and Associate Matrices

Let $A = [a_{ij}]$ be an $m \times n$ matrix. The transpose of A is denoted by $A^T = [a_{ji}]$ and is the nxm matrix obtained by interchanging columns and rows. The matrix A is *symmetric* if $A = A^T$, *skew-symmetric* if $A = -A^T$. Also, it can be seen that $(AB)^T = B^T A^T$. In the case where A contains complex elements, we let $\bar{A}$ be the *conjugate* of A whose elements are the conjugates of those of A. Matrices satisfying $A = \bar{A}^T$ are *Hermitian* and those satisfying $A = -\bar{A}^T$ are *skew-Hermitian.*

1.4 Determinants, Minors, and Cofactors

In this section we will consider square matrices only. The determinant of a square matrix A denoted by $det(A)$ or $\mid A \mid$ is a scalar-valued function of A and is given as follows for $n = 1, 2, 3$.

1. $n = 1$, then $\mid A \mid = a_{11}$
2. $n = 2$, then $\mid A \mid = a_{11}a_{22} - a_{12}a_{21}$
3. $n = 3$, then $\mid A \mid = a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{12}(a_{23}a_{31} - a_{21}a_{33}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$

Standard methods apply for calculating determinants. The determinant of a square matrix is 0, that of an identity matrix is 1, and that of a triangular or diagonal matrix is the product of all diagonal elements. Each element a_{ij} in an $n \times n$ square matrix A has associated with it a minor M_{ij} obtained as the determinant of the $(n-1) \times (n-1)$ matrix resulting form deleting the ith row and the jth column. From these minors, we can obtain the cofactors given by $C_{ij} = (-1)^{i+j} M_{ij}$.

Example 4 Given

$$A = \begin{pmatrix} 2 & 4 & 1 \\ 3 & 0 & 2 \\ 2 & 0 & 3 \end{pmatrix}$$

Then, $M_{12} = 5$, $C_{12} = -5$, $M_{32} = 1 = -C_{32}$.

$\triangle$

Some useful properties of determinants are listed below:

1. Let A and B be $n \times n$ matrices, then $\mid AB \mid = \mid A \mid . \mid B \mid$.
2. $\mid A \mid = \mid A^T \mid$
3. If A contains a row or a column of zeros, then $\mid A \mid = 0$.
4. If any row (or column) of A is a linear combination of other rows (or columns), then $\mid A \mid = 0$.
5. If we interchange any 2 rows (or columns) of a matrix A, the determinant of the resulting matrix is $- \mid A \mid$.
6. If we multiply a row (or a column) of a matrix A by a scalar α, the determinant of the resulting matrix is $\alpha \mid A \mid$.
7. Any multiple of a row (or a column) may be added to any other row (or column) of a matrix A without changing the value of the determinant.

1.5 Rank, Trace, and Inverse

The rank of an $m \times n$ matrix A denoted by $r_A = rank(A)$ is the size of the largest nonzero determinant that can be formed from A. Note that $r_A \leq min\{m, n\}$. If A is square and if $r_A = n$, then A is nonsingular. If r_A is the rank of A and r_B is the rank of B, and if $C = AB$, then $0 \leq r_C \leq min\{r_A, r_B\}$. The trace of a square matrix A was defined earlier as $Tr(A) = \sum_{i=1}^{n} a_{ii}$.

Example 5 xx

$\triangle$

If A and B are conformable square matrices, $Tr(A+B) = Tr(A)+Tr(B)$, and $Tr(AB) = Tr(BA)$. Also, $Tr(A) = Tr(A^T)$.

Example 6 Show that $Tr(ABC) = Tr(B^T A^T C^T)$. First, write $Tr(ABC) = Tr(CAB)$, then $Tr([CAB]^T) = Tr(CAB)$, thus proven.

$\triangle$

Note that $rank(A + B) \neq rank(A) + rank(B)$ and that $Tr(AB) \neq Tr(A)Tr(B)$.

Example 7 xx

$\triangle$

Next, we define the inverse of a square, nonsingular matrix A as the square matrix B of the same dimensions such that

$$AB = BA = I$$

the inverse is denoted by A^{-1} and may be found by

$$A^{-1} = \frac{C^T}{\mid A \mid}$$

where C is the matrix of cofactors C_{ij} of A. Note that for the inverse to exist, $\mid A \mid$ must be nonzero, which is equivalent to saying that A is nonsingular. We also write $C^T = Adjoint(A)$.

