A SERIES OF CLASS NOTES TO INTRODUCE LINEAR AND NONLINEAR PROBLEMS
TO ENGINEERS, SCIENTISTS, AND APPLIED MATHEMATICIANS

LINEAR CLASS NOTES:
A COLLECTION OF HANDOUTS FOR
REVIEW AND PREVIEW
OF LINEAR THEORY
INCLUDING FUNDAMENTALS OF
LINEAR ALGEBRA

CHAPTER 8

Inner Product and

Normed Linear Spaces

Since solving linear algebraic equations for a field **K** require only a finite number of exact algebraic steps, any field will do. However, actually carrying out the process usually involves approximate arithmetic and hence approximate solutions. In an infinite dimensional vector space the solution process often requires an infinite process that converges. Hence the vector space (or its field) must have additional properties. Physicists and engineers think of vectors as quantities which have **length** and **direction**. Although the notion of direction can be discusssed, the definition of a vector space does not include the concept of the length of a vector. The abstraction of the concept of length of a vector is called a **norm** and vector spaces (also called **linear spaces**) which have a norm (or length) are called **normed linear spaces**. Having the notion of length in a vector space gives us the notion of a **unit vector** (i.e. a vector of length one).

DEFINITION #1. A **normed linear space** is a real or complex vector space V on which a norm has been defined. A norm (or length) is a function $\|\cdot\|: V \to \mathbf{R}^+ = \{\alpha \in \mathbf{R} : \alpha \geq 0\}$ such that

1) $\|\vec{x}\| \geq 0, \quad \forall \vec{x} \in V$

   $\|\vec{x}\| = 0$ if and only if $\vec{x} = \vec{0}$

2) $\|\alpha\vec{x}\| = |\alpha|\|\vec{x}\| \quad \forall \vec{x} \in V \quad \forall$ scalars α

3) $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\| \quad \forall \vec{x}_1, \vec{x}_2 \in V$ (this is called the triangle in equality)

Note that the zero vector $\vec{0}$ is the only vector with zero length. For all other vectors $\vec{x}$ we have $\|\vec{x}\| > 0$. Hence for each non zero $\vec{x}$ we can define the unit vector

$$\vec{u} = \vec{u}(\vec{x}) = \frac{1}{\|\vec{x}\|}\vec{x} = \frac{\vec{x}}{\|\vec{x}\|}. \tag{1}$$

LEMMA #1. If $\vec{x} \in V$, $\|\vec{x}\| \neq 0$ $\in$V, and $\vec{u} = \vec{u}(\vec{x}) = \frac{1}{\|\vec{x}\|}\vec{x} = \frac{\vec{x}}{\|\vec{x}\|}$, then $\|\vec{u}\| = 1$.

Proof. Let $\vec{u} = \vec{u}(\vec{x}) = \frac{1}{\|\vec{x}\|}\vec{x} = \frac{\vec{x}}{\|\vec{x}\|}$. Then

Statement	Reason
$\|\vec{u}\| = \left\|\frac{1}{\|\vec{x}\|}\vec{x}\right\|$	Definition of $\hat{u}$
$= \left\|\frac{1}{\|\vec{x}\|}\right\|\|\vec{x}\|$	Property (2) above
$= \frac{\|\vec{x}\|}{\|\vec{x}\|}$	Algebraic Properties of **R**
$= 1$	Algebraic Properties of **R**.

THEOREM #1. If $\vec{x} \neq \vec{0}$, then $\vec{x}$ can be written as $\vec{x} = \|\vec{x}\|\vec{u}$ where $\vec{u}$ is a unit vector in the direction of $\vec{x}$ and $\|\vec{x}\|$ gives the length of $\vec{x}$.

Proof idea. That $\|\vec{x}\|$ is the length or norm of $\vec{x}$ follows from the definition of $\|\vec{x}\|$ as a norm or length. Since $\vec{x} \neq \vec{0}$, $\|\vec{x}\| \neq 0$ and we define $\vec{u}$ as before as the unit vector $\vec{u} = \vec{u}(\vec{x}) = \frac{1}{\|\vec{x}\|}\vec{x} = \frac{\vec{x}}{\|\vec{x}\|}$ we see that $\vec{u}$ is a positive scalar multiple of $\vec{x}$ so that it is pointed in the same "direction" as $\vec{x}$. To show that $\vec{x} = \|\vec{x}\|\vec{u}$ is left as an exercise. QED

The abstraction of the notion of **distance** between two points in a set (or vector space) is called a **metric**.

