Definitions/Concepts in ENSC801 - Linear Systems Theory

Mirza Faisal Beg

Vector Space Concepts and Normed Spaces

In this course, we start with discussion of a *set*, loosely understood to be a collection of elements. The key idea is that sets are well defined, in that either an element is in the set or not in the set. We denote a set by upper case alphabets such as X, Y, G, H, and denote elements in the set by lower case alphabet such as x, y, x_1, y_j etc. These elements take many forms such as points, lines, functions, matrices, sequences of points, even other sets. For example, an element x of the set $X = C[a, b] = \{$ functions $f : [a, b] \to \mathbb{R}$ $\}$ is a continuous real-valued function on domain $[a, b]$. We can then define operations among elements of the set and structures on sets. A vector space is a set on which the operation of addition (+) between any two elements has been defined, and addition of any two elements in such a set must result in another element in the set for it to be called a vector space. A group is a set on which has been defined a law of composition that allows one to *compose* two elements and get another element in the set(group). Structures on set are operations that can be applied to the elements in the set and give some information on these elements. For example, measuring the "size" of an element is possible in vector spaces with the definition of a suitable *norm* structure and these sets on which a norm is defined are called normed spaces. With the structure of an inner-product defined on a vector space, given any two elements, we can compute roughly the geometric notion of *angle* between the elements, with elements being called orthogonal or perpendicular when the inner-product is zero, extending our geometric intuition from high school geometry. The study of sets, and associated operations and structures on these sets in abstract form, forms the bulk of this course. As we will see, a unifying theme of this course is extending our geometric intuition from ordinary three dimensional space to sets whose elements are sequences of scalars or real-valued functions, and collecting apparently diverse results under one or few key principles.

Section 2.2: Vector Space Fundamentals

1. Definition of a vector space where elements support the operations of addition and scalar multiplication. The vector space is the name given to a set that contains all possible linear combinations of its elements.
2. Axioms of a vector space.
3. Real/Complex vector spaces.
4. Combining sets to create larger sets by forming the *Cartesian Product* $X \times Y$ of two vector spaces. $X \times Y = \{\text{ordered pairs } (x, y) : x \in X,\ y \in Y\}$.
5. Cartesian product set is also a vector space with operations $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ and $\alpha(x, y) = (\alpha x, \alpha y)$.
6. Product space $X^n = X \times X \times \ldots$ (n times) $\times X$.

Section 2.3: Subspaces, Linear Combinations and Linear Varieties

1. Definition of a subspace. It is a vector space in its own right.
2. A non-empty subset M of a vector space is called a subspace if $\alpha x + \beta y \in M$ when $x \in M$ and $y \in M$.
3. If a set does not contain the null-element, it cannot be a vector space or a subspace of a vector space.
4. The entire space X is its own subspace. A subspace not equal to the entire space is said to be a proper subspace.
5. Let M and N be a subspace of vector space X. Then the intersection $M \cap N$ is a subspace of X.
6. The sum of two subsets S and T in a vector space, denoted by $S + T$ consists of all vectors of the form $s + t$ where $s \in S$ and $t \in T$.
7. The sum $M + N$ of two subspaces M and N is also a subspace.
8. A *linear combination* of elements $x_1, x_2, \ldots, x_n$ in a vector space is a sum of the form $\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n$ where $\alpha_i \in \mathbb{R}$ are any scalars.
9. A trivial linear combination is where all the scalars are zero. Such a linear combination will give the null element, regardless of the elements x_i being combined.
10. A *convex combination* of elements $x_1, x_2, \ldots, x_n$ in a vector space is a sum of the form $\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n$ where α_i are scalars such that (1) $\alpha_i \geq 0$ and (2) $\sum_i \alpha_i = 1$.
11. A convex combination is a special type of linear combination.
12. The set $[S]$ consists of all linear combinations of vectors from subset S of a vector space X. $[S]$ is a subspace of vector space X, it is also called as the subspace generated by the set S or the span of S.
13. Subspace $[S]$ is the smallest subspace containing the set S. If M is any other subspace containing S, then M contains $[S]$.
14. The translation of a subspace is called a linear variety. If M is a subspace in vector space X, and $z \in X$ is any element in the vector space, then the linear variety is a subset $V = z + M$.
15. Linear variety is not a subspace, it does not contain the origin.
16. Linear varieties are also commonly called affine spaces.
17. Any vector $z \in V$ can act as the translating element and yet give the same linear variety V, however, the origin ($\theta \in M$) gets mapped to different elements $z \in V$.

Section 2.4: Convexity and Cones

1. A convex set K in a vector space X is such that if elements x_1 and x_2 lie in K, then all elements of the form $\alpha x_1 + (1 - \alpha) x_2$ also lie in K, where $0 \leq \alpha \leq 1$. (this basically means that the straight line segment joining any two elements in the set also lies completely in the set).
2. Subspaces are convex. The convex combination $\alpha x_1 + (1 - \alpha) x_2$ of two elements is a specific form of a linear combination $\alpha_1 x_1 + \alpha_2 x_2$ with $\alpha_2 = 1 - \alpha_1$. Subspaces contain all linear

combinations of their elements. Hence, they also contain all convex combinations of their elements.

3. Linear Varieties are convex. The empty set is convex.
4. If K and G are convex sets, then αK and $K + G$ are convex sets.
5. The intersection of an arbitrary collection of convex sets is also convex.
6. The convex cover(convex hull) of a set S in a vector space X is the smallest convex set containing S. It is denoted as $co(S)$.
7. A cone with vertex at origin is a set C such that for all elements $x \in C$, $\alpha x \in C$ for all $\alpha \geq 0$.
8. A cone with vertex p is a translation $p + C$ of cone C at origin.
9. A convex cone is both a convex set and a cone.
10. Subspaces and linear varieties are convex cones. Set of all non-negative continuous functions, for example, is a convex cone in the vector space of all continuous functions.

Section 2.5: Linear Independence and Dimension

1. A vector x is said to be linearly dependent on a set S of vectors if it can be expressed as a linear combination of vectors from S. When a vector x cannot be expressed as a linear combination of vectors from a set S, it is said to be linearly independent of S.
2. A set of vectors is a linearly independent set if each vector of the set is linearly independent of the rest of the set, i. e. it cannot be expressed as a linear combination of the other vectors.
3. The trivial linear combination (using zero for all scalars) will result in the null element regardless of the elements being combined in the sum. Sometimes, a non-trivial (where atleast one or more scalars are non-zero) linear combination can also give the null element. For example, if two elements in the combination are the same, then choice 1 and -1 for these two in the combination gives the null element. Stated another way, a set of elements is linearly independent if the only linear combination that gives the null element is the trivial linear combination with all scalars being zero.
4. A necessary and sufficient condition for the set S of vectors $x_1, x_2, \ldots, x_n$ to be linearly independent is that the only linear combination that gives the null-element is the zero(trivial) linear combination. In symbols, set $\{x_k\}_{k=1}^n$ is linearly independent if and only if

$$\sum_{k=1}^{n} \alpha_k x_k = 0 \implies \alpha_k = 0 \ \forall k.$$

5. Thus, a linearly independent set cannot contain the null element. If the null is part of the set of elements, then any non-zero scalar multiplying the null-vector in the linear combination will still give the result to be the null element, and hence the set will, by above test, fail linear independence.
6. A linearly independent set S is advantageous to use since it allows a unique expression of any element x in the span of S as a linear combination of vectors from S. If $x = \sum_{k=1}^n \alpha_k x_k = \sum_{k=1}^n \beta_k x_k$, then $\sum_{k=1}^n (\alpha_k - \beta_k) x_k = 0$ which from linear independence of $\{x_k\}_{k=1}^n$ gives that $\alpha_k = \beta_k$.

7. A *basis* for a vector space X is a set S of vectors(elements) chosen from the space X with two properties at once: (1) the set S of vectors is linearly independent and (2) the set of vectors in S span the space X (i. e. all possible linear combinations from the set S fill the whole space X, or alternatively stated, every element $xinX$ can be written as a linear combination of the elements in S).

8. (1) Linear independence of basis gives unique expression of elements in terms of the basis elements and (2) any vector in the space X can be written as a linear combination of (*finitely many*) elements of the basis set since it spans the space. Such as basis is called a Hamel basis (for infinitely many elements in the linear combination, we enter into the realm of infinite series where analysis of convergence of such series lead to basis formed by these infinite sets, also called the Schauder basis).

9. The number of elements in the basis is called the dimension of the vector space.

10. If the dimension is finite, the vector space is called finite-dimensional. Otherwise, it is infinite dimensional.

11. Example: Set $S = \{x_1, x_2\}$ where $x_1(t) = 1, x_2(t) = t,\ t \in [a, b]$ spans a two-dimensional subspace in the space of continuous functions defined on interval $[a, b]$. Any element in this subspace has the form $x = ax_1 + bx_2$ i.e. $x(t) = a + bt$ where $a, b \in \mathbb{R}$ are (any) scalars.

12. Any two bases for a finite dimensional vector space contain the same number of elements.

13. When the space $\mathbb{R}^n$ of n-tuple vectors is considered (each vector has n-components) with the special Euclidean basis $e_i = (0 \dots 1 \dots 0)^t$ where the one occurs at the i^{th} position, then this is called n-dimensional Euclidean space and often denoted by E^n.