Example 8 Show the following property

$$(AB)^{-1} = B^{-1}A^{-1}$$

$\triangle$

assuming of course that A and B are compatible and both invertible. There are other important relations involving inverses of matrices, one of which is the *Matrix Inversion Lemma*

$$(A_1 - A_2A_4^{-1}A_3)^{-1} = A_1^{-1} + A^{-1}A_2(A_4 - A_3A_1^{-1}A_2)^{-1}A_3A_1^{-1} \tag{4}$$

Example 9 Prove MIL

$\triangle$

1.6 Elementary Operations and Matrices

The basic operations called *elementary operations* are as follows

1. Interchange any 2 rows (or columns)
2. Multiply any row (or column) by a scalar α
3. Multiply any row (or column) by a scalar α and add the resulting row (or column) to any other row (or column)

Note that each of these elementary operations may be represented by a postmultiplication for column operations (or premultiplication for row operations) by elementary matrices which are nonsingular.

Example 10 Let

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & -1 \end{bmatrix}$$

Then, let use the following row operations:

1. Exchange rows 1 and 3
2. Multiply row 1 by -2 and add it to row 3
3. Multiply row 2 by 1/2 and add it to row 3
4. Multiply row 2 by 1/2
5. Multiply row 3 by 2/5
6. Multiply column 1 by -1 and add it to column 2
7. Multiply column 1 by 1 and add it to column 3
8. Multiply column 2 by -1/2 and add it to column 3

The end result is

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$\triangle$

Example 11 Given the following polynomial matrix

$$P(s) = \begin{bmatrix} s^2 & 0 \\ 0 & s^2 \\ 1 & s+1 \end{bmatrix}$$

then the corresponding row operations are performed:

1. interchange rows 3 and 1.
2. multiply row 1 by $-s^2$ and add it to row 3
3. multiply row 2 by $s+1$ and add it to row 3.

This corresponds to the following multiplications by the matrices given:

1.

$$P_1(s) = \begin{bmatrix} 1 & s+1 \\ 0 & s^2 \\ s^2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} P(s)$$

2.

$$P_2(s) = \begin{bmatrix} 1 & s+1 \\ 0 & s^2 \\ 0 & -s^2(s+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -s^2 & 0 & 0 \end{bmatrix} P_1(s)$$

3.

$$P_3(s) = \begin{bmatrix} 1 & s+1 \\ 0 & s^2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & s+1 & 0 \end{bmatrix} P(s)$$

$\triangle$

In general, for any matrix of rank r we can reduce it via column and row operations to one of the following normal forms

$$I_r,\ [I_r \mid 0],\ [I_r \mid 0]^T, or \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \tag{5}$$

Next, we discuss vector spaces.

2 Vectors and vector Spaces

In most of our applications, we need to deal with linear real and complex vector spaces which are defined subsequently.

Definition 1 *A* real linear vector space *(resp.* complex linear vector space *is a set V, equipped with 2 binary operations: the addition (+) and the scalar multiplication (.) such that*

1. $x + y = y + x, \ \forall x, y \in V$
2. $x + (y + z) = (x + y) + z, \ \forall x, y, z \in V$
3. *There is an element 0_V in V such that $x + 0_V = 0_V + x = x, \ \forall x \in V$*
4. *For each $x \in V$, there exists an element $-x \in V$ such that $x + (-x) = (-x) + x = 0_V$*
5. *For all scalars $r_1, r_2 \in R$ (resp. $c_1, c_2 \in C$), and each $x \in V$, we have $r_1.(r_2.x) = (r_1 r_2).x$ (resp. $c_1.(c_2.x) = (c_1 c_2).x$*
6. *For each $r \in R$ (resp. $c \in C$), and each $x_1, x_2 \in V$, $r.(x_1 + x_2) = r.x_1 + r.x_2$ (resp. $c.(x_1 + x_2) = c.x_1 + c.x_2$)*
7. *For all scalars $r_1, r_2 \in R$ (resp. $c_1, c_2 \in C$), and each $x \in V$, we have $(r_1 + r_2).x = r_1.x + r_2.x$ (resp. $(c_1 + c_2).x = c_1.x + c_2.x$)*
8. *For each $x \in V$, we have $1.x = x$ where 1 is the unity in R (resp. in C.*

Example 12 *The following are linear vector spaces with the associated scalar fields: R^n with R, C^n with C.*

△

Definition 2 *A subset M of a vector space V is a subspace if it is a linear vector space in its own right. One necessary condition for M to be a subspace is that it contains the zero vector.*

Example 13 xxx

△

2.1 Linear Independence

Consider a finite number of vectors $\{x_i\} = \{x_1, x_2, \cdots, x_n\}$ belonging to a vector space X. The set $\{x_i\}$ is said to be linearly independent if $a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0$ implies that all $a_i = 0$. Otherwise, $\{x_i\}$ is said to be linearly dependent.

Example 14 The set $\{e_i\}$ $i = 1, \cdots, n$ is linearly independent in R^n. On the other hand, the 2 vectors, $x_1^T = [1 \ -2]$; $x_2^T = [2 \ -4]$ are linearly dependent in R^2.

