Linear Algebra and Ordinary Differential Equations

Avery Karlin

Fall 2016

Contents

Secondary Textbook: Linear Algebra with Applications by Otto Bretscher Teacher: Dr. Robert Powers

"Matrices give you the courage to do calculations you may not be able to do otherwise, representing a large amount of data as a single letter"

1 Chapter 1 (Paul) - Basic Concepts

1.1 Definitions

1. Differential equations are those which contain either ordinary or partial derivatives, such that it is an ordinary differential equation if it only contains the former, and a partial derivative equation if it contains the latter
2. The order of an equation is given by the highest order derivative within it
3. Linear differential equations are those which can be written in the form: $f(x) = a_n(x)y^{(n)}(x) + a_{n-1}(x)y^{(n-1)}(x) + ... + a_1(x)y'(x) + a_0(x)y(x)$, where a(x) are coefficient functions or constants, while y(x) is the main function
    (a) Linear differential equations have no power on the main function or any of its derivatives
    (b) Nonlinear differential equations are those which cannot be written in this format
4. The solution on an open interval interval is a function y(x) which satisfies the equation on the interval
    (a) The interval restricts a variable value to prevent there from being issues with the solution, such as division by zero
5. Initial conditions are a set of conditions on the solution to determine which solution is the correct one, since there may be infinite possible solutions
    (a) Initial conditions are a value of y(x) or some degree of its derivation for some number
    (b) Not all equations will have any solutions, such that the existence question is if a solution exists or not, and even if there is a solution, it may not be possible to find the solution
    (c) The number of solutions to the equation, and the conditions needed to obtain a single equation are called the uniqueness question
6. Initial value problems are those with an appropriate number of initial conditions, based on the order of the equation (equal to the order of the equation)
7. The interval of validity is the largest possible continuous interval in an initial value problem, such that the solution is valid and contains the value of x within the initial conditions
8. The general solution is the the most general form a solution can take, ignoring initial conditions
9. The actual solution is the solution which satisfies the equation and initial conditions
10. An explicit solution is a solution in the form of y = y(x), such that y only appears on one side, without any exponents, while an implicit solution is non-explicit
    (a) Both the actual and general solution may exist in both forms, though all explicit forms of the implicit solution may not be real solutions

1.2 Direction Fields

1. Direction fields are a graph of arrows to show the value of y'(x) (the slope of y(x)) for different values of y(x) based on the differential equation as x increases
    (a) This is done by first finding each value of y(x) where y'(x) = 0, then testing each region around these values of y(x)

2. These provide a sketch of the solutions, as well as the long term behavior of y(x) as x increases
   (a) The family of solution curves, or set of integral curves, can be drawn based on this to show the graph of every possible solution curve
   (b) It also allows determination of how the solution varies based on the value of y(0), and the $\lim_{x\to\infty} y(x)$ based on the value of y(0)
3. If f'(x) is in terms of both x and f(x), then it must be plotted two dimensionally, first finding the lines where the derivative would be constant
   (a) The line such that f'(x) = 0 should be found, then the same techniques can be used to determine the rest of the field, using regional intervals

2   Chapter 2 (Paul) - First Order Equations

<u>Note:</u> Only contains the first four sections

2.1   Linear Equations

1. The equation must be put in the form of $y'(x) + a_1(x)y(x) = a_0(x)$, such that $a_1(x)$ and $a_2(x)$ are continuous
2. The integrating factor (u(x)) is a function such that $u(x)a_1(x) = u'(x)$
   (a) u(x) can be divided by both sides, then it can be changed to (ln(u(x))' = p(x), which can be integrated to get an equation for u(x)
3. The integrating factor can then be multiplied by the equation, use the product rule, and integrate to get an equation for y(x)
4. Constant variables can be changed to simpler constant variables, since they are an unknown constant regardless, as long as they are changed consistently

2.2   Separable Equations

1. Separable equations are those in the form of $N(y)\frac{dy}{dx} = M(x)$, which can be multiplied by dx and have the integral taken of it, to a solution in some form

2.3   Exact Equations

1. Exact equations are those in the form $M(x,y) + N(x,y)\frac{dy}{dx} = 0$, such that some function, $\Psi(x,y)$, can be found where $\Psi_x = M(x,y)$ and $\Psi_y = N(x,y)$
   (a) Thus, if the derivatives and function of Ψ are continuous, then $M_y = N_x$, while if the latter is not true, it is not exact
   (b) By the chain rule, $\frac{d}{dx}(\Psi(x,y(x))) = 0$ is the equation, such that $\Psi(x,y(x)) = c$ is the implicit solution of the equation

2.4   Bernoulli Differential Equations

1. Bernoulli equations are those in the form $\frac{dy}{dx} + p(x)y = q(x)y^n$, where p(x) and q(x) are continuous on the interval of validity, and n is a real number
2. //Finish

3   Chapter 1 (Bretscher) - Linear Equations

1. Geometrically, vectors are denoted as an arrow from the origin to the coordinates of the vector, written as $\vec{v}$
   (a) It is able to be translated from the origin if necessary (such that if $\vec{v} = (x, y)$ is translated (a, b), it will stretch from (a, b) to (a + x, b + y))
   (b) For some infinite set of vectors, it is represented only as the endpoint from the origin for simplicity sake
   (c) Vectors are assumed to be column vectors unless otherwise stated, to allow for matrix multiplication
2. Gauss-Jordan elimination is considered an algorithm, or a finite procedure written in a fixed symbolic vocabulary, with precise instructions, following a discrete set of steps, requiring no insight to complete
   (a) On the other hand, it is noted that rounding errors can occur with certain algorithms, limited the ability to automate it
3. **READ SECTION 1.3**

4   Chapter 2 - Matrices and Systems of Linear Equations

Note: Does not contain Sections 7-8

4.1   Definition and Notation of Matrices

1. Matrices are said to be m x n, with m horizontal rows and n vertical columns, generally denoted by an uppercase variable, where each value is called an entry, represented as an array within brackets
   (a) Index notation is such that in an m x n matrix, $A = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$
   (b) m x n is called the size of the matrix, while m and n are called the dimensions
2. Two matrices are said to be equal if they hve the same size, and all cooresponding elements within the matrices are equal
3. Matrices with a single row or column are called row/column n-vectors, where n is the lenght of the vector, or simply called row/column vectors
   (a) The elements of the matrix are called the components of the vector
   (b) Regular matrices can thus be broken up into a matrix of m row n-vectors or n column m-vectors
   (c) Lists/matrices of vectors must be in the same direction, such that they can be listed in the cooresponding direction
4. Matrices can be transposed, such that the row and column vectors of the matrix are interchanged, denoted by A^T, such that $a_{ji} = a_{ij}^T$
   (a) Square matrices, such that $A^T = A$, are called a symmetric matrix
   (b) Square matrices, such that $A^T = -A = [-a_{ij}]$, are called a skew-symmetric/anti-symmetric matrix

i. Skew-symmetric matrices must have the leading diagonal fully zero, by definition, since it maps onto itself
5. Square matrices are those with the same number of rows and columns
   (a) The leading diagonal of a square matrix is the set of all elements, a_{ii}, where $1 \leq i \leq n$ for some matrix n x n
      i. The trace of the matrix, denoted tr(A) for matrix A, is the sum of all elements within the leading diagonal
   (b) Square matrices are said to be lower triangular if $a_{ij} = 0 \forall i < j$, and upper triangular $\forall i > j$, such that all elements above/below the diagonal are 0
      i. Matrices are called diagonal matrices if they are both lower and upper triangular, able to be denoted as $A = diag(d_1, \cdot, d_n)$, where $d_i = d_{ii}$
6. Matrix functions are defined as matrices whose elements are single variable functions of $\mathbb{R} \to \mathbb{R}$, able to be similarly written as column or row matrix functions