14. For space $\mathbb{R}^n$, suppose the basis vectors were placed as columns in a matrix. Then an element $v = \sum_k \alpha_k x_k = \sum_k \beta_k y_k = \sum_k c_k e_k$ has coordinates $\alpha = (\alpha_k)^t$ in basis $\{x_k\}_{k=1}^n$, coordinates $\beta = (\beta_k)^t$ in basis $\{y_k\}_{k=1}^n$ and $c = (c_k)^t$ in the Euclidean basis $\{e_k\}_{k=1}^n$. In matrix notation,

$$v = X\alpha = Y\beta = Ic.$$

Section 2.6: Normed Linear Spaces

1. A linear (vector) space X that has a real-valued function $\forall x \in X, \|x\| \in \mathbb{R}$ called a norm defined on it, satisfying the properties:

$$(1)\quad \|x\| \geq 0,\quad \|x\| = 0 \Leftrightarrow x = 0 \text{ (Positivity)} \tag{1}$$
$$(2)\quad \|x + y\| \leq \|x\| + \|y\| \text{ (Triangle inequality)} \tag{2}$$
$$(3)\quad \|\alpha x\| = |\alpha|\|x\|,\ \forall \alpha \in \mathbb{R} \tag{3}$$

   (We will only deal with real vector spaces where the coefficients α are real numbers).

2. The norm is an abstraction of our intuitive notion of size or length.

3. Since $\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|$, therefore, $\|x - y\| \geq \|x\| - \|y\|$. Therefore,

$$\|x\| - \|y\| \leq \|x - y\| \leq \|x\| + \|y\|.$$

Section 2.7: Open and Closed Sets

The introduction of norm structure on vector spaces allows the specification of close and far. This gives a way of defining topological concepts of adjacency, proximity, close and far based on neighborhoods defined through the norm.

1. Open neighborhood (open ball or open sphere) $S(x, \epsilon) = \{y : \|x - y\| < \epsilon\}$ is a ball/sphere centered at point x of radius epsilon. Note that the inequality is strictly less than, i. e. the open ball does not contain the set of elements that are at distance ϵ from x, it contains only those elements that are within distance ϵ of x.
2. An interior point p of a subset A in normed space X is one such that there exists an $\epsilon > 0$ and **all** elements x satisfying $\|x - p\| < \epsilon$ are members of set A. Around an interior point $p \in A$, we can find an open sphere $S(p, \epsilon)$ of some positive radius ϵ that lies completely inside A.
3. The collection of all interior points is called the interior of the set A and denoted by $\mathring{A}$.
4. An open set is one where $A = \mathring{A}$ (every point is an interior point, can find an open ball around each point of radius greater than zero and the ball is completely inside the set).
5. A closure point p (also called a limit point) of a set P is one such that every sphere $S(p, \epsilon)$ of positive radius ϵ contains atleast one point that lies inside the set P. The difference with open sets is that in open sets, all points inside the open sphere $S(p, \epsilon)$ for some positive radius ϵ must lie inside the set, whereas here, for all positive radius ϵ values, at least one (or more) points in $S(p, \epsilon)$ will lie inside P.
6. Collection of all closure points of a set P is called the closure of a set. Its denote by $\bar{P}$.
7. A set P is closed if $P = \bar{P}$.
8. The complement of an open set is closed and the complement of a closed set is open.
9. Intersection of a finite number of open sets is open, union of an arbitrary (even countably infinite) number of open sets is open.
10. Union of finite number of closed sets is closed. Intersection of any arbitrary (even countably infinite) number of closed sets is closed.
11. The empty set as well as the entire space are both open and closed.
12. The interior of a convex set in a normed space is convex. To show this, consider two points x_0, y_0 in the interior i.e. they lie in open balls $S(x_0, \epsilon), S(y_0, \epsilon)$ completely inside the set. Choose any point x, y in these respective open balls, and since these are points inside the set, so is their convex linear combination $z = \alpha x + (1 - \alpha)y$. Show that $\|z - z_0\| < \epsilon$ where $z_0 = \alpha x_0 + (1 - \alpha)y_0$ is also inside the set since x_0, y_0 are. Therefore, we can find an open ball $S(z_0, \epsilon)$ completely contained in the set, hence the interior is also convex.
13. The closure of a convex set in a normed space is convex.

Section 2.8: Convergence

The ability to specify the topological concept of “close” between two elements via measuring the norm of the difference (or error) between the elements allows us to define convergence of a sequence to a limit as the case when the norm of the error (of each element from the limit element) converges to zero.

1. Convergence in the norm (also called *Strong* Convergence): In a normed linear space, an infinite sequence of vectors $\{x_n\}$ is said to converge to a vector x if the sequence $\{\|x_n - x\|\}$ of scalars converges to zero. This can also be written as:

$$\lim_{n\to\infty} \|x_n - x\| = 0.$$

   In that case, we write $x_n \to x$ and call x the limit of the sequence $\{x_n\}$.

2. When $\|x_n - x\| \to 0$, it follows that $\|x_n\| \to \|x\|$. This is easy to see from $\|x_n\| = \|x_n - x + x\| \leq \|x_n - x\| + \|x\|$, and $\|x\| = \|x - x_n + x_n\| \leq \|x - x_n\| + \|x_n\|$, so as $n \to \infty$, we get $\lim_{n\to\infty} \|x_n\| \leq \|x\|$, and $\|x\| \leq \lim_{n\to\infty} \|x_n\|$. Therefore, $\lim_{n\to\infty} \|x_n\| = \|x\|$.
3. If a sequence converges, its limit is unique. Suppose $x_n \to x$ and $x_n \to y$. Then, $\|x - y\| = \|x - x_n + x_n - y\| \leq \|x - x_n\| + \|x_n - y\| \to 0$. Therefore if the sequence is convergent to a limit, the limit is unique.
4. Closed Sets: A set F if closed if and only if every convergent sequence whose elements belong to F has a limit that also belongs to F. The limit of a convergent sequence is a closure-point(also called cluster point or limit-point) of the set, and hence to be closed, the set must contain its limit points. To show the other way, i.e. the F closed implies every convergent sequence has limit in F, show its contrapositive i.e. F not closed implies that a sequence in F can converge outside F. This is obvious since F not closed implies it does not contain some of its limit-points. Therefore, we can have a sequence from F (since we can find a point in the set within any given error to the limit-point, by definition of the limit-point) that converges to these limit points outside F.

Section 2.9: Transformations and Continuity

1. A transformation is a rule that specifies a mapping from an element $x \in X$ to element $y \in Y$. Domain of the transformation is the set of elements in X on which the rule is specified, range is the set of elements in Y to which the transformation can assign values. Transformations are also called as operators.
2. Linear transformation: if X and Y are vector spaces, then $T : X \to Y$ that follows $T(x+y) = T(x)+T(y)$ is said to be a linear transformation or linear operator. Note that the definition and test for linear transformation needs the transformation be defined between two vector spaces in order that $x + y$ (addition in X) and $T(x) + T(y)$, the addition in space Y are defined and valid operations.
3. Define one-to-one(injective), onto(surjective).
4. Functional: a transformation $T : X \to \mathbb{R}$ is called a functional.
5. Continuous Transformation. In order to measure continuous, we need a way to measure close and far, and intuitively, a transformation is continuous if it maps points "close" in the domain to points "close" in the range. We will work with normed vector spaces as these provide a convenient way to measure close and far via the norm.

6. A transformation $T : X \to Y$ between normed linear spaces X and Y is called continuous at point $x_0 \in X$ if for every error $\epsilon > 0$ in output side, there exists an error $\delta > 0$ in the input side such that for all $\|x - x_0\| < \delta \implies \|T(x) - T(x_0)\| < \epsilon$.
7. Continuous Transformation defined in terms of mapping of sequences: A transformation $T : X \to Y$ between normed linear spaced X and Y is called continuous at point $x \in X$ if and only if a sequence of points $x_n \in X$ converging to $x \in X$ is mapped to sequence of points $T(x_n) \in Y$ converging to $T(x) \in Y$ i.e. $x_n \to x \in X \Leftrightarrow T(x_n) \to T(x)$.

Section 2.10: l_p and L_p spaces

1. The space of infinite sequences where each element has the form $x = \{\xi_1, \xi_2, \ldots\}, \xi_i \in \mathbb{R}$ and the norm of each element defined to be

$$\|x\|_p = \|x\|_{l^p} = \left(\sum_{i=1}^{\infty} |\xi_i|^p\right)^{1/p}$$

   is finite ($\|x\|_p < \infty$) is denoted by the space l_p.
2. For $p = \infty$, the space l_∞ consists of elements that are infinite sequences having norm $\|x\|_\infty = \sup_i |\xi_i|$ that is finite in value. Since the sup of elements in the space is finite, then each component in every infinite sequence in this space is also finite (and hence bounded).
3. Space of infinite sequences l_p are a vector space, and therefore referred to as a normed linear(vector) space.
4. L_p space is analogous to l_p space, consists of real-valued, Lebesgue measurable functions $x : \Omega \to \mathbb{R}$ for which $|x(t)|^p$ is Lebesgue integrable. The norm is defined as

$$\|x\|_p = \|x\|_{L^p} = \left(\int_\Omega |x(t)|^p \mathrm{d}t\right)^{1/p}$$

   and is finitely valued.
5. In this space, functions can differ on a set of measure zero and still have the same (integral) norm value. In particular, $\|x\|_p = 0$ does not imply $x = \theta$ function, since a function x having zero norm could take non-zero values on a set of measure zero and still have the same value for the norm.
6. So, $x = y \implies \|x\| = \|y\|$ always, but it is not always true that $\|x\| = \|y\| \implies x = y$, an example being the space L_p under discussion. Therefore, in L_p, elements with same norm can differ on a set of measure zero and these elements are said to be equal almost everywhere i. e.

$$\|x\|_p = \|y\|_p \implies x = y \quad a.e. \text{ (almost everywhere) }.$$

   We will choose not distinguish elements that differ on a set of measure zero, and therefore $L_p(\Omega)$ is a normed linear space.
7. Thus in L_p, there is a set of elements that satisfy $\|x\|_p = 0$.
8. The advantage of L_p spaces over space of Riemann integrable functions is that space L_p is a complete space whereas that is not true of that space of Riemann integrable functions. A sequence of Riemann integrable functions could converge to a function that may not be Riemann integrable. This is never the case for a sequence of Lebesgue integrable functions.