$\triangle$

Note that any set of vectors containing the zero vector is linearly dependent. Also, if $\{x_i\}$ $i = 1, \cdots, n$ is linearly dependent, adding a new vector x_{n+1} will not make the new set linearly independent. Finally, if the set $\{x_i\}$ is linearly dependent, then, at least one of the vectors in the set may be written as a linear combination of the others.

How do we test for linear dependence/independence? Given a set of n vectors $\{x_i\}$ each having m components. Form the $m \times n$ matrix A which has the vectors x_i as its columns. Note that if $n > m$ then the set of vectors has to be linearly independent. Therefore, let us consider the case where $n \leq m$. Form the $n \times n$ matrix $G = A^T A$ and check if it is nonsingular, i.e. if $\mid G \mid \neq 0$, then $\{x_i\}$ is linearly independent, otherwise $\{x_i\}$ is linearly dependent.

Geometrically, linear independence may be explained in R^n. Consider the plane R^2 and suppose we have 2 vectors x_1 and x_2. If the 2 vectors are linearly dependent, then $x_2 = ax_1$ or both vectors lie along the same direction. If they were linearly independent, then they form the 2 sides of a parallelogram. Therefore, in the case of linear dependency, the parallelogram degenerate to a single line. We can equip a vector space with many functions. One of which is the inner product which takes two vectors in V to a scalar either in R or in C, the other one is the norm of a vector which takes a vector in V to a positive value in R.

2.2 Basis Vectors

Given a vector space X, we say that $\{u_i, i = 1, \cdots m\}$ span X if every vector $x \in X$ may be written as a linear combination of u_i, i.e.

$$x = \sum_{i}^{m} a_i u_i \tag{6}$$

for some scalars a_i. The scalars a_i are the components of x in the u_i directions. Note that nothing is said about u_i beyond the fact that there is enough of them to manufacture every vector in X.

Example 15 3 vectors in R^2

$\triangle$

A set of vectors $\{v_i, i = 1, \cdots n\}$ in X is said to be a basis of X if it spans X and it is linearly independent. This is basically the minimum number of vectors needed to span X. The dimension of X is the number of vectors in any basis of X. Thus R^n has dimension n.

Example 16 Let X be the vector space of all vectors $x = [x_1 \ \cdots \ x_n]^T$ such that all components are equal. Then X is spanned by the vector of all 1's. Therefore dim $(X) = 1$. On the other hand, if X is the vector space of all polynomials of degree $n-1$ or less, a basis is $\{1, t, t^2, \cdots, t^{n-1}\}$ which makes dim $(X) = n$.

$\triangle$

The following section discusses the inner products of vectors which is then followed by a section on norms.

2.3 Inner Products

An inner product is an operation between two vectors of a vector space which will allow us to define geometric concepts such as orthogonality and Fourier series, etc. The following defines an inner product.

Definition 3 *An inner product defined over a vector space V is a function $< .,. >$ defined from V to F where F is either R or C such that $\forall x, y, z, \in V$*

1. $< x, y > = < \bar{y, x} >$ *where the* $< \bar{.,.} >$ *denotes the complex conjugate.*
2. $< x, y + z > = < x, y > + < x, z >$
3. $< x, \alpha y > = \alpha < x, y >, \ \forall \alpha \in F$
4. $< x, x > \geq 0$ *where the* 0 *occurs only for* $x = 0_V$

Example 17 The usual dot product in R^n is an inner product.

$\triangle$

2.4 Norms

A *norm* is a generalization of the ideas of distance and length. As stability theory is usually concerned with the size of some vectors and matrices, we give here a brief description of some norms that will be used in this book. We will consider first the norms of vectors defined on a vector space X with the associated scalar field of real numbers R.

2.4.1 Vector Norms

We start our discussion of norms by reviewing the most familiar normed spaces, that is the spaces of vectors with constant entries. In the following, $\mid a \mid$ denotes the absolute value of a for a real a or the magnitude of a if a is complex.

Definition 4 *A norm $\| . \|$ of a vector x is a real-valued function defined on the vector space X such that*

1. $\| x \| > 0$ *for all* $x \in X$ *with* $\| x \| = 0$ *if and only if* $x = 0$.
2. $\| \alpha x \| = \mid \alpha \mid \| x \|$ *for all* $x \in X$ *and any scalar* α.
3. $\| x + y \| \leq \| x \| + \| y \|$ *for all* $x, y \in X$.

Example 18 The following are Common norms in $X = R^n$ where R^n is the set of nx1 vectors with real components.

1. 1-norm: $\| x \|_1 = \sum_{i=1}^{n} \mid x_i \mid$.

2. 2-norm: $\parallel x \parallel_2 = \sqrt{\sum_{i=1}^{n} \mid x_i^2 \mid}$, also known as the Euclidean norm

3. p-norm: $\parallel x \parallel_p = (\sum_{i=1}^{n} \mid x_i^p \mid)^{\frac{1}{p}}$.