DEFINITION #2. A **metric space** is a set S on which a metric has been defined. A **metric** (or distance between) on S is a function $\rho: S\times S \to \mathbf{R}^{+} = \{\alpha \in \mathbf{R} \alpha \geq 0\}$ such that

1) $\rho(x,y) \geq 0$ $\forall$f $x,y \in S$.
   $\rho(x,y) = 0$ if and only if $x=y$.
2) $\rho(x,y) = \rho(y,x)$.
3) $\rho(x,z) \leq \rho(x,y) + \rho(y,z)$ (this is also called the triangle in equality)

THEOREM #2. Every normed vector space is a metric space with the metric $\rho(\vec{x},\vec{y}) = \|\vec{x} - \vec{y}\|$.

However, we note that a metric space need not be a vector space. Geometrically in $\mathbf{R}^3$, ρ is the distance between the tips of the position vectors $\vec{x}$ and $\vec{y}$. A metric yields the notion of a **topology**, but we need not develop this more general concept. However, to discuss approximate solutions, we do need the notion of **completeness**. Although it could be developed in a more general context, we are content to discuss **complete vector spaces**.

DEFINITION #3. A **Cauchy sequence of vectors** is a sequence $\{\vec{x}_n\}_{n=1}^{\infty}$ in a normed linear space V such that for any $\varepsilon>0$, there exists N with $|| \vec{x}_n - \vec{x}_m || < \varepsilon$ whenever $m,n>N$.

Thus $\vec{x}_m$ and $\vec{x}_n$ get close together when n and m are large.

DEFINITION #4. A sequence of vectors $\{\vec{x}_n\}_{n=1}^{\infty}$ in a vector space V is convergent if there exist $\vec{x}$ such that for any $\varepsilon>0$, there exists N with $|| \vec{x}_n - \vec{x} || < \varepsilon$ whenever $n>N$.

DEFINITION #5. A vector space is **complete** if every Cauchy sequence of vectors converges. A **Banach Space** is a complete normed linear space.

The concepts of metric (or topology) and completeness are essential for computing limits; for example, in the process of computing approximate solutions and obtaining error estimates. Completeness does for a vector space what **R** does for **Q** (which are just special cases). It makes sure that there are no holes in the space so that Cauchy sequences (that look like they ought to converge) indeed have a vector to converge to. If we wish to solve problems in a metric space S and S is not complete, we can construct the completion of S which we usually denote by $\overline{S}$. Then, since $\overline{S}$ is complete, we can obtain approximate solutions.

Recall that to determine if two vectors in $\mathbf{R}^3$ are perpendicular, we compute the **dot product**. The abstraction of the notion of dot product in an abstract vector space is called an **inner product**. Vector spaces on which an inner product is defined are called **inner product spaces**. As we will see, in an inner product space we have not only the notion of two vectors being **perpendicular** but also the notions of **length of a vector** and a new way to determine if a set of vectors is **linearly independent**.

DEFINITION #1. An inner product space is a real or complex vector space V on which an inner product is defined. A inner product is a function $(\cdot,\cdot)$: V × V→**R** such that

1) $(\vec{x},\vec{y}) = \overline{(\vec{y},\vec{x})} \quad \forall \vec{x},\vec{y} \in V$

   (the bar over the inner product indicates complex conjugate. If V is a real vector space, it is not necessary and we see that the inner product is commutative for real vector spaces.)

2) $(\alpha\vec{x},\vec{y}) = \alpha(\vec{x},\vec{y}) \quad \forall \vec{x},\vec{y} \in V$ and scalars α.

3) $(\vec{x}_1 + \vec{x}_2,\vec{y}) = (\vec{x}_1,\vec{y}) + (\vec{x}_2,\vec{y}) \quad \forall \vec{x}_1,\vec{x}_2,\vec{y} \in V$.

   (Properties 2 and 3) say that the inner product is linear in the first slot.)