4.2 Matrix Algebra

1. Matrix addition is defined for matrices of the same size, such that if $A = [a_{ij}], B = [b_{ij}], A + B = [a_{ij} + b_{ij}]$
   (a) Matrix addition and subtraction are commutative and associative
2. Matrix scalar multiplication is defined such that $nA = [na_{ij}]$, where $n \in \mathbb{C}$
   (a) Thus, subtraction is defined as the sum of one matrix and the product of -1 and the other matrix
   (b) Scalar multiplication is associative, has a unit property (1A = A), is distributive over matrix addition, and scalar addition is distributive over matrix scalar multiplication
3. The zero matrix is denoted as 0_{mxn}, since all elements are zero, such that it is matrix addition identity, and such that zero scalar multiplication and the difference of a matrix and itself returns the zero matrix
4. For two vectors of the same size, each either row or column, the dot product is defined such that $a \cdot b = a_1 b_1 + \cdot + a_n b_n$
   (a) The matrix product of a column and a row vector of the same size is extended from this, such that it equals a single element matrix of the dot product value
   (b) For some column n-vector and an m x n matrix, the product is extended such that it is the column matrix of the column vector and each row, forming an column m-vector product
      i. As a result, for some m x n matrix multiplied by a column n-vector, the product is the sum of each column matrix multiplied by the next item in the column n-vector $(A_{mxn} c_{nx1} = c_1 a_{1[mx1]} + c_2 a_{2[mx1]} + \cdots + c_n a_{n[mx1]})$
      ii. Thus, the column-vector product is found by taking the linear combination of the column m-vectors of A
         A. The linear combination of a list of vectors and a list of scalars of the same size is the sum of the scalar product of each cooresponding item in both lists
   (c) For some m x n matrix A and n x p matrix B, the product is defined as the row matrix containing the respective resultant column matrix of each column matrix in B multiplied by A
      i. Thus, $(AB)_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj}$, forming an m x p resultant matrix
      ii. It can also be written in index form, such that $AB_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$ for $1 \leq i \leq$

$m, 1 \leq j \leq p$

    iii. Thus, $AB = [Ab_1, Ab_2, \ldots, Ab_p]$

  (d) Matrix multiplication is noncommutative, but it is associative, and has both left and right distributivity

  (e) Matrices multiplied by themselves are represented normally by exponents

5. The identity matrix of some size n, I_n, is represented as a matrix with 1 on the main diagonal, 0 elsewhere, such that $A_{m \times n} I_n = A_{m \times n}$ or $I_m A_{m \times n} = A_{m \times n}$

  (a) Identity matrices are also called unit matrices

6. The transpose operation has properties such that $(A^T)^T = A, (A+C)^T = A^T + C^T, (AB)^T = B^T A^T$

7. Triangular matrices have the property that the product of two lower/upper square matrices is of the same type of matrix

  (a) In addition, the product of two unit lower/upper square matrices is a unit matrix of the same type

8. All rules that apply to algebraic manipulation of matrices applies equally to both real and complex numbers and functions

  (a) Calculus operations are done similarly, taking the derivative or integrating for each term within the matrix

4.3 Systems of Linear Equations Terminology

1. The system of linear equations of size m x n is the set of m equations, such that $\{a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n = b_k\}_{k=1}^m$

  (a) b_i are the system constants and a_{ij} are the system coefficients

    i. If $b_i = 0 \forall 0 \leq i \leq m$, the system is homogeneous, otherwise nonhomogeneous

  (b) The solution to the system is set of ordered n-tuples of real or complex scalars, $(x_1, x_2, \ldots, x_n)$, such that the sides of the equation are equal

    i. Since the set of all ordered n-tuples of real numbers is $\mathbb{R}^n$, the solution to the system is a subset of that

    ii. Ordered n-tuples within the set of real or complex numbers can also be written as row or column n-vectors

    iii. As a result, operations on row/column n-vectors can be applied easily to elements of $\mathbb{R}^n$ or $\mathbb{C}^n$

  (c) The matrix of coefficients, A, is the matrix such that the coefficients of each equation form each row, beginning with the first equation

  (d) The augmented matrix, $A^\#$, is the matrix of coefficients with an additional final column containing the system constants

2. For some system of n linear equations, there may be no solution, a single solution, or infinite solutions

  (a) Consistent systems are those with at least one solution, while inconsistent systems are those with no solutions

  (b) For a set of equal equations, the solution is the equation itself, while otherwise, assuming there is a solution, the solution is some set of a lower dimension (point, line, plane, etc)

3. For some system of m linear equations, it can be written as the matrix of coefficients multiplied by a column n-vector of the variables of each equation, $x_1, x_2, \ldots, x_n$, set equal to the column m-vector of the system constants

(a) As a result, any system of linear equations can also be written as $A\vec{x} = \vec{b}$, where A is the m x n matrix of coefficients, $\vec{x}$ is the column n-vector of variables, and $\vec{b}$ the column m-vector of matrix constants

    i. The column n-vector of variables is also called the vector of unknwons, and the column m-vector of matrix constants is also called the right-hand-side vector

    ii. As a result, the vector of unknowns is an element of $\mathbb{R}^n$ and the right-hand-side vector is an element of $\mathbb{R}^m$

    iii. Thus, the solution of $A\vec{x} = \vec{B}$ is $\forall \vec{x} \in \mathbb{R}^n$ which satisfy the equation

    iv. This also signifies the linear transformation of a n dimensional vector space to an m dimensional vector space, used as an alternative definition of a vector, able to be used to generate mappings such as rotational matrices

4. For some system of m linear equations, in a column m-vector, $\vec{x}$, the derivative is equal to the derivative of each component equation, and $\vec{x}$ is considered a vector-valued function

5. The vector formulation of some system of first order linear differential equations is written as $\vec{x}'(t) = A\vec{x}(t) + \vec{b}(t)$

