 MODULE 7

Topics: Linear operators

We are going to discuss functions = mappings = transformations = operators from one vector space $V_1$ into another vector space $V_2$. However, we shall restrict our sights to the special class of linear operators which are defined as follows.

**Definition:** An operator $L$ from $V_1$ into $V_2$ is linear if

$$L(x + \alpha y) = Lx + \alpha Ly \quad \text{for all } x, y \in V_1 \quad \text{and } \alpha \in F.$$ 

In this case $V_1$ is the domain of $L$, and its range, denoted by $R(L)$, is contained in $V_2$.

**Examples:**

1) The most important example for us:

$$V_1 = \mathbb{R}_n \quad \text{(or } \mathbb{C}_n) , \quad V_2 = \mathbb{R}_m \quad \text{(or } \mathbb{C}_m)$$

and

$$Lx \equiv Ax$$

where $A$ is an $m \times n$ real (or complex) matrix.

2) Let $K(t, s)$ be a function of two variables which is continuous on the square $[0,1] \times [0,1]$.

Define $L$ by

$$(Lf)(t) \equiv \int_0^1 K(t, s)f(s)ds$$

then $L$ is a linear operator from $C^0[0,1]$ into $C^0[0,1]$. $L$ is called an integral operator.

3) Define $L$ by

$$(Lf)(t) \equiv \int_0^t f(s)ds$$

then $L$ is a linear operator from $C^0[0,1]$ into the subspace $M$ of $C^1[0,1]$ defined by

$$M = \{g : g \in C^1[0,1], \ g(0) = 0\}.$$ 

4) Define the operator $Df$ by

$$(Df)(t) \equiv f'(t)$$

then $D$ is a linear operator from $C^1[0,1]$ into $C^0[0,1]$. 

39
5) Define the linear operator \( Lu \) by

\[
(Lu)(t) \equiv \sum_{n=0}^{N} a_n(t)u^{(n)}(t)
\]

then \( L \) is a linear operator from \( C^N[a, b] \) into \( C^0[a, b] \). \( L \) will be called an \( N \)th order linear differential operator with variable coefficients.

**Definition:** The inverse of a linear operator is the operator which maps the element \( Lx \) in the range of \( L \) to \( x \) in the domain of \( L \).

**Theorem.** A linear operator can have an inverse only if \( Lx = 0 \) implies that \( x = 0 \).

**Proof.** If \( Lx = y \) then the inverse of \( L \) is the mapping which takes \( y \) to \( x \). Suppose now that \( Lx_1 = y \) and \( Lx_2 = y \). Then by linearity \( L(x_1 - x_2) = 0 \). If \( x_1 - x_2 \neq 0 \) then there is no function which maps every \( y \) in the range of \( L \) uniquely into the domain of \( D \), i.e., the inverse function does not exist.

As an illustration we consider examples 3 and 4. We see that for any \( f \in C^0[0, 1] \)

\[
(D(Lf))(t) \equiv \frac{d}{dt} \int_{0}^{t} f(s)ds = f(t).
\]

On the other hand,

\[
(L(Df))(t) \equiv \int_{0}^{t} f'(s)ds = f(t) - f(0).
\]

So \( D \) is the inverse of \( L \) on the range of \( L \) in the first case but \( L \) is not the inverse of \( D \) in the second case. Note that

\[
(Lf)(t) \equiv 0
\]

implies that \( f(t) \equiv 0 \) as seen by differentiating both sides, but

\[
(Df)(t) \equiv 0
\]

does not imply that \( f(t) \equiv 0 \) since any constant function would also serve. However, if we consider \( D \) as an operator defined on the space \( M \) defined in 3) above then \( f(0) = 0 \) and the integration denoted by \( L \) is indeed the inverse of the differentiation denoted by \( D \). These examples serve to illustrate that when we define an operator we also have to specify its domain.
Linear operators from $\mathbb{R}^n$ (or $\mathbb{C}^n$) into $\mathbb{R}^m$ (or $\mathbb{C}^m$)

We are now considering the case of

$$Lx = Ax$$

where $A$ is an $m \times n$ matrix with entries $a_{ij}$. It is assumed throughout that you are familiar with the rules of matrix addition and multiplication. Thus we know that if $Ax = y$ then $y$ is a vector with $m$ components where

$$y_i = \sum_{j=1}^{n} a_{ij}x_j$$

i.e., we think of dotting the rows of $A$ into the column vector $x$ to obtain the column vector $y$. However, this not a helpful way of interpreting the action of $A$ as a linear operator. A MUCH MORE useful way of looking at $Ax$ is the following decomposition

$$(7.1) \quad Ax = \sum_{j=1}^{n} x_j A_j$$

where $x = (x_1, \ldots, x_n)$ and $A_j$ is the $j$th column of $A$ which is a column vector with $m$ components. That this relation is true follows by writing

$$x = \sum_{j=1}^{n} x_j \hat{e}_j$$

where $\hat{e}_j$ is the $j$th unit vector, and by observing that $A\hat{e}_j = A_j$. The immediate consequence of this interpretation of $Ax$ is the observation that

$$R(A) = \text{span}\{A_1, \ldots, A_n\}.$$ 

Many problems in linear algebra revolve around solving the linear system

$$Ax = b$$

where $A$ is an $m \times n$ matrix and $b$ is a given vector $b = (b_1, \ldots, b_m)$. It follows immediately from (7.1) that a solution can exist only if $b \in \text{span}\{A_1, \ldots, A_n\}$. Moreover, if the columns of $A$ are linearly independent then $Ax = 0$ has only the zero solution so that the solution of
Ax = b would have to be unique. In this case the inverse of A would have to exist on \( R(A) \) even if A is not square.