4) $(\vec{x},\vec{x}) \geq 0 \quad \forall \vec{x} \in V$

   $(\vec{x},\vec{x}) = 0 \quad \text{iff } \vec{x} = \vec{0}$

If V is an inner product space we define a mapping, $\|\cdot\| : V \to \mathbf{R}^+ = \{\alpha \in \mathbf{R} : \alpha \geq 0\}$ by

$$\|\vec{x}\| = \sqrt{(\vec{x},\vec{x})}. \qquad (1)$$

THEOREM #1. $\|\vec{x}\| = \sqrt{(\vec{x},\vec{x})}$ given by (1) defines a norm on any inner product space and hence makes it into a normed linear space. (See the previous handout)

DEFINITION #2. A Hilbert space is a complete, inner product space.

Again, the concept of completeness in a vector space is an abstraction of what **R** does for **Q**. **R** is complete; **Q** is not.

EXAMPLE (THEOREM). Let V = $\mathbf{R}^n$ and define the inner product by $(\vec{x},\vec{y}) = \vec{x}^T\vec{y} = \sum_{i=1}^{n} x_i y_i$ where $\vec{x}^T = [x_1,\ldots,x_n]^T$. Note that we can define the inner product in $\mathbf{R}^n$ in terms of matrix multiplication. Note also $(\vec{x},\vec{y}) = \vec{y}^T\vec{x}$. We can then prove (i.e. verify) that $(\vec{x},\vec{y}) = \vec{x}^T\vec{y}$ defines an inner product (i.e. satisfies the properties in the definition of an inner product).

Proof of 2). Let $\vec{x}^T = [x_1,...,x_n]^T$, $\vec{y}^T = [y_1,...,y_n]^T$ then

Statement	Reason
$(\alpha\vec{x},\vec{y}) = (\alpha\vec{x})^T\vec{y}$	definition of inner product for $\mathbf{R}^n$.
$= (\alpha[x_1,...,x_n])\begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}$	notation and definition of transpose
$= [\alpha x_1,...,\alpha x_n]\begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}$	definition of scalar multiplication
$= (\alpha x_1)y_1 + (\alpha x_2)y_2 +\cdots+ (\alpha x_n)y_n$	definition matrix multiplication
$= \alpha(x_1y_1 + x_2y_2 +\cdots+ x_ny_n)$	properties real numbers
$= \alpha(\vec{x}^T\vec{y})$	definition of matrix multiplication
$= \alpha(\vec{x},\vec{y})$	definition of inner product

QED

Note that in $\mathbf{R}^n$ we have

$$\|\vec{x}\| = (\vec{x},\vec{x}) = \vec{x}^T\vec{x} = x_1^{\,2} + \cdots + x_n^{\,2} = \sum_{i=1}^{n} x_i^{\,2} \ . \tag{2}$$

In $\mathbf{R}^3$ we know that two vectors are **perpendicular** if their dot product is zero. We abstract this idea by defining two vectors in an inner product space to be **orthogonal** (rather than use the word perpendicular) if their inner product is zero.

DEFINITION #3. The row vectors $\vec{x} = [x_1,x_2,...,x_n]$ and $\vec{y} = [y_1,y_2,...,y_n]$ in $\mathbf{R}^n$ are **orthogonal** if (and only if) $(\vec{x},\vec{y}) = \vec{x}^T\vec{y} = 0$.

Pythagoras Extended. In $\mathbf{R}^3$ (or any real inner product space) we might define two vectors to be perpendicular if they satisfy the Pythagorean theorem; that is if

$$\|x\|^2 + \|y\|^2 = \|x - y\|^2 (= \rho(\vec{x},\vec{y})) \ . \tag{3}$$

Since $\vec{x} = \vec{y} + (\vec{x} - \vec{y})$, (3) may be rewritten as

$$(\vec{x},\vec{x}) + (\vec{y},\vec{y}) = (\vec{x}-\vec{y},\vec{x}-\vec{y})$$
$$= (\vec{x},\vec{x}) - 2(\vec{x},\vec{y}) + (\vec{y},\vec{y}) \qquad (4)$$

(since the inner product is commutative; $\mathbf{C}^n$ is different.)