4. ∞ - norm : $\parallel x \parallel_\infty = \max \mid \mathrm{x_i} \mid \ \forall \mathrm{i} = 1, \cdots, \mathrm{n}$.

$\triangle$

Example 19 Consider the vector

$$x = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$$

Then, $\parallel x \parallel_1 = 5$, $\parallel x \parallel_2 = 2$ and $\parallel x \parallel_\infty = 2$.

$\triangle$

We now present an important property of norms of vectors in R^n which will be useful in the sequel.

Lemma 1 *Let $\parallel x \parallel_a$ and $\parallel x \parallel_b$ be any two norms of a vector $x \in R^n$. Then there exists finite positive constants k_1 and k_2 such that*

$$k_1 \parallel x \parallel_a \leq \parallel x \parallel_b \leq k_2 \parallel x \parallel_a \ \ \forall x \in R^n$$

The two norms in the lemma are said to be equivalent and this particular property will hold for any two norms on R^n.

Example 20 *Note the following*

1. It can be shown that for $x \in R^n$

$$\begin{aligned} ||x||_1 &\leq \sqrt{n}||x||_2 \\ ||x||_\infty \leq ||x||_1 &\leq n||x||_\infty \\ ||x||_2 &\leq \sqrt{n}||x||_\infty \end{aligned}$$

2. Consider again the vector of Example 1.4.2. Then we can check that

$$\begin{aligned} ||x||_1 &\leq \sqrt{3}||x||_2 \\ ||x||_\infty \leq ||x||_1 &\leq 3||x||_\infty \\ ||x||_2 &\leq \sqrt{3}||x||_\infty \end{aligned}$$

$\triangle$

Note that a norm may be defined independently from an inner product. Also, we can define the generalized angle between 2 vectors in R^n using

$$x.y =< x, y >=\parallel x \parallel . \parallel y \parallel \cos\theta \tag{7}$$

where θ is the angle between x and y. Using the inner product, we can define the orthogonality of 2 vectors by

$$x.y =< x, y >= 0 \tag{8}$$

which of course means that $\theta = (2i + 1)\pi/2$.

3 Linear Equations

In many control problems, we have to deal with a set of simultaneous linear algebraic equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2 \\ \vdots &= \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= y_m \end{aligned} \tag{9}$$

or in matrix notation

$$Ax = y \tag{10}$$

for an $m \times n$ matrix A, an $n \times 1$ vector x and an m vector y. The problem is to find the solution vector x, given A and y. Three cases might take place:

1. No solution exist
2. A unique solution exists
3. An infinite number of solutions exist

In order to understand what is going on, we should look into the range space and the null space of the matrix A, which are defined as follows: The range of A written as $\mathcal{R}(A) = \{y : y = \sum_{i=1}^{n} x_i a_i\}$ or the set of vectors $y \in R^m$ made up of all possible linear combinations of the columns of A. On the other hand, the null space of A denoted $\mathcal{N}(A) = \{x : Ax = 0\}$ or the set of vectors $x \in R^n$ such that $Ax = 0_m$.

Example 21 Show that $\mathcal{R}(A)$ is a subspace of R^m and $\mathcal{N}(A)$ is a subspace of R^n.

$\triangle$

Now, looking back at the linear equation $Ax = y$, it is obvious that y should be in $\mathcal{R}(A)$ for a solution x to exist, in other words, if we form

$$W = [A \mid y] \tag{11}$$

then $rank(W) = rank(A)$ is a necessary condition for the existence of at least one solution. Now, if $z \in \mathcal{N}(A)$, and if x is any solution to $Ax = y$, then $x + z$ is also a solution. Therefore, for a unique solution to exist, we need that $\mathcal{N}(A) = 0$. That will require that the columns of A form a basis of $\mathcal{R}(A)$, i.e. that there will be n of them and that they will be linearly independent, and of dimension n. Then, the A matrix is invertible and $x = A^{-1}y$ is the unique solution.

4 Eigenvalues and Eigenvectors

Let A be an $n \times m$ matrix, and denote the corresponding identity matrix by I. Then, let x_i be a nonzero vector in R^n and λ_i be scalar such that

$$Ax_i = \lambda x_i \tag{12}$$

Then, λ_i is an eigenvalue of A and x_i is the corresponding eigenvector. There will be n eigenvalues of A (some of which redundant). In order to find the eigenvalues of A we rewrite the previous equation

$$(A - \lambda_i I)x_i = 0 \tag{13}$$

Noting that x_i can not be the zero vector, and recalling the conditions on the existence of solutions of linear equations, we see that we have to require

$$det(A - \lambda_i I) = \mid (A - \lambda_i I) \mid = 0 \tag{14}$$

We thus obtain an nth degree polynomial, which when set to zero, gives the characteristic equation

$$\begin{aligned} \mid (A - \lambda I) \mid &= \Delta(\lambda) = 0 \\ &= (-\lambda)^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0 \\ &= (-1)^n(\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_p)^{m_p} \end{aligned} \tag{15}$$

Therefore, λ_i is an eigenvalue of A of algebraic multiplicity m_i. One can then shows,

$$\begin{aligned} Tr(A) &= \sum_{i=1}^{n} = (-1)^{n+1}c_{n-1} \\ \mid A \mid &= \prod_{i=1}^{n} \lambda_i = c_0 \end{aligned} \tag{16}$$

Also, if λ_i is an eigenvalue, then so is $\bar{\lambda_i}$.