9. For the L_∞ norm, analogous to l_∞ space, this is roughly speaking $\sup_{t\in\Omega} |x(t)|$. However, since for the same norm, there is a set of equivalent elements differing only over set of measure zero, the definition as given is ambiguous, as now we have a supremum operation that gives a set of values. This is made precise by defining the norm to be:

$$\|x\|_\infty = \inf_{y(t)=x(t)\ a.e.} \left[\sup_{t\in\Omega} |y(t)|\right].$$

In words, for all functions y that are almost everywhere equal to the given function x, find their supremum value, giving a set of values, and then take the infimum value of this set. This quantity is called the essential supremum and is unique.

10. Hölder Inequality for l_p spaces. If sequence $x = \{\xi_1, \xi_2, \ldots\} \in l_p$ and sequence $y = \{\eta_1, \eta_2, \ldots\} \in l_q$ where $1/p + 1/q = 1$, and $1 \le p, q \le \infty$, then

$$\sum_{i=1}^{\infty} |\xi_i \eta_i| \le \|x\|_p \|y\|_q.$$

Equality holds iff $(|\xi_i|/\|x\|_p)^p = (|\eta_i|/\|y\|_q)^q$.

11. Hölder Inequality for L_p spaces: If $x, y \in L_p[a,b], L_q[a,b]$ respectively, and $1/p + 1/q = 1$, and $p, q > 1$, then

$$\int_a^b |x(t)y(t)| \mathrm{d}t \le \|x\|_p \|y\|_q.$$

Equality holds if and only if

$$\left(\frac{|x(t)|}{\|x\|_p}\right)^p = \left(\frac{|y(t)|}{\|y\|_q}\right)^q$$

almost everywhere on $[a, b]$.

12. Minkowski Inequality for l_p or L_p spaces. If x and y are in $l_p(L_p)$ with $1 \le p \le \infty$, then so is $x + y$ and $\|x + y\|_p \le \|x\|_p + \|y\|_p$.

Section 2.11: Banach Spaces

1. There are two concepts in the study of convergent sequences, the first is that the elements must get closer and closer and the second is that they may tend to a limit element. The first concept is covered under the definition of Cauchy sequences, and the second concept, ie the existence of the limit element in the space (which may not exist, for example, a sequence from the set of rational numbers that converges but may tend to an irrational number in the limit) says something about whether the space is complete or has ”holes”.

2. A Cauchy sequence $\{x_n\}_{n=1}^{\infty}$ is one where the sequence of reals $\|x_n - x_m\|$ converges to 0 as n, m tend to ∞. Alternatively, we say that a sequence $\{x_n\}_{n=1}^{\infty}$ is Cauchy if given error $\epsilon > 0$, there exists an index $N(\epsilon)$ such that $\forall m,\ n$ greater than N (key here is “for all $m, n > N$”), the norm of the difference $\|x_n - x_m\| < \epsilon$.

3. In a normed space, every convergent sequence is Cauchy. When we say sequence $\{x_n\}_{n=1}^{\infty}$ is convergent, we mean that its limit exists in the space and can be represented by element x also in the space and $\lim_{n\to\infty} \|x_n - x\| = 0$. Now, $\|x_n - x_m\| = \|x_n - x + x - x_m\| \le \|x_n - x\| + \|x - x_m\|$. Hence, as $n, m \to \infty$, we get that $\|x_n - x_m\| \to 0$, which is the Cauchy property.

4. Every Cauchy sequence may not be convergent. Sequences that are Cauchy have elements that get closer, but may not be convergent (limit element may not exist in the space) as the underlying space may have "holes" i. e. lacks completeness. A standard example is that we can create a sequence of rational numbers with elements that get closer and closer to each other, but the limit could be an irrational number(say the number π). Hence, the sequence may be Cauchy, but the limit may not exist in the space of rationals and hence we will say that such a sequence does not converge. This is due to incompleteness of rationals.
5. A normed linear vector space X is said to be complete if the limit of *every* Cauchy sequence from X exists in X. A complete normed linear vector space is called a Banach Space.
6. Completeness of space guarantees the existence of solution to optimization problem as the limit of a sequence of solutions that may be found, and also furnishes a criterion, the Cauchy criterion, to evaluate if the sequence of solutions is converging to a limit without explicitly knowing the limit.
7. A Cauchy sequence is bounded.

Section 2.12: Complete Subsets

1. A subset in a normed vector space is complete if every Cauchy sequence from the subset converges to limit in the subset.
2. In a Banach space, a subset is complete if and only if it is closed.
3. In a normed linear space, any finite-dimensional subspace is complete (closed).

Section 2.13: Extreme Values of Functionals and Compactness

1. In finite-dimensional spaces, Weierstrass theorem states that a continuous function defined on a closed and bounded (which implies compactness in finite dimensional spaces) set achieves a maximum and minimum.
2. Define upper semicontinuous functional.
3. Define compact set in arbitrary normed spaces.
4. (Weierstrass) An upper semicontinuous functional on a compact subset K of a normed linear space X achieves a maximum on K.

Section 2.14: Quotient Spaces

1. The linear varieties obtained by translating a subspace M of vector space X can be regarded as elements of a new vector space, called the quotient space of X modulo M and denoted as X/M. Thus, $X/M = \{V_x = x + M : x \in X, M \text{ subspace of } X\}$.
2. Let M be a subspace of vector space X. Two elements x_1 and x_2 from X are said to be equivalent module M if $x_1 - x_2 \in M$ (or $x_2 - x_1 \in M$). In this case, write $x_1 \equiv x_2$. If one were to place a copy of M on either location x_1 or x_2 to generate a linear variety of M, then the other element will be found to be in the linear variety so generated.
3. The equivalence relation partitions the space X into disjoint subsets that are the linear varieties - translates of subspace M. In each linear variety, all the elements are equivalent,

since their difference (suppose $x + m_1, x + m_2$ are two elements in V_x, then difference is $x + m_1 - (x + m_2)$) lies in M. The linear varieties are also called cosets of M.

4. The unique coset (distinct translates giving linear variety V_x) of M in which an arbitrary element x lies is denoted as $[x]$. Can think of the coset of M containing x as the linear variety $[x] = V_x = x + M$.
5. Addition of linear varieties $[x_1] = x_1 + M$ and $[x_2] = x_2 + M$ gives a new linear variety $[x_1 + x_2] = x_1 + x_2 + M$. Similarly, scalar multiplication of a linear variety merely translates it further as $\alpha[x] = \alpha x + \alpha M = \alpha x + M = [\alpha x]$.
6. The disjoint-cosets of the subspace M of vector space X can also be thought of distinct elements in their own right. This set of disjoint-cosets has itself the structure of a vector space with addition defined as $[x_1] + [x_2] = [x_1 + x_2]$ (addition of cosets of elements x_1 and x_2 is the same as the coset of the element $x_1 + x_2$), and also $\alpha[x] = [\alpha x]$.
7. Thus, the quotient space X/M consisting of elements that are the linear varieties of M is itself a vector space.
8. Can we give the quotient space X/M which above was defined with structure of a vector space the additional structure of a norm? What is the norm of an element $[x] = x + M \in X/M$? Define the norm to be
$$\|[x]\| = \inf_{m \in M} \|x + m\|,$$
i. e. the norm of any element (i.e a coset or linear variety) in the quotient space is taken to be the infimum of the norm of all the elements in that coset or linear variety. The assumption that M is closed ensures that $\|[x]\| > 0$ if $[x] \neq \theta \in X/M$.

Section 2.14: Denseness and Separability

1. A set D is said to be dense in a normed space X if for every element $x \in X$, and every error $\epsilon > 0$, there exists an element $d \in D$ approximating x within that chosen error i. e. $\|x - d\| < \epsilon$. The error $\epsilon > 0$ can be made as close to zero as desired and we can always find some element d within that error radius of $x \in X$.
2. This implies that if set D is dense in X, then for every element $x \in X$, we can find a sequence of elements in D that converge to x. Hence, the closure of D is X.
3. A set is countable (infinite set whose elements, although infinite in number, can at least be counted) if it can be labeled in a one-to-one correspondence with the set of integers (which is also a set with infinite number of elements).
4. A normed space is separable if it contains a dense set that is countable.
5. If a normed space X contains a sequence $\{v_n\}$ with property that for all $x \in X$, there is a unique sequence of scalars $\{\alpha_n\}$ such that $\|x - \sum_{i=1}^{n} \alpha_i v_i\| \to 0$ as $n \to \infty$, then $\{v_n\}$ is called a Schauder basis for X. The expansion of $x \in X$ in terms of this basis is written as $x = \sum_{i=1}^{\infty} \alpha_i v_i$.
6. Example of Schauder basis: Space l^p has a Schauder basis $e_n = \{\delta_{n-j}\}_{j=1}^{\infty}$ (an infinite sequence with a '1' in the n-th place and zeroes elsewhere.
7. If a normed space X has a Schauder basis, then X is separable. However, the converse is not true. A normed space X may be seperable (contain a countable subset that is dense in

the space) but not have a Schauder basis.

