Linear Algebra and Differential Equations Chapter Summaries

Chapter 9

The most important idea of Chapter 9 is that you can always construct a linear mapping from one (finite-dimensional) vector space to another by defining the mapping of a vector using its coordinates with respect to a basis. Nonnumerical vectors (polynomial functions in a vector space of polynomials of degree less than or equal to 3 is an example) may be associated with a vector of real numbers with respect to a particular basis. It follows that linear mappings between general vector spaces may be described with matrices of real numbers, just as linear mappings from $\mathbf{R}^n$ to $\mathbf{R}^m$ are. The rest of this chapter summary gives more details.

Linear mappings

A mapping $L : V \to W$ is a *linear mapping* if $L(\boldsymbol{u} + \boldsymbol{v}) = L(\boldsymbol{u}) + L(\boldsymbol{v})$ and $L(c\boldsymbol{v}) = cL(\boldsymbol{v})$ for all $\boldsymbol{u}, \boldsymbol{v} \in V$ and all $c \in \mathbf{R}$.

Coordinates with respect to a basis

A basis determines a unique coordinate system for a vector space. If $\boldsymbol{v} \in V$ and $\mathcal{B} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is a basis for V, then

$$\boldsymbol{v} = \alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_n \boldsymbol{v}_n$$

for a unique set of scalars $(\alpha_1, \ldots, \alpha_n) \in \mathbf{R}^n$. These scalars are called *the coordinates of* $\boldsymbol{v}$ *with respect to the basis* $\mathcal{B}$ and are denoted $[\boldsymbol{v}]_{\mathcal{B}}$.

Linear mappings based on coordinates

If V and W are vector spaces, $\mathcal{B} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is a basis for V, and $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ is any set of vectors in W, then the mapping $L : V \to W$ defined by $L(\boldsymbol{v}) = \alpha_1 \boldsymbol{w}_1 + \cdots + \alpha_n \boldsymbol{w}_n$ where $[\boldsymbol{v}]_{\mathcal{B}} = (\alpha_1, \ldots, \alpha_n)$ is the unique linear mapping from V to W for which $L(\boldsymbol{v}_i) = \boldsymbol{w}_i$ for $i = 1, 2, \ldots, n$. This result says that specifying the linear mapping for a basis uniquely determines the linear mapping, and that such a mapping exists for any basis and any set to which the basis is mapped.

If $V = \mathbf{R}^n$ and $W = \mathbf{R}^m$, the linear mapping from V to W will be equivalent to matrix multiplication by the $m \times n$ matrix

$$A = (\boldsymbol{w}_1 \mid \cdots \mid \boldsymbol{w}_n)(\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_n)^{-1}$$

The second matrix is always invertible because the basis elements are linearly independent.

Invertible mappings

The mapping $L : V \to W$ is an *invertible mapping* if there exists a mapping $M : W \to V$ so that the composition of the mappings is the identity mapping. In other words, $(M \circ L)(\boldsymbol{v}) = M(L(\boldsymbol{v})) = \boldsymbol{v}$ for all $\boldsymbol{v} \in V$.

The linear mapping $L : V \to W$ defined above by its action on the coordinates of $\boldsymbol{v}$ with respect to the basis $\mathcal{B}$ is invertible if and only if the set $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ is a basis for W.

The null space and the range of a linear mapping

Any linear mapping $L : V \to W$ has two associated spaces. The *null space* of the linear mapping L is the set of vectors in V which are mapped to the zero vector in W. The *range* of the linear mapping L is the set of vectors in

W which may be mapped to by some vector in V. The dimension of V is equal to the sum of the dimensions of the nullspace and the range of L.

Column space and row space of a matrix

The *column space* and the *row space* of an $m \times n$ matrix A are the spans of the columns and rows of A respectively. The column space is a vector subspace of $\mathbf{R}^n$ and the row space is a vector subspace of $\mathbf{R}^m$. The *column rank* and the *row rank* are the dimensions of these vector subspaces. The nonzero rows of the reduced row echelon form of the matrix A form a basis for the row space of A. The nonzero rows of the reduced ro echelon form of the matrix A^t form a basis for the column space of A. The rank, row rank, and column rank of A are all equal.

Change of basis

The coordinates of a vector with respect to one basis may be found with respect to another basis by a matrix multiplication. Here are a few examples.