How do we determine eigenvectors? We distinguish 2 cases:

1. All eigenvalues are distinct: In this case, we first find $Adj(A-\lambda I)$ with λ as a parameter. Then, successively substituting each eigenvalue λ_i and selecting any nonzero column gives all eigenvectors.

   Example 22 Let

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -18 & -27 & -10 \end{bmatrix}$$

   then, solve for

$$\mid A-\lambda I \mid = -\lambda^3 - 10\lambda^2 - 27\lambda - 18 = 0$$

   which has 3 solutions, $\lambda_1 = -1,\ \lambda_2 = -3,\ \lambda_3 = -6$. Now, let us solve for the eigenvectors using the suggested method

$$Adj(A-\lambda I) = \begin{bmatrix} \lambda^2 + 10\lambda + 27 & \lambda + 10 & 1 \\ -18 & \lambda^2 + 10\lambda & \lambda \\ -18\lambda & -27\lambda - 18 & \lambda^2 \end{bmatrix}$$

   so that for $\lambda_1 = -1$ we can see that column 1 is

$$x_1 = [18\ \ -18\ 18]^T$$

   Similarly, $x_2 = [7\ \ -21\ 63]^T$ and $x_3 = [1\ \ -6\ 36]^T$. There is another method of obtaining the eigenvectors from the definition by actually solving $Ax_i = \lambda_i x_i$

   △

   Note that once all eigenvectors are obtained, we can arrange them in an $n \times n$ modal matrix $M = [x_1\ x_2 \cdots x_n]$. Note that the eigenvectors are not unique since if x_i is an eigenvector, then so is any αx_i for any scalar α.

2. Some eigenvalues are repeated: In this case, a full set of independent eigenvectors may or may not exist. Suppose λ_i is an eigenvalue with an algebraic multiplicity m_i. Then, the dimension of the Null space of $A - \lambda_i I$ which is also the number of linearly independent eigenvectors associated with λ_i is the geometric multiplicity q_i of the eigenvalue λ_i. We can distinguish 3 cases

   (a) Fully degenerate case $q_i = m_i$: In this case there will be q_i independent solutions to $(A - \lambda_i I)x_i$ rather than just one.

   Example 23 Given the matrix

$$A = \begin{pmatrix} 10/3 & 1 & -1 & -1/3 \\ 0 & 4 & 0 & 0 \\ -2/3 & 1 & 3 & -1/3 \\ -2/3 & 1 & -1 & 11/3 \end{pmatrix}$$

Then, its characteristic equation is

$$\Delta(\lambda) = \lambda^4 - 14\lambda^3 + 72\lambda^2 - 160\lambda + 128 = 0$$

There are 4 roots, $\lambda_1 = 2$, $\lambda_2 = \lambda_3 = \lambda_4 = 4$. We can find the eigenvector associated with λ_1 as $x_1 = [-1\ 0\ -1\ -1]^T$. Then, we find

$$(4I - A)x = \begin{pmatrix} 2/3 & -1 & 1 & 1/3 \\ 0 & 0 & 0 & 0 \\ 2/3 & -1 & 1 & 1/3 \\ 2/3 & -1 & 1 & 1/3 \end{pmatrix} x$$

There are an infinite number of solutions, 3 of which are $x_2 = [1\ 0\ 0\ -2]^T$; $x_3 = [0\ 1\ 1\ 0]^T$; $x_4 = [0\ 1\ 0\ 3]^T$

△

Note that now if $x_1, \cdots x_m$ are eigenvectors for eigenvalue λ_i, then so is any $y = \sum_{i=1}^m \alpha_i x_i$.

(b) Simple degenerate case $q_i = 1$: Here we can find the first eigenvector in the usual means, then we have $m_i - 1$ *generalized eigenvectors*. These are obtained as

$$\begin{aligned} Ax_1 &= \lambda_i x_1 \\ (A - \lambda_i)x_2 &= x_1 \\ \vdots &= \vdots \\ (A - \lambda_i)x_{m_i} &= x_{m_i - 1} \end{aligned} \tag{17}$$

Note that all x_i found this way are linearly independent.