<u>THEOREM #2</u>. $\vec{x}$ and $\vec{y}$ in $\mathbf{R}^n$ are perpendicular iff they are orthogonal.

<u>DEFINITION #4</u>. $\hat{u} = \frac{\vec{x}}{\|\vec{x}\|}$ is a **unit vector** (i.e. a vector of length one) in the direction of the nonzero vector $\vec{x}$ ($\vec{0}$ has no direction). Hence any nonzero vector $\vec{x}$ can be written as $\vec{x} = \|\vec{x}\|\frac{\vec{x}}{\|\vec{x}\|}$ where $\|\vec{x}\|$ is the magnitude or length and $\hat{u} = \frac{\vec{x}}{\|\vec{x}\|}$ is a unit vector in the direction of $\vec{x}$.

<u>DEFINITION #5</u>. The <u>cosine of the **acute angle** θ</u> ($0 \le \theta \le \pi$) between two nonzero vectors $\vec{x}$ and $\vec{y} \in \mathbf{R}^n$ is

$$\cos(\theta) = \frac{(\vec{x},\vec{y})}{\|\vec{x}\|\|\vec{y}\|} = \frac{\vec{x}^T\vec{y}}{\|\vec{x}\|\|\vec{y}\|} \qquad (5)$$

Note: This is often used as the geometric definition of dot product in $\mathbf{R}^3$.
To show that (5) does yield θ in $\mathbf{R}^3$ we first extend the concept of **projection**.

<u>DEFINITION #6</u>. The **vector projection** of a vector $\vec{b}$ in the direction of a non-zero vector $\vec{a}$ is given by

$$p = \frac{(\vec{a},\vec{b})}{\|\vec{a}\|} = \|\vec{b}\|\cos\theta$$

scalar

$$\vec{p} = \left(\|\vec{b}\|\cos\theta\right)\left(\frac{\vec{a}}{\|\vec{a}\|}\right)$$

magnitude of $\vec{p}$ — unit vector giving direction of $\vec{p}$

$$= \frac{(\vec{a},\vec{b})}{\|\vec{a}\|}\frac{\vec{a}}{\|\vec{a}\|} = \frac{(\vec{a},\vec{b})}{\|\vec{a}\|^2}\vec{a}$$

The magnitude of $\vec{p}$ is called the **scalar projection** of $\vec{b}$ in the direction of $\vec{a}$

We first review the definition, a theorem, and the test for a subspace.

DEFINITION #1. Let W be a nonempty subset of a vector space V. If for any vectors $\vec{x}$, $\vec{y} \in$ W and scalars $\alpha,\beta \in$ **K** (recall that normally the set of scalars **K** is either **R** or **C**), we have that $\alpha\,\vec{x} + \beta\,\vec{y} \in$ W, then W is a **subspace** of V.

THEOREM #1. A nonempty subset W of a vector space V is a subspace of V if and only if for $\vec{x}$, $\vec{y} \in$ V and $\alpha \in$ **K** (i.e. α is a scalar) we have.

i) $\vec{x}$, $\vec{y} \in$ W implies $\vec{x} + \vec{y} \in$ W, and

ii) $\vec{x} \in$ W implies $\alpha\,\vec{x} \in$ W.

TEST FOR A SUBSPACE. Theorem #1 gives a good test to determine if a given subset of a vector space is a subspace since we can test the closure properties separately. Thus if $W \subset V$ where V is a vector space, to determine if W is a subspace,we check the following three points.

1) Check to be sure that W is nonempty. (We usually look for the zero vector since if there is $\vec{x} \in$W, then $0\,\vec{x} = \vec{0}$ must be in W. Every vector space and every subspace must contain the zero vector.)
2) Let $\vec{x}$ and $\vec{y}$ be arbitrary elements of W and check to see if $\vec{x} + \vec{y}$ is in W. (Closure of vector addition)
3) Let $\vec{x}$ be an arbitrary element in W and check to see if $\alpha\,\vec{x}$ is in W. (Closure of scalar multiplication).

DEFINITION #2 Let W_1 and W_2 be subspaces of a vector space V. Then the **sum of W_1 and W_2** is defined as $W_1 + W_2 = \{\vec{x}_1 + \vec{x}_2 : \vec{x}_1 \in W_1 \text{ and } \vec{x}_2 \in W_2\}$.