Example 1: Let $V = \mathbf{R}^3$. The standard basis is $\mathcal{E} = \{e_1, e_2, e_3\}$. Another basis is $\mathcal{B} = \{(1,1,1), (1,1,-1), (1,-1,-1)\}$. The vector $(1,2,3)$ in the standard basis is $[(1,2,3)]_{\mathcal{E}} = (1,2,3)$. To find $[(1,2,3)]_{\mathcal{B}}$, construct the matrix B whose columns are the elements of the basis $\mathcal{B}$. The matrix B maps the coordinates with respect to $\mathcal{B}$ to coordinates with respect to $\mathcal{E}$. The matrix B^{-1} maps the other way. So,

$$[(1,2,3)]_{\mathcal{B}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 0 & 1/2 & -1/2 \\ 1/2 & -1/2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ -1/2 \\ -1/2 \end{pmatrix}$$

To verify this, note that

$$2\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Example 2: Let $V = \mathcal{P}_2 = \{a + bt + ct^2 \mid a, b, c \in \mathbf{R}\}$, the vector space of polynomials of degree three or less. The natural basis is $\mathcal{E} = \{1, t, t^2\}$. Another basis is $\mathcal{B} = \{1+t, 1+t^2, t+t^2\}$. Find $[1 + 2t + 3t^2]_{\mathcal{B}}$.

First, note that the coefficients of the basis elements of $\mathcal{B}$ are $\{(1,1,0), (1,0,1), (0,1,1)\}$ with respect to the natural basis $\mathcal{E}$. Using these as columns of a matrix B, the matrix B^{-1} will map the coordinates with respect to $\mathcal{E}$ to those with respect to $\mathcal{B}$.

$$[1 + 2t + 3t^2]_{\mathcal{B}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 \\ -1/2 & 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

To verify this, note that

$$0(1+t) + 1(1+t^2) + 2(t+t^2) = 1 + 2t + 3t^2$$

Chapter 10

The main idea of this chapter is orthogonality, the generalization of perpendicularity to higher dimensions. Orthogonal and orthonormal bases have special properties of practical and theoretical importance. Any basis can generate an orthonormal basis by the process of Gram-Schmidt orthogonalization.

Orthogonal vectors

Nonzero vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ are *orthogonal* if their dot product is zero. Geometrically, this means that there is a right angle between the vectors if they are drawn from the same point. The projection of one vector onto a vector orthogonal to it is the zero vector.

A set of nonzero orthogonal vectors are linearly independent.

A set of orthogonal vectors is *orthonormal* if the length of each vector is 1.

Coordinates of vectors with respect to orthonormal bases are simple. If $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is an orthonormal basis of V, then $\boldsymbol{v} = \alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_n \boldsymbol{v}_n$ where $\alpha_i = \boldsymbol{v} \cdot \boldsymbol{v}_i$ is the dot product of $\boldsymbol{v}$ and $\boldsymbol{v}_i$.

Gram-Schmidt orthonormalization

Gram-Schmidt orthonormalization creates an orthonormal basis from a basis by removing from each vector its projection onto the vectors which precede it and then dividing the remainder by its length so its new length is one. Specifically, let $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ be a basis for the vector space W. Define $\boldsymbol{v}_1 = \boldsymbol{w}_1/\|\boldsymbol{w}_1\|$ and then

$$\begin{aligned} \boldsymbol{v}'_{j+1} &= \boldsymbol{w}_{j+1} - (\boldsymbol{w}_{j+1} \cdot \boldsymbol{v}_1)\boldsymbol{v}_1 - \cdots - (\boldsymbol{w}_{j+1} \cdot \boldsymbol{v}_j)\boldsymbol{v}_j \\ \boldsymbol{v}_{j+1} &= \boldsymbol{v}'_{j+1}/\|\boldsymbol{v}'_{j+1}\| \end{aligned}$$

Then $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is an orthonormal basis of W.

Least-squares fits

Let W be a subspace of $\mathbf{R}^n$ with basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ and let $\boldsymbol{x}_0$ be a vector in $\mathbf{R}^n$. The vector w_0 in W closest to x_0 may be thought of as the orthogonal projection of x_0 into W. It is formed as follows. Form the $n \times k$ matrix $A = (\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_k)$ whose columns are the basis vectors of W. Then $\boldsymbol{w}_0 = A\boldsymbol{b}$ where $\boldsymbol{b} = (A^tA)^{-1}A^t\boldsymbol{x}_0$. In other words, the matrix $A(A^tA)^{-1}A^t$ is a projection matrix from $\mathbf{R}^n$ into W. The vector $\boldsymbol{w}_0 = A(A^tA)^{-1}A^t\boldsymbol{x}_0$ is the projection of $\boldsymbol{x}_0$ into the subspace spanned by the columns of A.