4.4 Elementary Row Operations and Row-Echelon Matrices

1. For a system of linear equations, the equations can be permuted/reordered, multiplied by a nonzero constant, and added to another equation within the system without changing the solution

    (a) Thus, on the augmented matrix of the system, the rows can be reordered, multiplied by a nonzero constant, or added to another row

    (b) These are called the elementary row operations, denoted by P_{ij} for permuting row i and j, $M_i(k)$ for multiplying row i by k, and $A_{ij}(k)$ for adding row i multiplied by k to row j

        i. All row operations are reversible, such that there is some inverse to move to the original matrix $(P_{ji}, M_i(\frac{1}{k}), A_{ij}(-k))$

    (c) For some matrix A, if matrix B can be found by a finite sequence elementary row operations, $A \sim B$, such that the matrices are called row-equivalent

        i. The individual operations within a sequence to get B from A is denoted by $A \sim^n B$ for the nth operation

2. Row-echelon form of a matrix is such that all rows containing only zeroes are at the bottom of the matrix, the first nonzero element in any nonzero row is a 1 (all rows have a leading 1), and the leading 1 of any row below the first is to the right of the one above

    (a) This notably does not mean that the leading 1s have to be in a diagonal, but rather there may be a column without a leading one, such that for a reduced matrix, there does not need to be zeroes there in the other rows above

    (b) This allows the back substitution technique to be used, substituting the bottom equation into the one above it, and so forth, to get the solution

    (c) Thus, row-echelon form is an upper triangular matrix with all non-zero rows having a leading 1

3. All matrices have infinite row-equivalent row-echelon matrix which they can be converted to by at least one series of elementary operations

    (a) This is generally done by putting a one in the pivot position, or the topmost position of the first non-zero column/pivot column, followed by putting zeroes below, and moving on to the submatrix under and so forth, which then becomes the pivot column

      i. As a result, linear systems are consistent iff the last column of the augmented matrix row-echelon form is not a pivot column

  (b) Fractions in general must be avoided for simplicity when getting echelon-matrices

4. While there are infinite row-equivalent row-echelon matrices for any matrix, all must have the same number of non-zero rows, called the rank of the matrix, denoted by rank(A) for some matrix A

  (a) For some matrix of n rows, if there is 1 zero row, there must be some set of constants, $c_1\vec{a}_1 + \cdots + c_n\vec{a}_n = 0$, such that each can be written in terms of the others by the elementary operations

  (b) The rank of some matrix of size m x n must be $\leq m, n$, by the definition of row-echelon form

  (c) Rank(A) $= \vec{0}$ is true iff A is a zero matrix

5. Reduced row-echelon matrices are a special case such that each leading one has only zeroes in the column otherwise

  (a) Each m x n matrix is row-equivalent to a unique reduced row-echelon matrix

  (b) This is able to be created by the same procedure, subtracting to produce zeroes above the leading zero after each iteration

4.5 Gaussian Elimination

1. Gaussian elimination is the process of converting the augmented matrix to row-echelon form, and using back substitution to solve the system

  (a) Gauss-Jordan elimination is Gaussian elimination when converting it to reduced row-echelon form, not requiring back substitution, denoted for matrix A as rref(A)

  (b) If rank(A) $=$ rank($A^{\#}$), there is a solution to the matrix

     i. If for some matrix, A_{mxn}, $n > rank(A^{\#})$, (such that there are more columns than rows) there are infinite solutions, setting each surplus variable as a free variable/parameter, such that the rest are bound variables/parameters, in terms of the free ones

     ii. The number of free variables for an m x n matrix is equal to n - rank($A^{\#}$), due to there being n variables total

     iii. If n $=$ rank($A^{\#}$), then there is a single, unique solution

  (c) If $rank(A) < rank(A^{\#})$, there is no solution, since there is a false statement within the matrix

  (d) It can be found that for any matrix, A, $rank(A) = rank(A^{T})$

  (e) For some square matrix, A_{nxn}, if the rank(A) $=$ n, it is considered to be of full rank, while for a non-square matrix, A_{mxn} if either rank(A) $=$ m (full row-rank) or rank(A) $=$ n (full column-rank), it is considered full rank

2. For systems of equations with imaginary numbers, it is noted that it can be multiplied by the complex conjugate as well as purely real numbers, such that the imaginary terms can be removed

3. Homogeneous linear systems are those of the form $\vec{A}\vec{x} = \vec{0}$, such that the solution, $\vec{x} = \vec{0}$ is always valid, called the trivial solution

  (a) As a result, homogeneous linear systems must either have only the trivial solution, or infinite solution, since each plane must pass through the origin geometrically

  (b) In addition, if the number of unknowns, n $>$ m (number of linear equations), there are

infinite solutions by necessity, while if m $\geq$ n, it could have a single or infinite solutions based on the rank

4.6   Inverse of Square Matrices

1. The inverse of some matrix, A_{nxn}, is the matrix, A_{nxn}^{-1}, such that $AA^{-1} = I_n$
   (a) This is able to be used to isolate $\vec{x}$ for some system of linear equations, such that for $A\vec{x} = b$, $\vec{x} = A^{-1}b$ is the unique solution to the system for all b in $\mathbb{R}^n$ by the uniqueness of the matrix inverse
      i. In addition, the converse is true, such that if a matrix has a unique solution (such that rank(A) = n), it must be invertible as well
      ii. Thus, all homogeneous linear equation systems must have an inverse matrix if they only have the trivial solution
      iii. This can be furthered such that if the matrix is consistent for all $\vec{y}$, then it has an inverse matrix
   (b) Inverse matrices are unique, such that if $AB = I_n, AC = I_n$, then $B = C = A^{-1}$
   (c) Matrices which have an inverse matrix are called invertible or nonsingular, while those which are not are called singular
   (d) If A and B are invertible, then A^{-1} is invertible
      i. In addition, AB is invertible, and $(AB)^{-1} = B^{-1}A^{-1}$, as well as A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$
      ii. Further, we can say that if AB is invertible, then both A and B are invertible
   (e) While for any finite matrix, AB = I, then B is the right inverse of A, and CA = I, then C is the left inverse, such that B = C, for an infinite dimensional matrix, the left and right inverses may not be equal, or one/both may not exist
2. Inverse matrices can be found by column by column, solving the system of equations such that $AA_t^{-1} = (I_n)_{t=1}^n$, such that the reduced augmented row-echelon matrix can be solved for, giving the system constants as the column of the inverse matrix
   (a) This effectively is modifying $[AI_n]$ such that it changes to the form $[I_nA^{-1}]$, called the Gauss-Jordan Technique
3. Orthogonal matrices are defined as those such that $A^T = A^{-1}$