Chapter 3: Properties of Hilbert Spaces

ENSC 801 - Linear System Theory

This chapter introduced the structure of inner-product $< \cdot, \cdot >$ on vector spaces and the concept of *orthogonal* ($x \perp y \Leftrightarrow < x, y >= 0$) and *aligned* ($x = \alpha y \Leftrightarrow | < x, y > | = \|x\|\|y\|$). The definition of an inner-product then allows the calculation of a norm induced by this inner-product to be $\|x\| \doteq \sqrt{< x, x >}$, and therefore to topological concepts in inner-product spaces such open, closed, complete etc. that were previously defined using the norm structure. In particular, complete inner-product spaces are called Hilbert spaces. The projection theorem is the main tool from this chapter, that there exists an element in a closed subspace of Hilbert space that minimizes distance to any given element in the Hilbert space. This is a key tool that finds enormously widespread usage.

Section 3.2: Pre-Hilbert Spaces

1. Inner-product and its properties symmetry, linearity and positivity.
2. Norm induced by the inner-product.
3. Parallelogram Law
4. Complete inner-product spaces are Hilbert Spaces.
5. Continuity of inner-product: If $\{x_n\} \rightarrow x$ and $\{y_n\} \rightarrow y$, then $\langle x_n, y_n \rangle \rightarrow \langle x, y \rangle$.

Section 3.3: The Projection Theorem

1. *Orthogonal:* $x \perp y \Leftrightarrow < x, y >= 0$.
2. *Aligned* $|\langle x, y \rangle| = \|x\|\|y\| \Leftrightarrow x = \alpha y$.
3. Pythagorean theorem in inner-product spaces: if $x \perp y$, then $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.
4. Projection theorem in pre-Hilbert spaces (existence of minimizer not guaranteed).
5. Projection theorem in Hilbert spaces (minimizing vector exists and is unique).

Section 3.4: Orthogonal Complements

1. Orthogonal complement of a subset S in Hilbert space H is denoted by $S^\perp$ and is the set of all vectors orthogonal to S.
2. $S^\perp$ is closed subspace.
3. $S \subset S^{\perp\perp}$.
4. If $S \subset T$ then $T^\perp \subset S^\perp$.
5. $S^{\perp\perp\perp} = S^\perp$.
6. Direct Sum of two subspaces: vector space X is the direct sum of two subspaces M and N if every vector $x \in X$ has *unique* representation of form $x = m + n$ where $m \in M, n \in N$. Denote this as $X = M \oplus N$.

7. If a set M is a closed subspace in Hilbert space H, then its orthogonal complement contains enough vectors to generate the whole space via a direct sum $H = M \oplus M^\perp$ and $M = M^{\perp\perp}$.

Section 3.5: Gram-Schmidt Procedure

1. Orthonormal set of vectors are mutually orthogonal and unit norm.
2. Orthogonal set of non-zero vectors is linearly independent set.
3. Gram-Schmidt procedure: Given a sequence $\{x_i\}$ of linearly independent vectors in pre-Hilbert space X, compute an orthonormal sequence $\{e_i\}$ such that the span $[e_1, e_2, \ldots, e_n]$ of the first n orthonormal vectors so computed is the same as the span $[x_1, x_2, \ldots, x_n]$ of the first n vectors in the linearly independent set (for all n).

$$\text{Compute } z_n = x_n - \sum_{i=1}^{n-1} \langle x_n, e_i \rangle e_i, \quad e_n = \frac{z_n}{\|z_n\|}.$$

4. This computation can be viewed as a approximation of each element x_n in the subspace spanned by $[e_1, \ldots, e_{n-1}]$. The normalized error $x_n - \sum_{i=1}^{n-1} \langle x_n, e_i \rangle e_i$ is set as the next orthonormal element.
5. Two key properties emerge from this procedure.
   1. First is that span $[e_1, \ldots, e_k]$ = span $[x_1, \ldots, x_k]$ for all $k = 1, 2, \ldots, n$. Hence, the best approximation of vector x in a $k-1$ dimensional subspace is given by $\hat{x}_k = \sum_{i=1}^{k-1} \alpha_i e_i = \sum_{i=1}^{k-1} \beta_i x_i$. Here, since the set $\{e_i\}$ is orthonormal, the coefficients $\alpha_i = \langle x, e_i \rangle$ as the Gram matrix for this expression is the identity matrix.
   2. Second is that at each step k, the new orthonormal vector e_k is orthogonal to the subspace generated by the preceding $\{x_i\}$, $i = 1, \ldots, k-1$ i. e. $\langle e_k, x_i \rangle = 0 \ \forall i = 1, \ldots, k-1$.

Section 3.6: Approximation in subspace

1. Approximation problem: Suppose $y_1, \ldots, y_n$ are elements of Hilbert space H. These vectors generate a finite-dimensional (closed) subspace M of H. Given an arbitrary element $x \in H$, we want to approximate it by the *closest* element $\hat{x} \in M$. Since the best approximation is in M, its represented as $\hat{x} = \sum_{i=1}^{n} \alpha_i y_i$. Note that the representation is unique only if the elements $y_1, \ldots, y_n$ are a basis, which means they have to be linearly independent.
2. The projection theorem applies, the error $x - \hat{x}$ is orthogonal to M therefore to each $y_i, i = 1, \ldots, n$. Hence, $< x - \hat{x}, y_i >= 0$. This gives $< \hat{x}, y_i >= \sum_{j=1}^{n} \alpha_j < y_j, y_i >=< x, y_i >$. In matrix form, we get

$$\underbrace{\begin{pmatrix} < y_1, y_1 > & < y_2, y_1 > & \ldots & < y_n, y_1 > \\ < y_1, y_2 > & < y_2, y_2 > & \ldots & < y_n, y_2 > \\ \vdots & \vdots & \ldots & \vdots \\ < y_1, y_n > & < y_2, y_n > & \ldots & < y_n, y_n > \end{pmatrix}}_{\text{Transpose of Gram matrix } = G(y_1, \ldots, y_n)^t} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} < x, y_1 > \\ < x, y_2 > \\ \vdots \\ < x, y_n > \end{pmatrix}.$$

3. This system of linear equations is called the *normal equations*. The system is invertible when the vectors $y_1, \dots, y_n$ are linearly independent.
4. In case the vectors $y_1, \dots, y_n$ are linearly dependent, the gram determinant is zero as the Gram matrix is not of full-rank i. e. has a non-empty null space. Therefore, there exist a multiplicity of solutions.
5. If we view the problem as mapping of linear operator $A : E^n \to H$ such that $A\alpha = \sum_{i=1}^n \alpha_i y_i$, then the task is find the $\hat{\alpha}$ that minimizes $\|x - A\alpha\|$ Therefore, first, project $x \in H$ to the range M of A which is the right-hand side of above system. If elements $y_1, \dots, y_n$ are linearly dependent, then operator A has a null-space that has, besides the trivial null element, other elements that are non-zero. Hence, there is a multiplicity of solutions. As we see later, $G(y_1, \dots, y_n)^t = A^\star A$, where $A^\star$ is the adjoint operator of operator A.

Section 3.7: Fourier Series

1. Approximation of any vector $x \in H$ in a finite dimensional subspace M spanned by orthonormal vectors $\{e_i\}_{i=1}^n$ is given by $\hat{x} \in M$ closest to x where

$$\hat{x} = \sum_{i=1}^{n} \alpha_i e_i.$$

Because $\{e_i\}_{i=1}^n$ are orthonormal, the Gram matrix is the identity matrix and the normal equations are instantly solved to give $\alpha_i =< x, e_i >$. Hence, the best approximation of an element $x \in H$ in the subspace M spanned by orthonormal vectors $\{e_i\}_{i=1}^n$ is given by

$$\hat{x} = \sum_{i=1}^{n} < x, e_i > e_i.$$

2. If $x \in M$, then can write

$$x = \sum_{i=1}^{n} < x, e_i > e_i.$$

3. The *closed* subspace generated by a subset S in Hilbert Space H is $[\bar{S}]$.
4. If M is an infinite-dimensional closed subspace generated by orthonormal set $\{e_i\}_{i=1}^\infty$, then when can

$$x = \sum_{i=1}^{\infty} < x, e_i > e_i$$

be written for $x \in H$? Note that now, $n = \infty$ and we go from a finite dimensional subspace to an infinite dimensional one.

5. First issue is that upon forming an infinite series $\sum_{i=1}^\infty \zeta_i e_i$, we run into problems related to convergence - does the beast that consumes infinitely many terms in the sum converge to some limit? What if it keeps growing as more terms (it is an infinite sum after all!) are continually added?
6. Secondly, suppose this infinite series does converge, then write $\hat{x} = \sum_{i=1}^\infty \zeta_i e_i$ and $\hat{x}$ belongs to closed subspace generated by set $S = \{e_i\}$ and now when does it make sense to write that $\hat{x}$ approximates x?

7. How to define convergence of an infinite series of the form $\sum_{i=1}^{\infty} x_i$: An infinite sum of the form $\sum_{i=1}^{\infty} x_i$ converges (i.e. a limit element exists) if the sequence of partial sums $s_n = \sum_{i=1}^{n} x_i$ converges, and if x is the element to which s_n converges as $n \to \infty$, we write $x = \sum_{i=1}^{\infty} x_i = \lim_{n\to\infty} \sum_{i=1}^{n} x_i$ (again remembering that on the left is the limit of the expression on the right).

8. Given a set of scalars $\{\zeta_i\}$, consider a series formed from these scalars to be of the form $\sum_{i=1}^{\infty} \zeta_i e_i$ where $\{e_i\}_{i=1}^{\infty}$ is an orthonormal sequence in Hilbert space H. This series $\sum_{i=1}^{\infty} \zeta_i e_i$ converges to element $x \in H$ if and only if $\sum_{i=1}^{\infty} |\zeta_i|^2 < \infty$. Also, it follows that the scalars can be related to the element x by $\zeta_i =< x, e_i >$.