Chapter 13

The main idea in Chapter 13 is that every matrix is similar to a block diagonal matrix of special form, called the Jordan normal form. The similarity matrix for the transformation is made up of eigenvectors and generalized eigenvectors of the matrix.

Diagonalizable matrices

An $n \times n$ matrix A is *real diagonalizable* if it is similar to a real diagonal matrix D. An $n \times n$ matrix A is *complex diagonalizable* if it is similar to a diagonal matrix D whose elements may be real or complex. If A has distinct real eigenvalues, it is real diagonalizable. The converse is not true, however. The identity matrix is an example of a diagonal matrix whose eigenvalues are not distinct. A matrix is diagonalizable if and only if its eigenvectors are linearly independent. It is real diagonalizable if these linearly independent eigenvectors are real.

If $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ are the linearly independent eigenvectors of A and the matrix $T = (\boldsymbol{v}_1|\cdots|\boldsymbol{v}_n)$ is the matrix whose columns are the (possible complex) eigenvectors of A, then $D = T^{-1}AT$ is a diagonal matrix. A matrix is in real block diagonal form if it is a block diagonal matrix where each block is either a 1×1 real number or a 2×2 matrix of the form

$$B = \begin{pmatrix} \sigma & -\tau \\ \tau & \sigma \end{pmatrix}$$

These blocks correspond to complex cojugate pair eigenvalues $\sigma \pm i\tau$.

Multiplicities of eigenvalues

The *algebraic multiplicity* of an eigenvalue λ_0 of $n \times n$ matrix A is the power m of the term $(\lambda_0 - \lambda)^m$ in the factored characteristic polynomial of A. The *geometric mulitplicity* of an eigenvalue λ_0 is the dimension of the associated eigenspace, namely nullity$(A - \lambda_0 I)$. These multiplicities may be read off of a matrix in Jordan normal form defined below.

Jordan normal form

A matrix in *real Jordan normal form* is a real block diagonal matrix where each block is one of the following:

1. a 1×1 real matrix;
2. a $k \times k$ real upper triangular matrix with the same real number repeated on the main diagonal, ones on the super diagonal, and zeros elsewhere;
3. a 2×2 matrix of the form $B = \begin{pmatrix} \sigma & -\tau \\ \tau & \sigma \end{pmatrix}$
4. a $2k \times 2k$ block diagonal matrix where each 2×2 diagonal block is of form B for the same σ and τ and there is a 2×2 identity matrix to the right of each diagonal block B except the last.

The following example shows how to read the eigenvalues and their algebraic and geometric multiplicites from a Jordan real normal form matrix. Blank parts of the matrices are filled with zeros.

Example 1:

$$A = \begin{pmatrix}
0 & & & & & & & & & & & \\
 & -1 & & & & & & & & & & \\
 & & -1 & 1 & 0 & 0 & & & & & & \\
 & & 0 & -1 & 1 & 0 & & & & & & \\
 & & 0 & 0 & -1 & 1 & & & & & & \\
 & & 0 & 0 & 0 & -1 & & & & & & \\
 & & & & & & 2 & -5 & & & & \\
 & & & & & & 5 & 2 & & & & \\
 & & & & & & & & 2 & -5 & 1 & 0 \\
 & & & & & & & & 5 & 2 & 0 & 1 \\
 & & & & & & & & 0 & 0 & 2 & -5 \\
 & & & & & & & & 0 & 0 & 5 & 2
\end{pmatrix}$$

The matrix A has four distinct eigenvalues, 0, -1, $2+5i$ and $2-5i$. The eigenvalue 0 has algebraic and geometric multiplicity 1. The eigenvalue -1 has algebraic multiplicity 5 because it appears on the main diagonal five times. Its geometric multiplicity is 2 because it appears in only two Jordan blocks. The complex conjugate pair of complex eigenvalues $2 \pm 5i$ each have algebraic multiplicity 3 because there are three digaonal 2×2 blocks with the right form. Their geometric multiplicity is each 2 because one real 2×2 block is alone while the other two are part of a larger 4×4 Jordan real block with the identity matrix in the upper right corner.