5   Chapter 3 - Determinants

<u>Note:</u> Does not contain Sections 3-4

5.1   Definition of the Determinant

1. Permutations are an arrangement of discrete objects in a specific order, such that the number of permutations of a set of n objects is n!
   (a) The number of inversions of some permutation, $N(\vec{x})$ are the number of pairs of items that are not in natural order, found equal to the number of adjacent swaps needed to restore it

     i. Odd permutations are those with an odd parity, such that the number of inversions is odd, while even permutations are those with an even parity, or an even or zero number of inversions

     ii. The parity of a permutation, $\sigma(\vec{x})$ is denoted as 1 if even, -1 if odd, such that $\sigma(\vec{x}) = (-1)^{N(\vec{x})}$

     iii. The interchanging of any items of a permutation creates a parity opposite the original

2. The determinant is based on some value which must be nonzero in order for there to be an inverse matrix, such that the rank is equal to the dimensions of the matrix

   (a) $det(A_{1x1}) = a_{11}, det(A_{2x2}) = a_{11}a_{22} - a_{12}a_{21}, det(A_{3x3}) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$

   (b) $det(A) = \sum_{\vec{p}} \sigma(\vec{p}) a_{1(p_1)} a_{2(p_2)} n_{(p_n)}$, where $\vec{p}$ is each distinct permutation of $\{1, 2, 3, \ldots, n\}$, said to have an order of n

     i. This can be visually calculated by taking arrows from each of the top and bottom rows of the matrix diagonally to the right, going down from the top, getting positive terms for each, up from the bottom, getting negative terms for each, wrapped to the left

   (c) The determinant is also denoted by $|A|$, where A is the matrix

   (d) This can be geometrically found to be equivalent to the cross product of $\vec{a}, \vec{b}$ when written as $\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$

     i. Thus, the area of a parallelogram represented by two vectors is equal to the absolute value of the cross product/determinant similarly

     ii. In addition, for some parallelepiped represented by three vectors, the volume is equal to the absolute value of the determinant A, $\begin{vmatrix} \vec{a} \\ \vec{b} \\ \vec{c} \end{vmatrix}$

        A. Thus, if all three are in the same plane, the determinant would be zero, since the volume is 0, able to test if vectors are coplanar

3. The Wronskian matrix is the 3x3 matrix of three solutions to some differential equation in the first row, and their first and second derivatives below

4. The cofactor matrix is the matrix of the discriminants of the submatrices having removed the row and column of that position, alternating positives and negatives, with the main diagonal being positive (-1^{i+j})

   (a) The minor matrix is the matrix without the sign changes, while the adjoint matrix (adj(A)) is the transpose of the cofactor matrix

   (b) For some invertible matrix, A, $A^{-1} = \frac{1}{det(A)} adj(A)$

5.2 Properties of Determinants

1. For an upper or lower triangular matrix, A_{nxn}, $det(A) = \prod_{i=1}^{n} a_{ii}$

2. For some matrix A_{nxn}, for some matrix B obtained by permuting/switching two rows of A, the determinant is det(A) = -det(B)

   (a) For some matrix obtained by multiplying on row by a scalar, the determinant is $\frac{1}{k} * det(B)$

   (b) For some matrix obtained by adding a multiple of any row of A to some other row, the

determinant is det(B)

3. These elementary row operation multipliers can be used to modify the determinant matrix to upper/lower triangular form, such that it can be calculated easier

    (a) This is due to the determinant of a triangular matrix being the product of the diagonal

4. Furthermore for some square matrices A and B, $det(A^T) = det(A)$ and $det(AB) = det(A)det(B)$, such that the properties true for rows are similarly true for columns/elementary column operations

    (a) It follows as a result that $det(A^{-1}) = \frac{1}{det(A)}$

    (b) In addition, if A has a row or column of zeroes, $det(A) = 0$, and if two rows or columns are equal, then $det(A) = 0$

    (c) If $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ are the row/column vectors of A, if B and C are made up of the same vectors except the ith, and $a_i = b_i + c_i$, then $det(A) = det(B) + det(C)$, used to break up a matrix of sums by row/column

6   Chapter 4 - Vector Spaces

<u>Note:</u> Does not contain Sections 5, 10-12

6.1   Vector Space Definition

1. Vector spaces are nonempty sets, whose elements are called vectors, which are defined for an addition and scalar multiplication operation, with scalars in the set A, called a vector space over A, generally $\mathbb{R}$ (real vector space) or $\mathbb{C}$ (complex vector space)

    (a) It must also be closed under addition and scalar multiplication, such that for elements in the set undergoing those operations, the result is also in the set

    (b) It also must have commutative and associative properties of addition, associative properties of scalar multiplication ($c(d\vec{x}) = (cd)\vec{x}$), and distributive properties of scalar multiplication over both vector and scalar addition

    (c) It also must have an additive inverse in the set, as well as an identity vector of addition (zero vector, not necessarily 0)

    (d) Further, it must have the unit property, such that for a vector scalar multiplied by 1, it is equal to itself

2. The set $\mathbb{R}^n$ is a real vector space for all $n \geq 1$, and the set $\mathbb{C}^n$ for all $n \geq 1$ can be a real or complex vector space depending on restricting the scalars to either set

    (a) In addition, the set of all m x n matrices with real inputs is a real vector space with standard matrix addition and scalar multiplication, denoted $M_{mxn}(\mathbb{R})$ with $M_n(\mathbb{R})$ denoting the real vector space of square matrices

    (b) Other vector spaces include P_n, the set of all real-valued polynomials of degree $\leq n$ with real coefficients, and $C^k(I)$, the set of all real-valued functions which are C^k on the domain, I

3. By the definition of vector spaces, it can be proved that for all vector spaces, the zero vector is unique, and the zero property of scalar multiplication is true by a scalar or a vector of zero

    (a) The additive inverse of each element is unique, and for each element, the additive inverse is equal to the scalar product of -1 and the element

(b) Further, if the scalar product of some element and scalar is the zero vector, the element and/or the scalar are zero

6.2 Subspaces

1. Vector spaces are used to limit the solution set to some vector space of possible unknowns for a differential equations, where the solution set is some subset of the vector space
   (a) If the subset is non-empty and a vector space under the same operations as the greater vector space, it is considered a subspace of the greater vector space
   (b) Only closure under addition and scalar multiplication must be checked, since all other vector space axioms are inherited from the greater vector space or from the definition of closure
       i. To be closed under scalar multiplication, it is essential that it contains the zero vector to be able to form a subspace, such that it also has that axiom true as well
2. All vector spaces have a trivial subspace, or a subspace containing only the zero vector
3. The solution set to any homogeneous system of linear equations $(Ax = 0)$ is a subspace of $\mathbb{C}^n$ for any matrix, A_{mxn}
   (a) The solution set of this equation, as a result, is called the null space of the matrix A, denoted nullspace(A)
   (b) For matrices with only real elements, it is usually denoted as a subspace of $\mathbb{R}^n$ instead for simplicity
   (c) On the other hand, the solution set to any nonhomogeneous system of linear equations is not a subspace, due to not having the zero vector in the set
4. The solution set of the homogeneous linear differential equation, $y'' + a_1(x)y' + a_2(x)y = 0$ on some interval I, is a subspace of $C^2(I)$, called the solution space of the equation
   (a) It can be found that for a linear differential equation, there is a general solution as the combination of two linearly independent equations which span the solution set