Example 24

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 8 & -20 & -18 & -7 \end{pmatrix}$$

its characteristic polynomial is

$$\Delta(\lambda) = \lambda^4 + 7\lambda^3 + 18\lambda^2 + 20\lambda + 8 = 0$$

then, $\lambda_1 = -1$, $\lambda_2 = \lambda_3 = \lambda_4 = -2$. The eigenvector of λ_1 is easily found to be $x_1 = [-1\ 1\ -1\ 1]^T$. On the other hand, one eigenvector 0f -2 is found to be $x_2 = [0.125\ -0.25\ 0.5 - 1]^T$, then $x_2 = (A + 2I)x_3$ leading to $x_3 = [0.1875\ -0.25\ 0.25\ 0]^T$ also, $x_3 = (A + 2I)x_4$ so that $x_4 = [0.1875\ -0.1875\ 0.125\ 0]^T$

△

(c) General case $1 \leq q_i \leq m_i$: Here, a general top-down method should be used. Here we solve the problem by writing

$$\begin{aligned}
(A - \lambda_i I)x_1 &= 0 \\
(A - \lambda_i I)x_2 &= x_1 \Rightarrow (A - \lambda_i I)^2 x_2 = (A - \lambda_i I)x_1 = 0 \\
(A - \lambda_i I)x_3 &= x_2 \Rightarrow (A - \lambda_i I)^2 x_3 = (A - \lambda_i I)x_2 = x_1 \neq 0 \\
&\quad (A - \lambda_i I)^3 x_3 = (A - \lambda_i I)x_1 = 0
\end{aligned} \tag{18}$$

This approach continues until we reach the index k_i of λ_i. This index is found as the smallest integer such that

$$rank(A - \lambda_i I)^k = n - m_i \tag{19}$$

the index indicates the length of the longest chain of eigenvectors and generalized eigenvectors associated with λ_i.

Example 25

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

its characteristic polynomial is

$$\Delta(\lambda) = \lambda^4 = 0$$

so $\lambda_1 = 0$ with $m_1 = 4$. Next, form $A - \lambda_1 I = A$, and determine that $rank(A - \lambda_1 I) = 2$. There are then 2 eigenvectors and 2 generalized eigenvectors. The question is whether we have one eigenvector-generalized eigenvector chain of length 3, or 2 chains of length 2. To check that, note that $n - m = 0$, and that $rank(A - \lambda_1 I)^2 = 0$, therefore, the index is $k_1 = 2$. this then guarantees that one chain has length 2, making the length of the other chain 2. First, consider $(A - \lambda_1 I)^2 x = 0$ Any vector satisfies this equation but only 4 vectors are linearly independent. Let us choose $x_1 = [1\ 0\ 0\ 0]^T$. Is this an eigenvector or a generalized eigenvector? It is an eigenvector since $(A - \lambda_1 I)x_1 = 0$. Similarly, $x_2 = [0\ 1\ 0\ 0]^T$ is an eigenvector. On the other hand, $x_3 = [0\ 0\ 1\ 0]^T$ is a generalized eigenvector since $(A - \lambda_i I)x_3 = x_1 \neq 0$. Similarly, $x_4 = [0\ 0\ 0\ 1]^T$ is a generalized eigenvector associated with x_2.

$\triangle$

In summary each $n \times n$ matrix has n eigenvalues and n linearly independent vectors, either eigenvectors or generalized eigenvectors.

4.1 Jordan Forms

Based on the analysis above, we can produce the Jordan form of any $n \times n$ matrix. In fact, if we have n different eigenvalues, we can find all eigenvectors x_i, such that $Ax_i = \lambda_i x_i$, then let $M = [x_1 \; x_2 \; \cdots \; x_n]$, and $\Lambda = diag(\lambda_i)$. This leads to $AM = M\Lambda$. We know that M^{-1} exists since all eigenvectors are independent and therefore $\Lambda = M^{-1}AM$. The Jordan form of A is then Λ. If the eigenvectors are orthonormal, $M^{-1} = M^T$. In the case that $q_i = m_i$, we proceed the same way to obtain $J = \Lambda$. On the other hand, if $q_i = 1$, then using the eigenvectors and generalized eigenvectors as columns of M will lead to $J = diag[J_1 \; J_2 \; \cdots \; J_p]$ where each J_i is $m_i \times m_i$ and has the following form

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & \lambda_i \end{pmatrix}$$

Finally, the general case will have Jordan blocks each of size k_i as shown in the examples below.

Example 26 Let

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -18 & -27 & -10 \end{bmatrix}$$

the eigenvalues are $\lambda_1 = -1, \; \lambda_2 = -3, \; \lambda_3 = -6$. The eigenvectors are

$$x_1 = [1 \; -1 \; 1]^T$$

Similarly, $x_2 = [1 \; -3 \; 9]^T$ and $x_3 = [1 \; -6 \; 36]^T$. Then,

$$M = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -3 & -6 \\ 1 & 9 & 36 \end{bmatrix}$$

then

$$M^{-1} = \begin{bmatrix} 1.8 & 0.9 & 0.1 \\ -1 & -1.167 & -0.167 \\ 0.2 & 0.2667 & 0.067 \end{bmatrix}$$

then $\Lambda = M^{-1}AM$.