9. In order now to start with an element $x \in H$ and write it as an infinite series, we have to (1) make sure that the infinite series converges, (2) the orthonormal set $S = \{e_i\}$ is complete (so that closed subspace generated by these elements is the whole space H).

10. *Bessel's* Inequality - sum of squares of projections of a vector $x \in H$ onto a set of mutually perpendicular directions cannot exceed the square of the length of the vector itself i.e

$$\sum_{i=1}^{n\leq\infty} |< x, e_i >|^2 \quad \leq \quad \sum_{i=1}^{\infty} |< x, e_i >|^2 \quad \leq \quad \|x\|^2.$$

11. Let $x \in H$ and given $\{e_i\}_{i=1}^{\infty}$ an orthonormal sequence in H. Then, series $\sum_{i=1}^{\infty} < x, e_i > e_i$ converges (following the Bessel's inequality that guarantees finiteness of the sum of the coefficients), and the element $\hat{x}$ to which it converges is in the closed subspace M generated by the $\{e_i\}$'s. Hence we denote this fact by writing $\hat{x} = \sum_{i=1}^{\infty} < x, e_i > e_i$ where we remember that on the left hand side is the convergent limit of the series on the right hand side. The approximation is not yet good enough to fully represent the given element $x \in H$, there is still an error(difference) $x - \hat{x}$ that is orthogonal to M.

12. Next step is to discuss when $M = H$ i.e. the concept of a complete orthonormal sequence.

Section 3.8: Complete Orthonormal Sequences

1. If the closed subspace M generated by the orthonormal sequence $\{e_i\}_{i=1}^{\infty}$ is H, then any element in the space can be written as $x = \sum_{i=1}^{\infty} < x, e_i > e_i$.

2. An orthonormal sequence $\{e_i\}$ in a Hilbert space H is complete if the closed subspace generated by $\{e_i\}$ is H. In this case, the inequality in the Bessel's inequality becomes an equality for all $x \in H$ and also $x = \sum_{i=1}^{\infty} < x, e_i > e_i$.

3. In other words, a sequence $\{e_i\}$ is complete in Hilbert space H if the space spanned by $\{e_i\}$ is dense in H. The closure of this space is H i. e if we denote $V = [\{e_i\}]$, then $\bar{V} = \overline{[\{e_i\}]} = H$.

4. Given an orthonormal sequence, how to show its complete? Show that the only vector orthogonal to each of the e_i's is the null-vector i. e. $M^{\perp} = \{\theta\}$.

5. To show this, let V be the subspace generated by the given orthonormal set $M = \{e_i\}$. Assume that the span of M which we denote here as $V = [M]$ is dense in H i. e. $\bar{V} = H$. Let $x \in M^{\perp}$. Hence $x \in \bar{V} = H$ (because $M^{\perp}$ is also in H). Since x is a closure point of V, there is a sequence $x_n \in V$ converging to x i. e. $x_n \to x$. Since $x \in M^{\perp}$, this implies that $< x_n, x >= 0$ as x_n's are in V and $M^{\perp} \perp V$. Since $x_n \to x$, then by continuity of

inner-product, we get $< x, x >= 0$.This implies $\|x\| = 0$, which in turn implies $x = \theta$. Hence, since this hold for any $x \in M^\perp$, we conclude that all the elements of $M^\perp$ must be equal to the null element giving that $M^\perp = \{\theta\}$.

Section 3.9: Approximations and Fourier Series

1. Gram-Schimdt procedure can be interpreted as a way of inverting the gram-matrix.
2. Given a set of linearly independent vectors $\{y_i\}_{i=1}^n$ spanning subspace M in the Hilbert Space H, then instead of solving the normal equations for the $\hat{x} \in M$ thats closest to $x \in H$, first use Gram-Schmidt procedure to obtain an orthonormal set $\{e_i\}_{i=1}^n$ generating M and then use the Fourier series representation $\hat{x} = \sum_{i=1}^n < x, e_i > e_i$.

Section 3.10: Dual Approximation Problem

1. Understand the modified projection theorem. In a linear variety $V = x + M$ (with M being a closed subspace in Hilbert space), there exists a unique element $x_0 \in V$ of minimum norm such that $x_0 \perp M$.
2. Let N be the subspace spanned by elements $\{y_i\}, i = 1, \ldots, n$ in H, then the set of elements that satisfy $< x, y_i >= 0, i = 1, \ldots, n$ is the subspace $N^\perp$.
3. The set of elements that satisfy $< x, y_i >= c_i, i = 1, \ldots, n$ is a translation of $N^\perp$ which we denote as $V = z + N^\perp$ for some $z \in H$. This set is said to be of codimension n, since the orthogonal complement of the subspace producing the given linear variety has dimension n.
4. Among all possible solutions of these set of constraints i. e. the set of solutions that are the elements in V, is there a unique element of minimum norm that we can identify?
5. Yes, by the modified projection theorem (since $N^\perp$ is closed subspace), there exists a unique particular element $x_p \in V$ of minimum norm and this element x_p is perpendicular to the subspace $N^\perp$ that generates the linear variety V. Thus, $x_p \in N^{\perp\perp}$. Since N is closed (as it is finite dimensional), $N^{\perp\perp} = N$. Thus, $x_p \in N$ and therefore has the form $x_p = \sum_{i=1}^n \beta_i y_i$. We can now solve for the unknowns $\{\beta_i\}$ using the given conditions $< x, y_i >= c_i, i = 1, \ldots, n$. This is the minimum norm (or the least squares solution) for this problem.

Chapter 5: Properties of Dual Spaces

Characterizing the optimal element (error of minimum norm) in a closed subspace from a given element in Hilbert space was done via the projection theorem. This result was possible due to the space having the structure of the inner-product allowing definition of *orthogonal* ($x \perp y \Leftrightarrow < x, y >= 0$) and *aligned* ($| < x, y > | = \|x\|\|y\| \Leftrightarrow x = \alpha y$). In normed spaces, the concept of *orthogonal* and *aligned* are generalized via the dual space of bounded linear functionals. These lead to theorems equivalent to the projection theorem but valid in normed linear spaces.

Section 5.2: Basic Concepts

1. Functional is a mapping from a vector space X to the scalars i. e. $f : X \to \mathbb{R}$.
2. Linear functional in additional follows $f(ax + by) = af(x) + bf(y)$ where $a, b \in \mathbb{R}$ and $x, y \in X$.
3. Continuous functional on normed vector spaces at point $x_0 \in X$. Given favorite error $\delta > 0$, there exists $\epsilon > 0$ such that $\|x - x_0\|_X < \epsilon \Rightarrow |f(x) - f(x_0)|_{\mathbb{R}} < \delta$.
4. If a functional on a normed space X is continuous at a single point, and if it is linear, then it is continuous throughout X.
5. Bounded Functionals: A functional $f : X \to \mathbb{R}$ is bounded if there is constant M such that $|f(x)| \le M\|x\|$ for all $x \in X$.
6. Norm of a bounded linear functional: The smallest constant M for which $|f(x)| \le M\|x\|$ for all $x \in X$ is called the norm of f and denoted by $\|f\|$. Thus,

$$\begin{aligned}
\|f\| &= \inf_M \{M : |f(x)| \le M\|x\|, \ \forall x \in X\} \\
&= \sup_{x \neq \theta} \frac{|f(x)|}{\|x\|} \\
&= \sup_{\|x\|=1} |f(x)| \\
&= \sup_{\|x\|\le 1} |f(x)|
\end{aligned}$$

   and therefore, $|f(x)| \ \le \ \|f\| \ \|x\|$ is also valid.
7. A linear functional on a normed space is bounded if and only if it is continuous. Basically, if the functional is continuous at any one point, then it is also continuous over the whole space, as well as bounded.
8. **Algebraic Dual Space** of Vector Space X. Linear functionals on a space are themselves regarded as elements of a vector space, and for any two functionals say f_1 and f_2, define the addition in the vector space of functionals to be $(f_1 + f_2)(x) \doteq f_1(x) + f_2(x)$. This vector space of linear functionals is called the *algebraic dual* of normed space X.
9. **Normed Dual Space** $X^\star$ of Vector Space X is the subspace of the algebraic dual space and consists of all bounded (i. e. continuous) linear functionals that follow the properties of a vector space (defined as above) as well as having a norm. Thus, for each element (bounded linear functional) $f \in X^\star$, the norm $\|f\|_{X^\star}$ is defined and available.

10. Elements of the normed dual space $X^\star$ are sometimes denoted as $x_1^\star, x_2^\star$ etc. Thus, the value of the linear functional $x_1^\star$ at point $x \in X$ is given by $x_1^\star(x) \in \mathbb{R}$. This is also denoted by the notation $x_1^\star(x) =< x, x_1^\star >$. Note that this is not the inner-product, as the two entries inside the $<,>$ are from different spaces, one from $X^\star$ supplying the linear functional $x_1^\star$, and the other from vector space X supplying the point x on which the linear functional is evaluated.

11. However, there is some similarity in the properties of the notation $x_1^\star(x) =< x, x_1^\star >$ to an inner-product, and therefore, gives good reason to use this notation. For example, (1) it is bi-linear (linear in both arguments) and, (2)$x^\star(x) =< x, x^\star >\leq | < x, x^\star > | \leq \|x^\star\|_{X^\star}\|x\|_X$ which has form similar to the Cauchy-Schwarz inequality.