6.3 Spanning Sets

1. Linear combinations are expressions of the form, $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k$, for some vector space V, where $v_i \in V, c_i \in \mathbb{C} \forall 1 \le i \le k$, such that the resultant expression $\in V$
   (a) If all vectors within vector space V can be written as a linear combination of some set of vectors, $\{v_1, \ldots, v_k\}$, V is said to be spanned/generated by that set, and the set is aid to span V, or be the spanning set for V
   (b) $\mathbb{R}^3$ is spanned by the axis unit vectors, $\vec{i}, \vec{j}, \vec{k}$ as a result, furthered such that any three noncoplanar vectors span it
   (c) It can be shown that vectors span a set by proving that there is at least one solution for the equations formed by the linear combination
       i. Thus, for some matrix of vectors, $A_{kx1}, [\vec{v}_1, \ldots, \vec{v}_n]$, it spans some vector space V iff the system $A\vec{x} = \vec{v}$ is consistent for all $\vec{v} \in V$
2. For some set of vectors, v, within a vector space, V, the set of all possible linear combianations of v is a subset of V, denoted span(v)
   (a) This can be extended further such that it is noted that span(v) is a subspace of V
   (b) span(v) can also be referred to as the subspace of V spanned by v as a result
   (c) It is noted that the trivial case of the span states that $span(\emptyset) = 0$

6.4 Bases and Dimension

1. Linearly independent finite sets of vectors in a vector space are those such that the linear combination of the vectors is equal to zero only if the constant multipliers are all zero
   (a) Linear dependent finite sets of vectors are those that can equal zero without all zero constants
   (b) For some matrix of vectors, $A = [v_1, v_2, \ldots, v_k]$ where $\vec{v_i} \in \mathbb{R}^n$, $\{v_1, v_2, \ldots, v_k\}$ is linearly independent if $A\vec{x} = 0$ has only a trivial solution
      i. Further, if k > n, then the set is linearly dependent as a result, and if k = n, the set is linearly dependent iff $\det(A) = 0$
2. Minimal spanning sets are spanning sets containing the minimum number of vectors needed to span the vector space, such that none of the vectors can be formed by the linear combination of the other two
   (a) As a result, for all minimal spanning sets within a nontrivial vector space V, the spanning set must be linearly independent
   (b) In addition, any nontrivial, finite linearly dependent vector set contains a linearly independent subset with the same linear span as the dependent set
3. A set of finite vectors in a vector space V is called a basis for V if the vectors are linearly independent and span V, such that it is the minimal spanning set of V
   (a) For infinite-dimensional vector spaces, such as $C^n(I), n \geq 1$, it is impossible to find a finite set of vectors which spans the vector space, such that those which have a finite set are called finite-dimensional vector spaces
   (b) The standard basis of $\mathbb{R}^n$ are the set of n unit vectors of the Cartesian coordinate system
4. For any finite-dimensional vector space with a basis of m vectors, any set of more than m vectors is linearly dependent, such that all bases have the same number of vectors
   (a) As a result, any spanning set must contain at least m vectors for that vector space
   (b) Further, the dimension, dim(V), of the vector space V is the number of vectors in any basis of V
      i. For the trivial vector space, the dimension is defined as 0
      ii. It can be found that $dim(\mathbb{R}^n) = n, dim(M_{mxn}(\mathbb{R})) = mn, dim(P_n) = n + 1$ as well
   (c) If dim(V) = n, then any set of n linearly independent vectors in V, or any set of n vectors that span V are a basis of V
      i. This allows the dimension of the solution set of a differential equation to allow any linearly independent set of n equations to act as the general solution, spanning the set
      ii. As a result, if div(V) = n and S is some set of n vectors in V, if either S is a basis of V, S is linearly independent, or S spans V, then the other two statements are true
   (d) For some subspace of vector space V, the dimension of the subspace must be $\leq$ that of V, where if it is equal to the dimension of V, then the subspace must be equal to V
      i. Further, the basis for the subspace is part of the basis for the greater vector space, V, due to there being some set of linearly independent vectors not spanned by the subspace basis, which can be added to form the greater basis, called extending a basis

6.5   Change of Basis

1. For any vector within a finite vector space, for each basis of the vector space, each vector can be expressed as a unique linear combination of the basis vectors
   (a) The converse is also true, such that if each vector in a vector space can be expressed as a linear combination of a set of vectors, the set is a basis of the space
2. An ordered basis B of some vector space V is an ordered set, $(\vec{v}_1, \ldots, \vec{v}_n)$, such that the n-tuple, $(c_1, c_2, \ldots, c_n)$ for some $\vec{v}$ in V, such that $\vec{v} = c_1 v_1 + \cdots + c_n v_n$, is called the components of $\vec{v}$ relative to the ordered basis B
   (a) The column vector of the components, denoted $[\vec{v}]_B$, is called the component vector of $\vec{v}$ relative to the ordered basis B
3. For some ordered basis B in vector space V, if $\vec{x}, \vec{y} \in V$, then $[\vec{x} + \vec{y}]_B = [\vec{x}]_B + [\vec{y}]_B$ and $[c\vec{x}]_B = c[\vec{x}]_B$
4. For some ordered bases B, $\{\vec{v}_1, \ldots, \vec{v}_n\}$, and C, $\{\vec{w}_1, \ldots, \vec{w}_n\}$, in n-dimensional vector space V, $P_{B \to C} = [[\vec{v}_1]_C, \ldots, [\vec{v}_n]_C]$, where $P_{B \to C}$ is the change of base matrix for B to C
   (a) As a result, for some $\vec{v}$, $[\vec{v}]_C = P_{B \to C}[\vec{v}]_B$
   (b) It follows from this definition that $P_{B \to C}$ and $P_{C \to B}$ are inverses of each other
      i. This can be extended, such that for ordered bases A, B, and C, $P_{A \to C} = P_{B \to C} P_{A \to B}$
      ii. In addition, since for the standard basis E, $[\vec{v}]_E = \vec{v}$, such that each free variable has a basis vector with itself once only, $P_{B \to C} = (P_{C \to E})^{-1} P_{B \to E}$
      iii. Thus, the inverse of the base matrix can be used to get the change of base matrix from a standard basis, such that $P_{EtoB} = [\vec{v}_1 \vec{v}_2 \vec{v}_3]$, when $\vec{v}_n$ are the basis vectors of B in terms of E, such that P_{BtoE} is the inverse