However, we usually cannot tell by inspection whether the columns of A are linearly independent or whether \( b \) belongs to the range of A. That question can only be answered after we have attempted to actually solve the linear system. But how do we find the solution \( x \) of \( Ax = b \) for an \( m \times n \) matrix \( A \)?

**Gaussian elimination**

It is assumed that you are familiar with Gaussian elimination so we shall only summarize the process. We subtract multiples of row 1 of the system from the remaining equations to eliminate \( x_1 \) from the remaining \( m - 1 \) equations. If \( a_{11} \) should happen to be zero then this process cannot get started. In this case we reorder the equations of \( Ax = b \) so that in the new coefficient matrix \( a_{11} \neq 0 \). The process then starts over again on the remaining \( m - 1 \) equations in the \( m - 1 \) unknowns \( \{x\}_j=2 \). Eventually the system is so small that we can find its solution or observe that a solution cannot exist. Back-substitution yields the solution of the original system, if it exists.

**The LU decomposition of \( A \)**

For non-singular square matrices Gaussian elimination is equivalent to factoring \( A \) (or a modification \( PA \) of \( A \) obtained by interchanging certain rows of \( A \)). A consistent approach to Gaussian elimination for an \( n \times n \) matrix is as follows:

1) Let \( a_{ij}^{(1)} = a_{ij} \) (the original entries of \( A \)). Then for \( k = 1, \ldots, n - 1 \) compute the multipliers

\[
m_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, \quad i = k + 1, \ldots, n
\]

(where we have assumed that we do not divide by zero) and overwrite \( a_{ij}^{(k)} \) with

\[
a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik}a_{kj}^{(k)} \quad i = k + 1, \ldots, n \quad j = 1, \ldots, n.
\]

We denote by \( A^{(k)} \) the matrix with elements \( a_{ij}^{(k)}, \ i = k, \ldots, n; \ j = 1, \ldots, n \).

2) Let \( L \) be the lower triangular matrix with entries

\[
L_{ii} = 1
\]
\[ L_{ij} = m_{ij} \quad j < i. \]

Let \( U \) be the upper triangular matrix with entries

\[ U_{ij} = a^{(i)}_{ij} \quad j \geq i. \]

\((U \text{ is actually the matrix } A^{(m)}, \text{ but in computer implementations of the } LU \text{ factorization the zeros below the diagonal are not computed. In fact, the elements of } L \text{ below the diagonal are usually stored there in } A^{(m)}.\)

**Theorem:** \( A = LU. \)

**Proof.** \((LU)_{ij} = (m_{i1}, m_{i2}, \ldots m_{i,i-1}, 1, 0, \ldots, 0) \cdot (u_{1j}, u_{2j}, \ldots, u_{jj}, 0, \ldots, 0) = \sum_{k=1}^{i-1} m_{ik} a^{(k)}_{kj} + a^{(i)}_{ij} = \sum_{k=1}^{i-1} [a^{(k)}_{ij} - a^{(k+1)}_{ij}] + a^{(i)}_{ij} = a^{(1)}_{ij} = a_{ij}.\) We see that under the hypothesis that all elements of \( L \) can be found the original matrix has been factored into the product of two triangular matrices. This product allows an easy solution of

\[ Ax = b. \]

Let \( y \) be the solution of

\[ Ly = b \]

then \( x \) is the solution of

\[ Ux = y \]

because \( Ax = LUx = Ly = b.\) An example may serve to clarify this algorithm. Consider the system

\[
Ax = \begin{pmatrix} 4 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 4 \end{pmatrix} x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.
\]

For \( k = 1 \) we obtain

\[
L_{11} = 1 \\
L_{21} = m_{21} = \frac{2}{4} \\
L_{31} = m_{31} = \frac{1}{4}
\]

and

\[
A^{(2)} = \begin{pmatrix} a^{(1)}_{11} & a^{(1)}_{12} & a^{(1)}_{13} \\ a^{(2)}_{21} & a^{(2)}_{22} & a^{(2)}_{23} \\ a^{(2)}_{31} & a^{(2)}_{32} & a^{(2)}_{33} \end{pmatrix} = \begin{pmatrix} 4 & 2 & 1 \\ 0 & 3 & \frac{3}{2} \\ 0 & \frac{3}{2} & \frac{15}{4} \end{pmatrix}
\]
For \( k = 3 \) we obtain \( L_{32} = m_{32} = \frac{1}{2} \) and

\[
A^{(3)} = \begin{pmatrix}
4 & 2 & 1 \\
0 & 3 & \frac{3}{2} \\
0 & 0 & 3
\end{pmatrix}
\]