THEOREM #2 Let W_1 and W_2 be subspaces of a vector space V. Then the sum of W_1 and W_2 is subspace of V.

DEFINITION #3 Let W_1 and W_2 be subspaces of a vector space V such that $W_1 \cap W_2 = \{\vec{0}\}$. Then the sum of W_1 and W_2 defined as $W_1 + W_2 = \{\vec{x}_1 + \vec{x}_2 :$ $\vec{x}_1 \in W_1 \text{ and } \vec{x}_2 \in W_2\}$ is a **direct sum** which we denote by $W_1 \oplus W_2$

THEOREM #3 Let W_1 and W_2 be subspaces of a vector space V and $V = W_1 \oplus W_2 = \{\vec{x}_1 + \vec{x}_2 : \vec{x}_1 \in W_1 \text{ and } \vec{x}_2 \in W_2\}$. Then for every vector $\vec{x}$ in V there exist unique vectors $\vec{x}_1 \in W_1$ and $\vec{x}_2 \in W_2$ such that $\vec{x} = \vec{x}_1 + \vec{x}_2$.

DEFINITION #4 Let W_1 and W_2 be subspaces of a real inner product space V with inner product $(\cdot,\cdot)$. W_1 and W_2 are said to be **orthogonal** to each other if $\forall$ + $\vec{y}$ x $\in W_1$ and $\forall$ x + $\vec{y}$ $\in W_2$, we have $(\vec{x}, \vec{y}) = 0$. We write $W_1 {}_2 W_2$.

THEOREM #4. If $A \in \mathbf{R}^{m\times n}$, then its row space is orthogonal to its null space and its column space is orthogonal to its left null space. We write $R(A^T) {}_2 N(A)$ and $R(A) {}_2 N(A^T)$.

DEFINITION #5 Let W be a subspace of a real inner product space V with inner product $(\cdot,\cdot)$. The **orthogonal complement** of W is the set $W^2 = \{ \vec{y} \in V: (\vec{x}, \vec{y}) = 0 \ \forall \ \vec{x} \in W\}$.

THEOREM #5. Let W be a subspace of a real inner product space V with inner product $(\cdot,\cdot)$. Then the orthogonal complement of W, $W^2 = \{ \vec{y} \in V: (\vec{x}, \vec{y}) = 0 \ \forall \ \vec{x} \in W\}$, is a subspace.

THEOREM #6. Let W be a subspace of a real inner product space V with inner product $(\cdot,\cdot)$. Then the orthogonal complement of W^2 is W. We write $(W^2)^2 = W$.

THEOREM #7. If $A \in \mathbf{R}^{m\times n}$, then its row space is the orthogonal complement of its null space and the null is the orthogonal complement to the row space. We write $R(A^T)^2 = N(A)$ and $N(A)^2 = R(A^T)$. Similarly, $R(A)^2 = N(A^T)$ and $N(A^T)^2 = R(A)$.

THEOREM #8. Let W be a nonempty subset of a real inner product space V with inner product $(\cdot,\cdot)$. Then V is the direct sum of W and W^2, $V = W \oplus W^2$.

THEOREM #9. If $A \in \mathbf{R}^{m\times n}$, then $\mathbf{R}^n = R(A^T) \oplus N(A)$ and $\mathbf{R}^m = R(A) \oplus N(A^T)$.

In vector spaces where the concept of length or **norm** (as well as direction) is available, we can talk about approximate solutions to the mapping problem

$$T(\vec{u}) = \vec{b} \tag{1}$$

where T is a linear operator from V to W; $T:V\to W$ where V and W are normed linear spaces. Let $\vec{u}_a \in V$ be an **approximate solution** of (1) and $\vec{u}_e \in V$ be the **exact solution** which we assume to exist and be unique. A measure of how good a solution $\vec{u}_a$ is is given by the **norm** or length of the **error vector** in V,

$$\vec{E}_v = \vec{u}_e - \vec{u}_a\,;$$

that is,

$$E_V = \|\vec{E}_v\| = \|\vec{u}_e - \vec{u}_a\|\,.$$

If T is invertible (e.g., if $T:\mathbf{R}^n\to\mathbf{R}^n$, is defined by a matrix, $T(\vec{x}) = A\vec{x}$ and $\det A \neq 0$), then