Chapter 15

Systems of linear differential equations

The system $\cdot X = AX$ where A is an $n \times n$ matrix of real numbers can be solved by finding all the eigenvalues of the matrix A and by finding the eigenvectors corresponding to the eigenvalues. If all the eigenvectors are linearly independent, then for each eigenvalue λ_i and its associated eigenvector $\boldsymbol{v}_i$, the function $X(t) = e^{\lambda_i t}\boldsymbol{v}_i$ is a solution and the general solution is a linear combination of these n functions. In addition, the unique solution to the initial value problem $\cdot X = AX$ with $X(0) = X_0$ is $X(t) = e^{tA}X_0$ where e^{tA} is the matrix exponential function.

nth order linear, homogeneous, differential equations (of a single function) with constant coefficients

The general form of this problem type is

$$\frac{d^n x}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \cdots + a_1\frac{dx}{dt} + a_0 = 0$$

The general solution to this equation is the linear combination of n linearly independent functions, each of which is a solution to the equation. To find n such functions, we find the characteristic polynomial of the equation

$$p(r) = r^n + a_{n-1}r^{n-1} + \cdots + a_1 r + a_0$$

Factor this polynomial into linear and quadratic factors with real coefficients. If $(r - r_i)$ is a factor of $p(r)$ of multiplicity m_i, then $t^j e^{r_i t}$ is a solution for j going from 0 to $m_i - 1$ (m_i functions in all).

If $(r^2 - 2sr + (s^2 + c^2))$ is a factor of $p(r)$ of multiplicity k, then $t^j e^{st}\cos ct$ and $t^j e^{st}\sin ct$ are solutions for j going from 0 to $k-1$ ($2k$ functions). Every solution to the differential equation is a linear combination of functions constructed in this way. The unique solution to an initial value problem associated with this general equation is found by imposing the given initial conditons $x(0) = x_0$, $x'(0) = x_1$, $\ldots$, $x^{(n-1)}(0) = x_{n-1}$ and solving for the coefficients in the linear combination.

Differential Operators

Let $\mathcal{C}^\infty$ be the set of real functions which have derivatives of all orders for all values of t, and define the linear mapping $D : \mathcal{C}^\infty \to \mathcal{C}^\infty$ by $D(f(t)) = f'(t)$. We have shown that this is a linear mapping on the vector space $\mathcal{C}^\infty$. We call it the differential operator. D^2 is the composition of D with itself, D^3, D^4, etc are defined as compositions. If a is a real number aD^k is also a linear mapping on $\mathcal{C}^\infty$ and sums of such operators are also linear mappings on $\mathcal{C}^\infty$.

The expression

$$\frac{d^n x}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \cdots + a_1\frac{dx}{dt} + a_0$$

may be reexpressed as

$$(D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0)x$$

in differential operator notation.

The function x is a solution to the homogeneous differential equation if the operator maps x to the zero function. In this case we say that the operator *annihilates* x.

If a is a real number $D - a$ annihilates e^{at} and all multiples of e^{at}. If s and c are real numbers, $D^2 - 2sD + (s^2 + c^2)$ annihilates $e^{st}\cos ct$ and $e^{st}\sin ct$ and all linear combinations of these functions.

The equation

$$\frac{d^n x}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \cdots + a_1\frac{dx}{dt} + a_0 = 0$$

has characteristic polynomial $p(r) = r^n + a_{n-1}r^{n-1} + \cdots + a_1 r + a_0$ which may also be written with D as the argument $p(D) = D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0$.

nth order linear, non-homogeneous, differential equations (of a single function) with constant coefficients

The general form of an equation of this type is

$$\frac{d^n x}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \cdots + a_1\frac{dx}{dt} + a_0 = g(t)$$

The general solution to this equation consists of the general solution to the associated homogeneous equation, as discussed above, and a particular solution to the non-homogeneous part. We need to find a function $x_p(t)$ which is a particular solution.

Method of undetermined coefficients

This method can only be used if $g(t)$ is itself a solution to some nth order linear differential equation with constant coefficients. Rewrite the equation using differential operator notation:

$$(D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0)x = g(t)$$

Find a differential operator (of smallest possible degree) which annihilates $g(t)$ — call it $q(D)$. Apply $q(D)$ to the equation:

$$q(D)(D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0)x = q(D)g(t) = 0 \qquad (*)$$

The specific solution to $(D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0)x = g(t)$ must be one of the solutions to (*).

The general solution to (*) includes the general solution to $(D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0)x = 0$ This part can be ignored. Consider only new functions which are solutions to (*). Operate on these new functions by $(D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0)$ and set the result equal to $g(t)$ and solve for the coefficients — thus determining their values!