12. Normed Dual space $X^\star$ is Complete i.e. a Banach Space. In order to verify this, we will need to show that all Cauchy sequences from this space are convergent (i.e. the limit exists and lies in the space). This is equivalent to saying that the limit of a series of bounded linear functionals is also a bounded linear functional. Steps in the proof: (1) Let $\{x_i^\star\}$ be a Cauchy sequence from $X^\star$ i. e. $\|x_n^\star - x_m^\star\| \to 0$ as $m, n \to \infty$. (2) For any $x \in X$, sequence $\{x_n^\star(x)\}$ is a sequence of scalars, and is also Cauchy (since $|x_n^\star(x) - x_m^\star(x)| \leq \|x_n^\star - x_m^\star\|\|x\|$ which goes to zero as $m, n \to \infty$). (3) Since $\mathbb{R}$ is complete, therefore, the Cauchy sequences of scalars converges to element $x^\star(x) \in \mathbb{R}$. (4) Therefore, define a functional $x^\star : X \to \mathbb{R}$ such that $x^\star(x) \doteq \lim_{n\to\infty} x_n^\star(x)$. (5) Show that $x^\star$ is linear (easy). (6) Show $x^\star$ is bounded. Therefore, conclude that $x^\star \in X^\star$, the space of bounded linear functionals, and hence $X^\star$ is complete.

Section 5.3: Duals of some common Banach Spaces

1. Dual space for $X = E^n$ is itself in the sense that $X^\star$ consists of all functionals of the form $f(x = \{\zeta_i\}) = \sum \eta_i\zeta_i$ and the norm of $f \in X^\star$, by the Cauchy Schwarz inequality for finite sequences, is found to be $\|f\|_{X^\star} = \|\eta\|_{\mathbb{R}^n}$.

2. Dual space for l_p is l_q. Dual for l_1 is $l_\infty = \{x = \{\zeta_1, \zeta_2, \ldots\} : \|x\|_{l_\infty} = sup_i|\zeta_i| < \infty\}$.

3. Dual of l_∞ is not l_1. Dual of c_0 is l_1 where c_0 is the space of all infinite sequences of scalars with entries converging to zero i.e. $x \in c_0 \Rightarrow x = \{\zeta_1, \zeta_2, \ldots\}$ such that $\zeta_i \to 0$ as $i \to \infty$. .

4. In any Hilbert space, fixing an element $y \in X$(Hilbert), we can define a functional $f(x) =< x, y >_X$ which is linear, and bounded since from C-S, we can see that $| < x, y >\leq \|x\|_X\|y\|_X$ and since equality holds for $x = y$, we get that $\|f\|_{X^\star} = \|y\|_X$.

5. Riesz-Frechet theorem shows that dual of Hilbert space is itself i. e. not only are functionals defined on Hilbert spaces as above of form $f(x) =< x, y >_X$ linear and bounded, it turns out that **any** bounded and linear functional defined on a Hilbert space must of of this form i.e. for X a Hilbert space and $f \in X^\star$ be given as a bounded linear functional on this space, then there exists a unique representor $y \in X$ for the given bounded linear functional $f \in X^\star$ such that $f(x) =< x, y >_X$, and $\|f\|_{X^\star} = \|y\|_X$. Thus there is a one-one correspondence between the Hilbert space and the space of bounded linear functionals on the Hilbert space.

Section 5.4: Extension of Linear Functionals - Hahn Banach Theorem

1. Extension of linear functional: The Hahn-Banach theorem states that a bounded linear functional f defined on a subspace M of a normed space can be extended to a bounded

linear functional F on the entire space with the norm equal to the norm of f on M.

2. Sublinear functional: A real-valued function p defined on a real vector space is said to be a *sublinear functional* on X if

   1. $p(x_1 + x_2) \leq p(x_1) + p(x_2), \quad \forall x_1, x_2 \in X$, and
   2. $p(\alpha x) = \alpha p(x), \quad \forall \alpha \geq 0, \forall x \in X$.

   Any norm is also a sublinear functional.

3. Hahn-Banach Theorem (existence of extension)
4. Corollary 1: Norm of the extension on whole space equals the norm of the of Bounded Linear Functional being extended (i. e. there exists an extension of minimum norm).
5. Corollary 2: Every element $x \in X$ (NLS) defines a bounded linear functional on X. Let $x \in X$, a normed space. Then, there exists a non-zero bounded linear functional $F : X \to \mathbb{R}$ such that $F(x) = \|F\|\|x\|$.
6. Converse of corollary 2 not true, not every bounded linear functional on X associated to element of X (unlike Hilbert spaces as shown by Riesz-Frechet.).

Section 5.5: Dual Space on the Normed Linear Space (NLS) $C[a, b]$

1. Dual of $C[a, b]$.
2. Riesz Representation Theorem: Dual of (NLS) $C[a, b]$ is the space $BV[a, b]$ of functions of bounded variation. Let f be a bounded linear functional on $X = C[a, b]$. Recall that $\|x\|_{X=C[a,b]} = sup_{a \leq t \leq b}|x(t)|$. Then, there exists a function $v \in BV[a, b]$ such that $\forall x \in X$,

   1. $$f(x) = \int_a^b x(t)dv(t),$$
   2. and, $\|f\| = T.V.(v)$.

   The space $BV[a, b]$ of functions is the space of functions of bounded total variation

   $$T.V.(v) = \int_a^b |dv(t)|$$

   with norm

   $$\|v\| = |v(a)| + T.V.(v).$$

3. Space $NBV[a, b]$.

Section 5.6: Second Dual Space

1. Notation $x^\star(x) =< x, x^\star >$. Note that this is not the inner-product notation as $x, x^\star$ belong to different spaces. However, the notation indicates that it has properties similar to inner-product and can be interpreted as a generalization for normed spaces.
2. Space of *all* bounded linear functional on $X^\star$ is the second dual space, denoted $X^{\star\star}$. Given $x \in X$, $f(x^\star) =< x, x^\star >$ defines a linear functional on $X^\star$.

3. Mapping $\varphi : X \rightarrow X^{\star\star}$ defined by $x^{\star\star} = \varphi(x)$ such that $x^{\star\star}(x^\star) = \varphi(x)(x^\star) \triangleq < x, x^\star >$ is the natural mapping of X into $X^{\star\star}$.
4. Natural mapping is linear and norm-preserving, it is not in general onto.
5. A normed space X is reflexive if the natural mapping $\varphi : X \rightarrow X^{\star\star}$ is onto. If so, write $X = X^{\star\star}$.
6. Let $X = l_p$. Then $X^\star = l_q$, and $X^{\star\star} = l_q^\star = l_p$. Spaces l_p, L_p are reflexive.
7. Any Hilbert space is reflexive.
8. For any normed space X, every element $x \in X$ defines a non-zero, bounded, linear functional $x^\star \in X^\star$ given by $x^\star(x) = \|x^\star\|\|x\|$ (corollary 2 of H-B theorem, pg 112). The converse is true only in reflexive normed spaces, in which every $x^\star \in X^\star$ is associated to an element $x \in X$ such that $x^\star(x) =< x, x^\star >= \|x\|\|x^\star\|$.

Section 5.7: Alignment and orthogonal complements

1. In Hilbert spaces, Cauchy Schwarz inequality $| < x, y > | \leq \|x\|\|y\|$ holds with equality iff $y = \alpha x$ or x, y are *aligned*. For X being a Hilbert space, every $x^\star \in X^\star$ is associated to some $y \in X$ (by Riesz-Frechet) so that $x^\star(x) =< x, x^\star >=< x, y >_X$. Thus, $x^\star$ is said to be aligned with x if its representor $y \in X$ is aligned with x i.e. $y = \alpha x$.
2. Given a normed linear space X and its dual $X^\star$, elements $x \in X$ and $x^\star \in X^\star$ are aligned if $< x, x^\star >= \|x\|\|x^\star\|$. Thus, for normed spaces, alignment is a relation between two distinct vector spaces.
3. If X is Hilbert space, then $< x, y >= 0$ implies $x \perp y$ and since $X = X^\star$ for such spaces (by Riesz-Frechet), then this implies that for some $x^\star \in X^\star$, we get $x^\star(x) =< x, y >= 0$ to give that $x^\star \perp x$.
4. Similarly, if X is a normed space, then $< x, x^\star >= 0$ gives that vectors $x \in X$ and $x^\star \in X^\star$ are orthogonal. This is a generalization of the definition of orthogonal to normed spaces.
5. Orthogonal complement $S^\perp$ of a set S in a normed linear space X is the set of all elements $x^\star \in X^\star$ orthogonal to every vector in S.
6. Given a subset $U \in X^\star$, its orthogonal complement is in $X^{\star\star}$. Since there is a natural mapping of X into $X^{\star\star}$, therefore can think of $U^\perp$ as a subset in X. This is the orthogonal complement ${}^\perp U \subset X$ as the set of all elements in X orthogonal to every element in U.
7. If M is a closed subspace of a normed space X, then ${}^\perp[M^\perp] = M$.

Section 5.8: Minimum norm problems

1. Theorem for characterizing the minimim norm representation analogous to projection theorem in Hilbert Space.
2. Duality Principle: equivalence of problem formulated in normed space and its dual problem in the dual of the normed space.
3. Theorem 1: Let x be element in real normed space X and let d be its distance from subspace M. Then
$$d = \inf_{m \in M} = \max_{x^\star \in M^\perp, \|x^\star\| \leq 1} < x, x^\star >$$

where the maximum on the right is achieved for some $x_0^\star \in M^\perp$. If the infimum on the left is achieved for some $m_0 \in M$, then $x_0^\star$ is aligned with $x - m_0$.