6.6   Rank-Nullity Theorem

1. rowspace(A) of a matrix is the subspace spanned by the vector set of each row of A, while the colspace(A) is the subspace spanned by the vector set of each column
   (a) It is found that for two row-equivalent matrices, the rowspaces of each are equal, and thus the set of nonzero rows of row-echelon form of a matrix is the basis for its rowspace, such that the basis can be found by converting row-echelon
   (b) $colspace(A) = rowspace(A^T)$ by the definition of transposing, such that this can be used to find the basis of a colspace
      i. In addition, the row-echelon form of a matrix can be used to produce simplified vectors, where each column with a leading one is corresponds to a column of the original matrix, making up a basis of the colspace by extension of the previous method
   (c) It follows that rank(A) = dim(rowspace(A)) = dim(columnspace(A)), although the rowspace and colspace may be in different dimensions
2. The nullity of some matrix A, denoted nullity(A), is defined as dim(nullspace(A))
3. The Rank-Nullity Theorem states that for some matrix A_{mxn}, rank(A) + nullity(A) = n, such that the nullity of a matrix is equal to the number of free variables in the solution of $A\vec{x} = 0$
   (a) Thus, for some homogeneous linear system of equations $A\vec{x} = 0$ where A_{mxn}, if rank(A) = n, then there is only the trivial solution, such that nullspace(A) = $\{0\}$
      i. In addition, if $rank(A) = r < n$, then there are infinite solutions, all found by

$\vec{x} = c_1 \vec{x}_1 + \cdots + c_{n-r} \vec{x}_{n-r}$, where $\{\vec{x}_1, \ldots, \vec{x}_{n-r}\}$ is any linearly independent set of n - r solutions

    ii. This is called the general solution of the system

(b) For some nonhomogeneous linear system of equations, $A\vec{x} = \vec{b}$ where A_{mxn}, if b is not in colspace(A), then the system is inconsistent

    i. If b is in colspace(A), and dim(colspace(A)) = n, then there is a unique solution

    ii. If b is in colspace(A) and $dim(colspace(A)) < n$, then there are an infinite solutions, since the set of columns is not a basis, giving infinite combinations to form $\vec{b}$

    iii. Since nonhomogeneous systems are not vector spaces, the general solution is of the form, $\vec{x} = c_1\vec{x}_1 + \cdots + c_{n-r}\vec{x}_{n-r} + \vec{x}_p$, where $\{\vec{x}_1, \ldots, \vec{x}_{n-r}\}$ is a basis for nullspace(A), and x_p is some particular solution

       A. This is due to the difference of two solutions being in the nullspace(A), able to be formed by a linear combination, such that the first is a general solution

7   Chapter 5 - Linear Transformations

<u>Note:</u> Does not contain Sections 2-4, 7, 10

7.1   Linear Transformation Definition

1. For some vector spaces V and W, a mapping T from V into W is some rule which assigns each vector $\vec{v}$ in V to precisely one vector $\vec{w} = T(\vec{v}) \in W$, denoted $T : V \rightarrow W$

(a) It is notable that the same vector in W can have multiple vectors in V mapped to it, but for each vector in V, there may only be one mapping in W

2. Linear transformations T are mappings such that $\forall \vec{u}, \vec{v} \in V, T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$ and $\forall c \in \mathbb{C}, T(c\vec{v}) = cT(\vec{v})$

(a) These are called the linearity properties, V is called the domain of T, and W is called the codomain of T

(b) It is noted that the addition and scalar multiplication operations on each side of the properties are for different vector spaces, and are not required to be the same operations

(c) These properties are used to find the transformation in terms of a different original base without a change of base matrix

(d) Mappings which do not fulfill the linearity properties are called nonlinear transformations

(e) As a result, a mapping is a linear transformation iff $T(c_1\vec{v}_1 + c_2\vec{v}_2) = c_1T(\vec{v}_1) + c_2T(\vec{v}_2)$

    i. Further, since this can be extended for any number of $c_k\vec{v}_k$, for some basis set of those k vectors, it proves that the entire vector space transforms

(f) Further, it is established that $T(\vec{0}_v) = \vec{0}_w$, where each are the respective zero vectors from that vector space, and $T(-\vec{v}) = -T(\vec{v}) \forall \vec{v} \in V$

3. For some m x n real matrix, A, T: $\mathbb{R}^n \rightarrow \mathbb{R}^m$, where T = $A\vec{x}$, then T is a linear transformation, called a matrix transformation

(a) The converses is also true, such that for all linear transformations, T: $\mathbb{R}^n \rightarrow \mathbb{R}^m$, it can be expressed as a matrix transformation, T = $A\vec{x}$, where A is the matrix of the transformation of column standard basis vectors in $\mathbb{R}^n$

    i. This is due to the fact that any vector, $\vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n$, such that by the properties of linear transformations, it is shown to be the same as the transformation

(b) As a result, for the linear transformation T: $\mathbb{R}^n \to \mathbb{R}^m$, $A_{mxn} = [T(\vec{e}_1), \ldots, T(\vec{e}_n)]$, where $\vec{e}_k$ is a standard basis column vector in $\mathbb{R}^n$, is called the matrix of T

    i. This can also be used to model any transformation able to be described by real number standard basis vectors, such as polynomials

(c) A projection transformation is some linear transformation such that $T^2 = T$, such that applying it a second time produces no modification

7.2 The Matrix of a Linear Transformation

1. Generalizing the transformation matrix to all vector spaces, for some vector spaces V and W with ordered bases $B = \{\vec{v}_i\}_{i=1}^n$ and $C = \{\vec{w}_i\}_{i=1}^n$, for linear transformation $T : V \to W$, then $[T]_B^C = [[T(\vec{v}_1)]_C, \ldots, [T(\vec{v}_n)]_C]$ is the matrix representation of T relative to bases B and C

  (a) For some transformation which does not change the vector space or the base, it is called the matrix representation of T relative to the basis B

  (b) Thus, the change-of-basis matrix and the $\mathbb{R}^n$ conversion matrix are just special cases of this matrix

  (c) As a result, it can be used to determine the transformation by the transformation matrix, by the formula stating that $[T(\vec{v})]_C = [T]_B^C[\vec{v}]_B$, such that the general vector in B can be converted to the corresponding vector in C of the new vector space

    i. Thus, using $[T(\vec{V})]_C$ and C itself, $T(\vec{v})$ can be found for some general vector, $\vec{v}$

  (d) This is due to the multiplier acting to transform the given vector first, in terms of the original base, followed by converting it into the correct base rather than leaving it in terms of the transformations of the original base, by the properties of ordered bases and linear transformations

2. It is found by the change of base equations and linear transformation matrix equations, that $[T]_C^C = P_{B \to C}[T]_B^B P_{C \to B}$, giving a relationship between each of the matrix representations of T relative to some base