Thus

\[
LU = \begin{pmatrix}
1 & 0 & 0 \\
\frac{2}{4} & 1 & 0 \\
\frac{1}{4} & \frac{1}{2} & 1
\end{pmatrix}
\begin{pmatrix}
4 & 2 & 1 \\
0 & 3 & \frac{3}{2} \\
0 & 0 & 3
\end{pmatrix} = A.
\]

In order to solve \( Ax = b \) we now solve

\[
\begin{pmatrix}
1 & 0 & 0 \\
\frac{1}{2} & 1 & 0 \\
\frac{1}{4} & \frac{1}{2} & 1
\end{pmatrix} y = \begin{pmatrix}
1 \\
2 \\
3
\end{pmatrix},
\]

yielding \( y = (1, \frac{3}{2}, 2) \), and

\[
\begin{pmatrix}
4 & 2 & 1 \\
0 & 3 & \frac{3}{2} \\
0 & 0 & 3
\end{pmatrix} x = \begin{pmatrix}
1 \\
\frac{3}{2} \\
2
\end{pmatrix}.
\]

We find that the solution of \( Ax = b \) is

\[
x = (0, \frac{1}{6}, \frac{2}{3}).
\]

Of course, it is possible that the diagonal element \( a_{kk}^{(k)} \) is zero. In this case we interchange row \( k \) of the matrix \( A^{(k)} \) with row \( i \) for some \( i > k \) for which \( a_{ik}^{(k)} \neq 0 \). If \( A \) is non-singular this can always be done. In fact, computer codes for the \( LU \) decomposition of a matrix \( A \) routinely exchange row \( k \) with row \( i \) for that \( i \) for which \( |a_{ik}^{(k)}| \geq |a_{jk}^{(k)}|, j = k, \ldots, n \). This process is called partial pivoting. It assures that \( |m_{kj}| \leq 1 \) and stabilizes the numerical computation. Any text on numerical linear algebra will discuss the \( LU \) factorization and its variants in some detail. In this course we shall hand over the actual solution to the computer.

For subsequent modules we shall retain the following observations:

1) Gaussian elimination can also be applied to non-square system of the form

\[
Ax = b.
\]

If \( m < n \) then the last equation obtained is a linear equation in \( \{x_m, \ldots, x_n\} \). If \( m > n \) then the the last \( m - n \) equations all have zero coefficients. A solution of \( Ax = b \) can
exist only if the last \( m - n \) terms of the source term \( b' \) generated during the elimination likewise vanish.

2) If \( A \) is a non-singular \( n \times n \) matrix then there always exists an \( LU \) decomposition of the form

\[
LU = PA
\]

where \( P \) is a non-singular matrix which permutes the rows of \( A \). If no partial pivoting has to be carried out then \( P = I \).
Module 7 - Homework

1) Define \( Lx \equiv Ax \) where \( A \) is an \( m \times n \) complex matrix. What are the domain and range of \( L \)?

2) Define

\[
(Lf)(t) \equiv f'(t) + \int_0^1 K(t, s)f(s)ds
\]

where \( K \) is continuous in \( s \) and \( t \) on the unit square.

i) What is a suitable domain for \( L \)? What is the corresponding range?

ii) Let \( f(t) \equiv \cos t \) and \( K(x, y) \equiv e^{x-y} \). Find \( Lf \).

3) Let \( L : V_1 \rightarrow V_2 \) denote the following operator and spaces:
   i) \( V_1 = \{ f : f \in C^2[0,1], \ f(0) = f(1) = 0 \} \)
   \( V_2 = C^0[0,1] \)
   \( (Lf)(t) \equiv f''(t) + f(t) \).

   ii) \( V_1 = \{ f : f \in C^2[0,1], \ f(1) = 0 \} \)
   \( V = C^0[0,1] \)
   \( (Lf)(t) \equiv f''(t) + f(t) \).

In both cases prove or disprove: \( Lf = 0 \) if and only if \( f \equiv 0 \).

4) Compute the \( LU \) decomposition of

\[
A = \begin{pmatrix}
3 & 2 & 1 \\
6 & 6 & 3 \\
0 & 2 & 2
\end{pmatrix}.
\]

Use the \( LU \) decomposition to solve

\[
Ax = \hat{e}_1.
\]

5) Let \( P^{ij} \) be the matrix obtained from the \( m \times m \) identity matrix \( I \) by interchanging rows \( i \) and \( j \). Let \( A \) be any \( m \times n \) matrix. What is the relation between \( P^{ij}A \) and \( A \)?

Let

\[
A = \begin{pmatrix}
1 & 2 & 3 \\
1 & 2 & 4 \\
2 & 6 & 1
\end{pmatrix}.
\]

Apply the \( LU \) factorization to \( A \) and to \( P^{23}A \).