4. Corollary 1: Let x be element in real normed linear space and let M be a subspace of X. A vector $m_0 \in M$ satisfies $\|x - m_0\| \leq \|x - m\|$ for all $m \in M$ if and only if there is a non-zero vector $x^\star \in M^\perp$ aligned with $x - m_0$.
5. Theorem 2: Let M be a subspace in a real normed space X. Let $x^\star \in X^\star$ be a distance d from $M^\perp$. Then,
$$d = \min_{m^\star \in M^\perp} \|x^\star - m^\star\| = \sup_{x \in M, \|x\| \leq 1} < x, x^\star > .$$
6. Minimum norm problems in normed linear spaces must be formulated in the dual space if one is to guarantee the existence of a solution.

Section 5.10: Weak Convergence

1. Understand and appreciate various notions of Convergence - strong convergence (in the norm) and weak (with respect to elements in dual). Strong convergence is related to the norm and intuitively, can think of weak convergence as being convergence with respect to the generalized inner-product in normed spaces.
2. Weak Convergence: sequence $\{x_n\}$ in normed linear space X converges weakly to $x \in X$ if $\forall x^\star \in X^\star$, we have $< x_n, x^\star > \to < x, x^\star >$.
3. Strong convergence ($\|x_n - x\| \to 0$) $\implies$ weak convergence ($< x_n, x^\star > \to < x, x^\star >$) since $| < x_n, x^\star > - < x, x^\star > | = | < x_n - x, x^\star > | \leq \|x^\star\| \|x_n - x\| \to 0$ if $\|x_n - x\| \to 0$.
4. There are sequences that converge weakly but not strongly. As an example, consider $X = X^\star = l_2$ and consider $x_n = e_n = \{0, 0, \ldots, 1, 0, \ldots\}$ with the 1 in the n-th place. Then, for any $x^\star = \{\eta_1, \eta_2, \ldots\} \in X^\star = l_2$, $x^\star(x) = \eta_n$ which, because $\|x^\star\|_{l_2} < \infty$ must tend to 0 as $n \to \infty$. Hence, $x_n \to \theta$ weakly. But $\|x_n\|_{l^2} = 1$, hence $x_n \nrightarrow \theta$ strongly (in the norm).
5. For Hilbert spaces $X^\star = X$, hence weak convergence in Hilbert spaces becomes: sequence $\{x_n\}$ in Hilbert space X converges weakly to $x \in X$ if $\forall x^\star \equiv y \in X^\star = X$, we have $< x_n, y > \to < x, y >$.
6. Define weak convergence of sequence in $X^\star$ in terms of functionals in $X^{\star\star}$.
7. Alternative notion of convergence in $X^\star$ is through elements in X (viewed as embedded in $X^{\star\star}$ by the natural mapping). This is weak-star convergence - sequence $\{x_n^\star\} \to x^\star \in X^\star$ if for every $x \in X$, we get $< x, x_n^\star > \to < x, x^\star >$.
8. There are thus three notions of convergence of sequences: strong (in the norm), weak (through elements in dual space) and weak-star (only for sequences in dual space through elements in the second dual space). Strong $\implies$ weak, and weak $\implies$ weak-star. However, weak-star does not imply weak and weak does not imply strong convergence.
9. Compactness with strong convergence (every infinite sequence has a convergent (strong) subsequence) is severe requirement that is satisfied in few special cases. For compact sets, continuous functionals on a compact set achieve maximum and minimum in the set.
10. Weak and weak-star convergence provide alternative definitions of compactness.
11. Weak-star compact set $K \subset X^\star$: if every infinite sequence from K has a weak-star convergent

subsequence.

12. Similarly, we define weak continuity (for functionals on X) and weak-star continuity (for functionals on $X^\star$).
13. Theorem: A weak-star continuous real-valued functional on a weak-star compact subset S of $X^\star$ is bounded and achieves it maximum on S.

Section 5.11: Hyperplanes and Linear Functionals

1. A hyperplane is a linear vector space X is a maximal proper linear variety i. e. a linear variety H such that $H \neq X$ and if V is any other linear variety containing H, then either $V = H$ or $V = X$.
2. Hyperplanes are level sets of linear functionals.
3. If H is a hyperplane in space X, then there is a linear functional f on X and a constant c such that $H = \{x : f(x) = c\}$. Conversely, if f is a non-zero linear functional on X, the set $\{x : f(x) = c\}$ is a hyperplane in X.
4. For any hyperplane H in space X, there is a unique linear functional such that $H = \{x : f(x) = 1\}$.
5. A hyperplane in normed space X must either be closed or dense in X i. e. either $\bar{H} = H$ or $\bar{H} = X$.
6. Closed hyperplanes in a normed space X correspond to bounded linear functionals on X i. e. there is a unique correspondence between a closed hyperplane and members of dual space $X^\star$, the space of bounded linear functionals on X.

Chapter 6: Linear Operators and Adjoint Operators

Previously, we have looked at *functionals* which are functions whose domain is a linear space and range is the set of scalars $\mathbb{R}$. Now we look at functions whose domain and range are both vector space i.e $A : X \to Y$ where both X and Y are some abstract vector spaces. Such functions are called transformations or operators. The functionals then are a special case of transformations where the range $Y = \mathbb{R}$. As seen in case of minimum norm problems on a normed linear space, the optimal is characterized with respect to the dual space via alignments. This duality is extended to arbitrary spaces by the study of the adjoint operator $A^\star$ for a linear bounded operator A. We also study in this chapter the role of the adjoint operator in connecting, along with operator A, the four fundamental spaces $R(A), N(A), R(A^\star), N(A^\star)$. For Hilbert spaces, since the dual is the Hilbert space itself, the operator and its adjoint act on the same spaces and this unifies the concepts of solving $Ax = y$ when (1) $y \notin R(A)$ and hence there is no solution to the given equation (2) $y \in R(A)$ but the null space $N(A)$ is non-empty leading to multiple solutions and (3) combination of (1) and (2) where neither $y \in R(A)$ nor there are unique solutions. These problems need as appropriate the approach of projecting $y \notin R(A)$ to $R(A)$ so that error is least-norm and the concept of minimum-norm (energy) solution from among the multiple solutions that are available.

Section 6.2: Fundamentals

1. The words *transformation* and *operator* are used interchangeably as well as the symbols T and A. Also, often we write $T(x)$ as Tx.
2. Transformation $T : X \to Y$ where X and Y are vector spaces is a rule that maps an element of space X to some element of space Y.
3. The transformation may be in some cases defined on a subset $D \subseteq X$ of elements in the space X. The set of elements $D \subseteq X$ over which the transformation T is defined is called the *domain* of the transformation.
4. The set of elements in space Y that are connected under the given transformation to elements in space X are called the range of T. Can also write $R(T) = \{y \in Y : y = T(x) = Tx \text{ for some } x \in X\}$
5. Given a set $S \subseteq X$, the *image* of S under T is the subset $T(S) \subseteq Y$ of elements in space Y that are connected to elements inside subset S in space X. Similarly, given a subset $P \subseteq Y$, we get the definition of inverse-image of P as the subset of elements in space X that are mapped under T to set P in space Y. This subset in X is denoted as $T^{-1}(P)$.
6. Linear transformation or operator is one for which $A(x_1 + x_2) = A(x_1) + A(x_2)$.
7. *Range* $R(A)$ of linear operator $A : X \to Y$ is a subspace of Y (show that A must map θ_X to θ_Y and that if $y_1, y_2 \in R(A)$ then, it must be that $y_1 + y_2 \in R(A)$).
8. *Null Space* of a linear operator A is the set of all elements in space X that get mapped to $\theta_Y \in Y$, i. e. $N(A) = \{x \in X : Ax = \theta_Y \in Y\}$.
9. A linear operator on a normed space X is continuous at every point in X if it is continuous at a single point. Linear operator is bounded if and only if it is continuous.
10. Bounded operator is one for which $\|A(x)\|_Y = \|Ax\|_Y \leq M\|x\|_X, \ \forall x \in X$.
11. Vector space of operators can be defined with addition of two operators being defined as $(A_1 + A_2)(x) \doteq A_1(x) + A_2(x)$.

12. For most of these definition, we see they are obvious extensions from the corresponding definitions of functionals. For functionals, the absolute value gave a norm on range $\mathbb{R}$, here instead the norm of space Y is used. Often, the space for the norm is not explicitly indicated but obvious from the context.
13. The norm of $\|A\|$ an operator A is the smallest number M such that $\|Ax\|_Y \leq M\|x\|_X,\ \forall x \in X\}$. Alternatively, it is the largest value of $\|Ax\|_Y/\|x\|_X,\ x \neq \theta_X$. It also follows that these following definitions are equivalent:

$$\begin{aligned}
\|A\| &= \inf_M \{M : \|Ax\|_Y \leq M\|x\|_X,\ \forall x \in X\} \\
&= \sup_{x \neq \theta_X} \frac{\|Ax\|_Y}{\|x\|_X} \\
&= \sup_{\|x\|=1} \|Ax\|_Y
\end{aligned}$$

14. $\|Ax\|_Y \leq \|A\|\|x\|_X$.
15. The space of bounded operators between two vector spaces thus endowed with linear structure and norm lead to the space $B(X, Y)$, the normed space of all bounded linear operators mapping normed space X to normed space Y.
16. $B(X, Y)$ is complete when Y is complete (recall that similar proof for the normal dual space $X^\star$ of bounded linear functionals being complete used the completion of the range $\mathbb{R}$ of bounded linear functionals).
17. Product operators $TS(x) \triangleq T(S(x))$ where $S : X \to Y$, $T : Y \to Z$ and $TS : X \to Z$.
18. Let X, Y, Z be normed spaces and suppose $S \in B(X, Y)$, $T \in B(Y, Z)$. Then, $\|TS\| \leq \|T\|\|S\|$.