$\triangle$

Example 27 Given the matrix

$$A = \begin{pmatrix} 10/3 & 1 & -1 & -1/3 \\ 0 & 4 & 0 & 0 \\ -2/3 & 1 & 3 & -1/3 \\ -2/3 & 1 & -1 & 11/3 \end{pmatrix}$$

Then, its characteristic equation is

$$\Delta(\lambda) = \lambda^4 - 14\lambda^3 + 72\lambda^2 - 160\lambda + 128 = 0$$

There are 4 roots, $\lambda_1 = 2$, $\lambda_2 = \lambda_3 = \lambda_4 = 4$. We can find the eigenvector associated with λ_1 as $x_1 = [-1\ 0\ -1\ -1]^T$. Then, we find $x_2 = [1\ 0\ 0\ -2]^T$; $x_3[0\ 1\ 1\ 0]^T$; $x_2 = [0\ 1\ 0\ 3]^T$ Therefore,

$$M = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ -1 & 0 & 1 & 0 \\ -1 & -2 & 0 & 3 \end{pmatrix}$$

then

$$M^{-1} = \begin{pmatrix} -1/3 & 1/2 & -1/2 & -1/6 \\ 1/6 & 1/2 & -1/2 & -1/6 \\ -1/3 & 1/2 & 1/2 & -1/6 \\ 1/3 & 1/2 & -1/2 & 1/6 \end{pmatrix}$$

then

$$J = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

$\triangle$

Example 28 Now consider

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 8 & -20 & -18 & -7 \end{pmatrix}$$

its characteristic polynomial is

$$\Delta(\lambda) = \lambda^4 + 7\lambda^3 + 18\lambda^2 + 20\lambda + 8 = 0$$

then, $\lambda_1 = -1$, $\lambda_2 = \lambda_3 = \lambda_4 = -2$. Find eigenvectors as before and form

$$M = \begin{pmatrix} -1 & 1/8 & 0.1875 & 0.1875 \\ 1 & -1/4 & -1/4 & -0.1875 \\ -1 & 1/2 & 1/4 & 1/8 \\ 1 & -1 & 0 & 0 \end{pmatrix}$$

Then,

$$J = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

$\triangle$

Example 29

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

its characteristic polynomial is

$$\Delta(\lambda) = \lambda^4 = 0$$

so $\lambda_1 = 0$ with $m_1 = 4$. Then,

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

therefore making

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

$\triangle$

5 The Cayley-Hamilton Theorem

Note that $A^i = A \cdots A$ i times. Then, we have the Cayley-Hamilton theorem

Theorem 1 Let A be an $n \times n$ matrix with a characteristic equation

$$\begin{aligned} |\,(A - \lambda I) &= \Delta(\lambda) \\ &= (-\lambda)^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0 \\ &= 0 \end{aligned} \tag{20}$$

Then,

$$\begin{aligned} \Delta(A) &= (-1)^n A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I \\ &= 0 \end{aligned} \tag{21}$$

■

Example 30 Use the Cayley-Hamilton theorem to find the inverse of a matrix A

$$0 = (-1)^n A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I$$

then,

$$0 = (-1)^n A^{n-1} + c_{n-1}A^{n-2} + \cdots + c_1 I + c_0 A^{-1}$$

$$A^{-1} = \frac{1}{c_0}[(-1)^n A^n + c_{n-1}A^{n-1} + \cdots + c_1 I]$$

△

6 More on Matrices

The expression $< y, Ax >= y^T Ax$ is a bilinear form. When $y = x$ we have a quadratic form $x^T Ax$. Every matrix A can be written as the sum of a symmetric and skew symmetric matrices if it is real, and of a Hermitian and skew-Hermitian if it is complex. Then note that $< x, Ax >=< x, A_s x >$ if A is real, and $< x, Ax >=< x, A_H x >$ if A is complex.

6.1 Positive-Definite Matrices

Given an $n \times n$ matrix Q, then

1. Q is positive-definite, if and only if $< x, Qx > > 0$ for all $x \neq 0$.
2. Q is positive semidefinite if $< x, Qx >\geq 0$ for all x.

3. Q is indefinite if $< x, Qx > > 0$ for some x and $< x, Qx >< 0$ for other x.

Q is negative definite (or semidefinite) if $-Q$ is positive definite (or semidefinite). If Q is symmetric then all its eigenvalues are real.