3. Linear transformation matrices can multiplied for the composition of linear transformations, such that for $T_1 : U \to V, T_2 : V \to W, T_3 : U \to W$, then $[T_3]_A^C = [T_2 T_1]_A^C = [T_2]_B^C[T_1]_A^B$, providing a single matrix for the total linear transformation

4. For some linear transformation $T : V \to W$, with ordered bases B for V and C for W, for some $\vec{v} \in V$, then $\vec{v} \in Ker(T)$ iff $[\vec{v}]_B \in nullspace([T]_B^C)$

  (a) In addition, for some $\vec{w} \in W$, then $\vec{w} \in Rng(T)$ iff $[\vec{v}]_B \in colspace([T]_B^C)$

  (b) The kernel of a transformation T: $V \to W$, Ker(T), is the set of all vectors, $\vec{v} \in V$, such that $T(\vec{v}) = \vec{0}$

  (c) The range of a transformation T: $V \to W$, Rng(T), is the subset of W, consisting of all vectors formed by $T(\vec{v}) \forall V \in \vec{v}$

5. For some linear transformation, $T : V \to W$ with ordered bases B for V and C for W, then T is one-to-one iff $nullspace([T]_B^C) = \{\vec{0}\}$

  (a) In addition, T is onto iff $colspace([T]_B^C) = \mathbb{R}^n$, where n = dim(W)

  (b) Further, $T : V \to V$ is invertible iff T is one-to-one and onto

    i. As a result, it is found that T is invertible iff $[T]_B^C$ is an invertible matrix for all ordered bases B and C, which is equivalent to $[T]_B^C$ is invertible for some ordered bases B and C in V

    ii. Thus, $[T^{-1}]_C^B[T]_B^C = [I]_B^B = I_{dim(B)}$, such that $([T]_B^C)^{-1} = [T^{-1}]_C^B$, allowing T^{-1} to be determined

(c) One-to-one is defined for some transformation, $T : V \to W$, if each distinct element in V is mapped to a distinct element in W, such that for all $\vec{v}_1 \neq \vec{v}_2 \in V$, then $T(\vec{v}_1) = T(\vec{v}_2)$

(d) Onto is defined for some transformation, $T : V \to W$, such that if W = Rng(T)

7.3 The Eigenvalue/Eigenvector Problem

1. For some nxn matrix, A, any values of a constant, λ, such that $A\vec{v} = \lambda\vec{v}$ has non-trivial solutions, are called eigenvalues of A, while the corresponding nonzero vectors are eigenvectors of A

   (a) This is based off being used to solve a system of linear first order differential equations, such that $\frac{d\vec{x}}{dt} = A\vec{x}(t)$, where is a 2-column vectors, such that $\vec{x}(t) = e^{\lambda t}\vec{v}$

   (b) These are also called characteristic values and vectors of A respectively

   (c) The matrix can also be viewed as some matrix of a linear transformation, $T : \mathbb{C}^n \to \mathbb{C}^n$, such that $T(\vec{v}) = A\vec{v} = \lambda\vec{v}$, such that it is a transformation that purely stretches the vector

      i. As a result, it is found that if $\vec{v}$ is an eigenvector of A, then $c\vec{v}$ is for all nonzero scalars c

   (d) Further, by definition, it can be thought that eigenvalues of A are those such that $(A - \lambda I_n)\vec{v} = 0$, such that it must have nontrivial solutions

      i. As a result, by the fact that lack of invertibility implies infinite solutions, $det(A - \lambda I) = 0$ if λ is an eigenvalue of A

      ii. $p(\lambda) = det(A - \lambda I_n)$ is called the characteristic polynomial of A of degree n by the definition of the determinant, and $p(\lambda) = 0$ is called the characteristic equation of A

      iii. Thus, it is found that a matrix A is invertible iff 0 is not an eigenvalue of A

   (e) Matrices are defective if the number of linearly independent eigenvectors are not the same number as the dimensions of the matrix

   (f) Complex eigenvalues and eigenvectors for a real matrix are those found by an imaginary root of the characteristic polynomial, such that it follows that the conjugate of λ and $\vec{v}$ are also eigenvalues and vectors

      i. The conjugate of a complex vector is the conjugate of each component within the vector

      ii. It is noted that the transformation dilation of an eigenvector is a complex multiplier, such that n vectors are able to span the complex vector space of $\mathbb{C}^n$

   (g) For some matrix A with eigenvalue λ and eigenvector $\vec{v}$, then λ^2 is an eigenvalue of A^2 and $\vec{v}$ is an eigenvector of A^2

7.4 Diagonalization

1. For some equation $\vec{x}' = A\vec{x}$, for some linear change of variables, $\vec{x} = S\vec{y}$, then $y' = S^{-1}AS\vec{y} = B\vec{y}$

   (a) Thus, for some square matrices A and B, A and B are similar if some invertible matrix S exists, such that $B = S^{-1}AS$

   (b) Similar matrices can be found to have the same eigenvalues and characteristic polynomial/equations

2. The simplest diagonal matrix is such that $B = diag(\lambda_1, \ldots, \lambda_n)$

(a) For any square matrix A, it is similar to a diagonal matrix iff it is nondefective, such that for some n linearly independent eigenvectors, $\vec{v}_1, \ldots, \vec{v}_n$, S $= [\vec{v}_1, \ldots, \vec{v}_n]$

(b) Matrices which are similar to a diagonal matrix are called diagonalizable, equivalent to being nondefective

(c) This allows the system of differential equations to be uncoupled, such that $y_1' = \lambda_1 y_1$ and $y_2' = \lambda_2 y_2$, such that $\vec{y} = y_1 + y_2$ as the vector span

    i. As a result, for the original system of functions x(t), the identity $\vec{x} = S\vec{y}$ can be used to get the solution

7.5   Matrix Exponential Function

1. The matrix exponential function is defined for some square matrix A_{nxn} as $e^{At} = I_n + \frac{1}{2!}(At)^2 + \frac{1}{3!}(At)^3 + \ldots$, found to converge for all square matrices and values of t to a square matrix

(a) It is also true that for some matrices, A, B, where AB = BA, $e^{(A+B)t} = e^{At}e^{Bt}$

(b) For all square matrices, A, $(e^{At})^{-1} = e^{(-A)t} = e^{-At}$, such that it follows that the inverse of e^{At} is e^{-At}

(c) It is also found that for some diagonal matrix, $A = diag(d_1, \ldots, d_n)$, then $e^{At} = diag(e^{d_1 t}, \ldots, e^{d_n t})$

2. It can be found that for some square nondefective matrix, A, $e^{At} = Se^{Dt}S^{-1}$, where D is the diagonalization, such that the function result can be found by the diagonal

7.6   Jordan-Canonical Form

1. The Jordan-Canonical form is an approximation of diagonalization, relying on the matrix being able to have complex entries

(a) This compensates for the lack of enough linearly independent eigenvalues to form the diagonalization matrix