Section 6.3: Inverse Operators

1. Conditions on when $Ax = y$ has unique solution x for a given y. The inverse operator is such that $x = A^{-1}y$.
2. If linear operator has an inverse, the inverse is linear since $A^{-1}(Ax + Ay) = x + y = A^{-1}Ax + A^{-1}Ay$.

Section 6.4: Banach Inverse Theorem

1. Given a continuous linear operator mapping normed spaces $A : X \to Y$, then if A has inverse, it is linear. However, the inverse need not be continuous.
2. If X and Y are Banach spaces, the inverse of a continuous linear operator is not only linear but also continuous.

Section 6.5: Adjoints

1. Let $A : X \to Y$ mapping between normed spaces X and Y be a bounded linear operator. The adjoint operator maps between the dual spaces $A^\star : Y^\star \to X^\star$ is defined as the operator $A^\star$ that satisfies $< Ax, y^\star >=< x, A^\star y^\star >$.

2. Fixing a $y^\star \in Y^\star$, we get that $y^\star(Ax)$ is a scalar for each $x \in X$, hence a functional on X. This functional on X is called $x^\star = A^\star y^\star \in X^\star$ and defined to be $x^\star = A^\star y^\star = y^\star A$.
3. Since
$$|x^\star(x)| = |y^\star(Ax)| = | < Ax, y^\star > | \quad \le \quad \|y^\star\|\|Ax\| \le \|y^\star\|\|A\|\|x\|,$$
it follows that the functional $y^\star A$ is bounded and therefore justified to be an element in $X^\star$.
4. Hence, $A^\star : Y^\star \to X^\star$ is given by $A^\star y^\star = y^\star A$ so that $\forall x \in X$, we get $A^\star y^\star(x) = y^\star(Ax)$.
5. The adjoint operator $A^\star$ of operator $A \in B(X,Y)$ is linear.
6. The adjoint operator is bounded.
7. $\|A^\star\| = \|A\|$.
8. If I is the identity operator on normed space X, then $I^\star = I$.
9. If $A_1, A_2 \in B(X,Y)$, then $(A_1 + A_2)^\star = A_1^\star + A_2^\star$ (vector space of adjoint operators).
10. If $A_1 \in B(X,Y)$ and $A_2 \in B(Y,Z)$ then $(A_2A_1)^\star = A_1^\star A_2^\star$.
11. If $A \in B(X,Y)$ has bounded inverse, then $(A^{-1})^\star = (A^\star)^{-1}$.
12. If X and Y are real Hilbert spaces, then they are their own dual, and hence the adjoint $A^\star : Y \to X$ if $A : X \to Y$ and satisfies
$$< x, A^\star y >_X = < Ax, y >_Y$$
13. Self-adjoint operator $A^\star = A$.
14. A self-adjoint linear operator A on Hilbert space H is said to be positive-semidefinite if $\langle x, Ax\rangle \ge 0, \forall x \in H$.

Section 6.6: Relations between Range and NullSpace

1. Let $A \in B(X,Y)$ where X, Y are normed linear spaces. Then, $[R(A)]^\perp = N(A^\star)$.
2. Let $A \in B(X,Y)$ where X, Y are Banach spaces and $R(A)$ is closed. Then, there is a constant K such that for each $y \in R(A)$ there is an x satisfying $Ax = y$ and $\|x\| \le K\|y\|$.
3. Let $A \in B(X,Y)$ where X, Y are Banach spaces and $R(A)$ is closed. Then, $R(A^\star) = [N(A)]^\perp$.
4. Let A be a bounded linear operator acting between two real Hilbert spaces. Then
    1. $[R(A)]^\perp = N(A^\star)$.
    2. $\overline{R(A)} = [N(A^\star)]^\perp$.
    3. $[R(A^\star)]^\perp = N(A)$.
    4. $\overline{R(A^\star)} = [N(A)]^\perp$.

Section 6.8: Geometric Interpretation of Adjoints

1. $A : X \to Y$ and $A^\star : Y^\star \to X^\star$.
2. Let M be a subspace in X. Then, $AM = \{Ax \in Y : x \in M\}$ is a subspace in Y. If $H = x_0 + M$, then the operator A maps it to another linear variety $L = Ax_0 + AM$ in Y.

3. Each hyperplane H and L define unique bounded linear functionals $x_i^\star \in X^\star$ and $y_1^\star \in Y^\star$ through the relations $H = \{x :< x, x_1^\star >= 1\}$ and $L = \{y :< y, y_1^\star >= 1\}$. $A^\star$ maps $y_1^\star$ back to $A^\star y_1^\star \in X^\star$. In fact, $A^\star y_1^\star = x_1^\star$. Hence, the adjoint can be interpreted as mapping the linear variety $L \in Y$ back to the hyperplane $H \in X$.

Section 6.9: Optimization in Hilbert Space

1. Let $A \in B(G, H)$ where G, H are Hilbert spaces. Let fixed $y \notin R(A)$ be given. Find $x \in G$ that minimizes $\|y - Ax\|$.
2. For normed space Y, the orthogonal complement of a set $S \subset Y$ was a subset $S^\perp \in Y^\star$ in the normed dual space $Y^\star$ i. e. $S^\perp = \{y^\star \in Y^\star : y^\star(s) = 0 \ \forall s \in S\}$. Since the adjoint operator $A^\star : Y^\star \to X^\star$, therefore we showed before that in fact for operator $A : X \to Y$, where X, Y are normed spaces, the orthogonal complement of the subspace $R(A)$ was a subspace in the dual space i. e. $[R(A) \subset Y]^\perp = N(A^\star) = \{y^\star \in Y^\star : A^\star y^\star = \theta^\star_{X^\star}\}$.
3. When we consider Hilbert spaces, they are their own dual $X^\star = X, Y^\star = Y$. Hence, $A^\star : Y^\star = Y \to X^\star = X$ and therefore, the orthogonal complement of a subspace $R(A) \subset Y$ is a subspace $N(A^\star) \subset Y = \{y \in Y : A^\star y = \theta_X\}$.
4. Then, for the above optimization problem, $y - Ax \in [R(A)]^\perp = N(A^\star)$ i.e. $A^\star(y - Ax) = A^\star y - A^\star Ax = \theta_X$. Thus, get that such $x \in G$ satisfy $A^\star Ax = A^\star y$.
5. Existence when $R(A)$ closed.
6. Uniqueness when $N(A)$ contains only the null-element (also in this case we say $N(A)$ is "empty") i.e. $A^\star A$ is invertible. Then, $x = (A^\star A)^{-1} A^\star y$.

Section 6.10: Dual Problem

1. Let $A \in B(G, H)$ where G, H are Hilbert spaces. Let fixed $y \in R(A)$ be given, $R(A)$ assumed closed. Let the subspace $N(A)$ also not be empty, hence there are multiple solutions. The vector $x \in G$ of minimum norm solving $Ax = y$ is desired.
2. Since $N(A)$ is non-empty, therefore the set of solutions is a linear variety $V = x + N(A)$, where $x \in R(A^\star)$.
3. The minimum norm solution is, according to modified projection theorem, the element perpendicular to the subspace generating the linear variety. Thus, optimal (minimum norm) solution $x \in R(A^\star)$. Any other solution $x_1 = x + u, u \in N(A)$ will have higher norm $\|x_1\|^2 = \|x\|^2 + \|u\|^2$ and therefore sub-optimal.
4. Hence minimum norm solution $x = A^\star z$ for some $z \in H$.
5. Solve $AA^\star z = y$ first. Then, $x = A^\star z$.
6. When $AA^\star$ is invertible, then, $x = A^\star(AA^\star)^{-1} y$.

Section 6.11: Pseudoinverse operators

1. Let $A \in B(G, H)$ where G, H are Hilbert spaces. Let fixed $y \notin R(A)$ be given, $R(A)$ assumed closed. $N(A)$ is also not empty, hence there are multiple solutions. The vector $x \in G$ of minimum norm solving $Ax = y$ is desired.

2. Among all vectors $x_1 \in G$ satisfying

$$\|Ax_1 - y\| = \min_x \|Ax - y\|,$$

let x_0 be the unique vector of minimum norm. The pseudo-inverse operator $A^\dagger$ of operator A is the operator mapping y into the corresponding x_0 as y varies over H.

3. $N(A)$ is closed, so $G = N(A) \oplus N(A)^\perp$.
4. If $R(A)$ is closed, then $H = R(A) \oplus R(A)^\perp$.
5. Operator A is one-to-one between $N(A)^\perp = R(A^\star)$ and $R(A)$, it is also onto.
6. Between $N(A)^\perp$ and $R(A)$, operator A has a linear inverse, which according to Banach inverse theorem is continuous(bounded).
7. The minimum norm solution $x \in N(A)^\perp$, hence the pseudo-inverse operator has range equal to $N(A)^\perp = R(A^\star)$. Its domain is $R(A)$ and that is extended to be the whole of H by setting $A^\dagger y = 0 \quad \forall y \in R(A)^\perp$.
8. Look at the SVD of matrix A as $A = U\Sigma V^t$ in this context of giving the pseudoinverse.

References

1. (Text for ENSC 801) Optimization by vector space methods, David Luenberger, ISBN 0471018117-X.

2. Introductory Real Analysis, Kolmogorov and Fomin, ISBN 0-486-61226-0.

3. Introductory Functional Analysis with Applications, Erwin Kreyszig, ISBN 0-471-50459-9.

4. Applied Functional Analysis, D. H, Griffel, ISBN 0-486-42258-5.