Example 31 xxx

$\triangle$

7 MATLAB Software

8 Notes and References

---

9 Problems

Problem 1 Prove the Matrix Inversion lemma:

$$(A_1 - A_2 A_4^{-1} A_3)^{-1} = A_1^{-1} + A^{-1} A_2 (A_4 - A_3 A_1^{-1} A_2)^{-1} A_3 A_1^{-1}$$

Problem 2 Show that a set of vectors is linearly dependent if and only if one of the vectors is a linear combination of the other vectors in the set.

Problem 3 Show that any linear combination of vectors from $\mathcal{L}$ is also in $\mathcal{L}$.

Problem 4 Are the vectors in the basis for a particular subspace unique?

Problem 5 Given a basis for a subspace $\mathcal{L}$, can every vector in $\mathcal{L}$ be uniquely represented as a linear combination of these basis vectors?

Problem 6 Show that if A is a symmetric matrix (i.e., $A = A^T$, where the superscript T denotes transpose), then

$$x^T A y = y^T A x$$

for all vectors x and y of the appropriate dimension.

Given

$$f(x) = \frac{1}{2} x^T A x + b^T x + \alpha$$

where A is an $n \times n$ symmetric matrix, b is an n-dimensional vector, and α a scalar. Show that

$$\nabla_x f(x) = Ax + b$$

and

$$H = \nabla_x^2 f(x) = A$$

Problem 7 Consider the following matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 4 \end{bmatrix}$$

1. Is this matrix invertible? Justify your answer!
2. Find all eigenvalues and the corresponding eigenvectors of this matrix. Are the eigenvectors linearly independent or linearly dependent? Justify your answer!
3. Use the Cayley-Hamilton theorem to find the inverse of A (if A is invertible).

Problem 8 Consider the same matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 4 \end{bmatrix}$$

1. Is this matrix positive-definite? negative-definite? indefinite? Justify your answer!
2. Can 2 different matrices have the same eigenvalues? Justify your answer!

Problem 9 Let A be an $n \times n$ constant matrix. Then, we can define the matrix exponential $exp(At)$ as an $n \times n$ matrix, solution to the matrix differential equation

$$\frac{dF(t)}{dt} = AF(t)$$

where $F(0) = I = n \times n$ identity matrix. In other words, $exp(At)$ satisfies the above matrix differential equation and also $exp(A0) = I$.

1. Use the facts given so far and the property of the Laplace transforms

$$\mathcal{L}\{\dot{F}(t)\} = s\mathcal{L}\{F(t)\} - F(0)$$

   to show that

$$e^{At} = \mathcal{L}^{-1}[sI - A]^{-1}$$

2. Even if you don't prove the previous relation, use it to find e^{At} if

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Problem 10 Let A be a matrix from R^n to R^m i.e. A has m rows and n columns, so that if $x \in R^n$, then $y = Ax \ \in R^m$.

1. Show that the Null space of A, $\mathcal{N}(A)$ is a subspace of R^n and that the range of A, $\mathcal{R}(A)$ is a subspace of R^m.

Hence, $\mathcal{N}(A)$ and $\mathcal{R}(A)$ have well-defined dimensions. In fact, the dimension of $\mathcal{R}(A)$ is the rank of A denoted by $rank(A)$ which is the number of linearly independent columns in A. Denote the dimension of $\mathcal{N}(A)$ by $nl(A)$. Then consider any basis $\{x_i\}_{i=1}^{nl(A)}$ of $\mathcal{N}(A)$, and complete this basis into a basis of R^n by adding $n - nl(A)$ linearly independent vectors $\{x_i\}_{i=nl(A)+1}^{n}$. Therefore, any $x \in R^n$ may be written as a linear combination of x_i so that

$$x = \sum_{i=1}^{n} \alpha_i x_i$$

1. By forming Ax show that the vectors $\{y_i\}_{i=nl(A)+1}^{n} = \{Ax_i\}_{i=nl(A)+1}^{n}$ span $\mathcal{R}(A)$ and are linearly independent. Therefore, show that

$$rank(A) + nl(A) = n$$

Problem 11 Suppose x is a vector in $\mathcal{N}(A)$. Show that x is orthogonal to every vector in $\mathcal{R}(A^T)$.

Problem 12 Given a set of vectors $\{x_1, \cdots, x_N\}$. Show how to find a set of *orthonormal* vectors $\{y_1, \cdots, y_M; M \leq N\}$ spanning the same space that was spanned by $\{x_i\}$. Recall that an orthonormal set of vectors is a set of linearly independent vectors each having norm of 1.

Verify that complex numbers can be represented by matrices via

$$z = a + ib \sim \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

with the usual rules of complex numbers and of matrix manipulations, i.e. check addition, and multiplication.

Problem 13 Given a top-companion matrix

$$A = \begin{bmatrix} -a_1 & -a_2 & \cdots & -a_{n-1} & -a_n \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

Show that if λ is an eigenvalue of A then $p = [\lambda^{n-1} \ \cdots \ \lambda \ \ 1]^T$ is the corresponding eigenvector.