2. Generalized eigenvectors are found by solving such that $(A - \lambda I)^p \vec{v} = 0$, where p is some positive number

(a) As a result, for some smallest positive integer p, such that for all $\vec{v}$ where $(A - \lambda I)^p \vec{v} = 0$, $(A - \lambda I)^{p-1}\vec{v}$ is a regular eigenvector of A

(b) This can be used to get the remaining linearly independent vector to form a base for a defective matrix

    i. This is based on the fact that for any square matrix A with entries $\in \mathbb{C}^n over \mathbb{C}$, a basis of generalized eigenvectors can be found

(c) For some generalized eigenvector of the lowest possible power, p, $\vec{v}$, the ordered set of the cycle of generalized eigenvectors cooresponding to λ is $((A - \lambda I)^{p-1}\vec{v}, \ldots, (A - \lambda I)\vec{v}, \vec{v})$

    i. $\vec{v}$ is the initial vector, while $(A - \lambda I)^{p-1}\vec{v}$ is the terminal vector

    ii. The only regular/general eigenvector within the cycle is the terminal vector

    iii. The cycle of generalized eigenvectors is linearly independent, proven by multiplying by the matrix to some power for each

    iv. Further, for some matrix where $(A - \lambda I)^p = 0_n$, any vector such that $(A - \lambda I)^{p-1}\vec{v} \neq 0$ can be used as the initial vector of a cycle

3. Jordan blocks corresponding to λ is a square matrix of some size, with λ throughout the leading diagonal, 1s throughout the superdiagonal (diagonal above the leading), and 0s elsewhere

(a) Jordan-Canonical form is thus defined as a square matrix with Jordan blocks of some size on the leading diagonal and zeroes elsewhere
  i. All Jordan-Canonical matrices are upper triangular
  ii. It is noted that the size of the individual blocks do not have to be equal
(b) It can be found that every square matrix is similar to a unique Jordan canonical matrix, denoted JCF(A)
(c) It is found that the S multipler matrix to gain Jordan-Canonical form is made up of each cycle of generalized eigenvectors for each eigenvalue, starting from the terminal to initial vectors
  i. The number of Jordan blocks as a result is the number of linearly independent regular eigenvectors, each equal in size to the number of vectors in the cooresponding generalized cycle
  ii. The order of the vector sequences must be the same as the order of the Jordan blocks in the JCF(A) matrix as well
4. For some system of linear differential equations, $\vec{x}' = A\vec{x}$, it can thus be converted to Jordan-Canonical form, allowing the final system to be solved, using back substitution

8 Chapter 7 - Systems of Differential Equations

8.1 General Results of First Order Linear Systems

1. For some ordered set of n-vector valued functions $X = x_1, \ldots, x_n, the Wronskian of the functions is defined as.$ $= \det(\mathrm{x}_1(t), x_2(t), \ldots, x_n(t))$
  (a) For some point t_0 in I, on which the functions are continuous, if $W[X](t_0) \neq 0$, X is linearly independent on I
    i. This is due to there being at some point that the functions are linearly independent, such that the matrix of the Wronkian is invertible at that point and over the interval, are linearly independent
2. The Existence and Uniqueness Theorems state that for some initial value problem, $\vec{x}'(t) = A(t)\vec{x}(t) + \vec{b}(t), \vec{x}(t_0) = \vec{x}_0$, where A and B are continuous on some interval I, the problem has a unique solution on I
3. For some homogeneous vector differential equations ($\vec{x}'(t) = A(t)\vec{x}(t)$, the set of all solutions for some square-n matrix A, continuous on interval I, is a vector space of dimension n
  (a) This is due to the fact that for some functions $D(\vec{x}) = \vec{x}', T(\vec{x}) = A\vec{x}$, such that the solution set is the kernel of $(D - T)(\vec{x})$
  (b) The initial value problem is taken such that for n solutions, $\vec{x}_i(t_0) = e_i$, where e_i is the standard basis vector
    i. Thus, the $det([x_1, \ldots x_n](t_0)) = 1$, such that the matrix of solutions to the equation is linearly independent
    ii. Since this forms the standard basis vectors, any linear combination of these forms an initial value problem for those scalar multipliers
  (c) It follows that for some $x(t)$ made up of the linear combinations of the standard basis functions, it forms the unique solution
    i. Thus, the solution set is spanned by/has a basis of the n functions that at t_0, such that the function formed by the linear combination is the general solution

4. Thus, for some linear homogeneous first order differential equation, $x' = Ax$, any set of n linearly independent solutions on I is called a fundamental solution set on I
   (a) The matrix, $X(t) = [\vec{x}_1, \ldots, \vec{x}_n]$ is called the fundamental matrix for the equation as a result
   (b) As a result, $\vec{x})t) = \vec{X}(t)\vec{c}$, where $\vec{c} = [c_1, \ldots, c_n]^T$
5. The converse of the Wronkian linear independence definition can be proven, such that if they are linearly independent on an interval, the Wronkian is nonzero for all points in the interval
   (a) This is due to the fact that otherwise, for some IVP, $\vec{x}(t_0) = \vec{0}$ by some linear combination of nonzero functions,

8.2 Vector Differential Equations: Nondefective Coefficient Matrices

1. For the system $\vec{x}' = A\vec{x}$, it becomes easier to solve for some square matrix A of real constants
   (a) By the theory of eigenvalues, it was shown that the solution was of the form $\vec{x}(t) = e^{\lambda t}\vec{v}$, where $\vec{v}$ is a constant vector, where $A\vec{v} = \lambda\vec{v}$
   (b) As a result, for some nondefective matrix with real eigenvalues, the solutions are linearly independent, such that the general solution can be found as a result
      i. This is proven by the Wronkian for the solution set, assuming linearly independent eigenvectors
   (c) The solution can then be taken from the linear combination form, to get a vector-valued function solution form
2. It can also be found for real-valued eigenvalues by diagonalizing the matrix, such that $A = SDS^{-1}$, such that $y' = Dy$
   (a) This allows the general solution to be found from the solution of y, by the identity that $\vec{x} = S\vec{y}$, where S is the matrix of linearly independent eigenvectors
3. For a matrix with complex eigenvalues, since the matrix is a real matrix, the eigenvalues and vectors are in conjugate pairs
   (a) For two real value vector functions, $\vec{u}, \vec{v}$, where the two complex solutions are $\vec{u}(t) \pm i\vec{v}(t)$, then $\vec{u}(t)$ and $\vec{v}(t)$ are real solutions of the system
      i. This is due to the fact that any linear combination of the eigenvectors is a multiple of one of the functions
   (b) Thus, the complex solutions can be found by the formulas above, after which the two real valued vector parts are the solutions
   (c) It is then found that all real-valued solutions of the complex eigenvectors are linearly independent to each other and to all real eigenvector solutions found

8.3 Vector Differential Equations: Defective Coefficient Matrices