$$E_V = \|\vec{E}_v\| = \|\vec{u}_e - \vec{u}_a\| = \|T^{-1}(\vec{b}) - T^{-1}(\vec{b}_a)\| = \|T^{-1}(\vec{b} - \vec{b}_a)\|$$

where $\vec{b}_a = T(\vec{u}_a)$. (The inverse of a linear operator, if it exists, is a linear operator.) By a well-known theorem in analysis,

$$\|T^{-1}(\vec{b} - \vec{b}_a)\| \le \|T^{-1}\|\,\|\vec{b} - \vec{b}_a\|$$

where $||T^{-1}||$ is the **norm of the operator** T^{-1} which we assume to be finite. If an a priori "estimate" (i.e., bound) of $||T^{-1}||$, say $||T^{-1}|| \le C$, can be obtained, then an "estimate of" (i.e., bound for) E_V can be obtained by first computing $\|\vec{b} - \vec{b}_a\|$. Even without an estimate of (bound for) $||T^{-1}||$, we may use

$$E = E_W = \|\vec{b} - \vec{b}_a\| = \|\vec{b} - T(\vec{u}_a)\|. \quad ||b - b_a|| = ||b - T(u_a)||$$

where

$$\vec{E}_W = \vec{b} - \vec{b}_a = \vec{b} - T(\vec{u}_a)\,;$$

is a measure of the error for $\vec{u}_a$. After all, if $\vec{u}_a = \vec{u}_e$, then $T(\vec{u}_a) = \vec{b}$ so that $E_W = 0$. We call E_W the **error vector** in W. Note that this error vector (and hence E) can always be computed whereas $\vec{E}_V$ usually can not. (If $\vec{E}_V$ is known, then the exact solution $\vec{u}_e = \vec{u}_a + \vec{E}_V$ is known and there is no need for an approximate solution.) In fact, E can be computed independent of whether (1) has a unique solution or not. We refer to a solution that minimizes $E = E_W$ as a **least error solution**. If (1) has one or more solutions, then these are all least error solutions since they all give E = 0. On the other hand, if (1) has no solution, then choosing $\vec{u}$ to minimize E gives a "best possible" solution. Under certain conditions, it is unique.

<u>DEFINITION #1</u>. Let S = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_k$} ⊆ W ⊆ V where W is a subspace of the inner product space V. Then S is said to be (pairwise) orthogonal if for all i,j we have { $\vec{x}_i$, $\vec{x}_j$) = 0 for i≠j. If S is a basis of W, it is called an orthogonal basis. (This requires that S does not contain the zero vector.) An orthogonal basis is said o be othonormal if for alli, $\|\vec{x}_i\| \neq 0$

If B = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_k$} ⊆V is an orthogonal basis for the inner product space V, then the coordinates for $\vec{x} \in V$ are particularly easy to compute. Let

$$\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_k\vec{x}_k \in V \quad . \tag{1}$$

To find c_i, take the inner product of both sides with $\vec{x}_j$.

$$\begin{aligned}
(\vec{x}_j,\vec{x}) &= (\vec{x}_j, c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_k\vec{x}_k) \ . \\
(\vec{x}_j,\vec{x}) &= (\vec{x}_j, c_1\vec{x}_1) + (\vec{x}_j, c_2\vec{x}_2) + \cdots + (\vec{x}_j, c_k\vec{x}_k) \\
(\vec{x}_j,\vec{x}) &= c_1(\vec{x}_j,\vec{x}_1) + c_2(\vec{x}_j,\vec{x}_2) + \cdots + c_k(\vec{x}_j,\vec{x}_k) \\
(\vec{x}_j,\vec{x}) &= c_j(\vec{x}_j,\vec{x}_j)
\end{aligned} \tag{2}$$

so that

$$c_j = \frac{(\vec{x}_j,\vec{x})}{(\vec{x}_j,\vec{x}_j)} \ . \tag{3}$$

The concepts of a basis and orthogonal basis can be extended to infinite dimensional spaces. We first extend the concepts of linear independence and spanning sets.

<u>DEFINITION #2</u>. An infinite set S in a vector space V is linearly independent if every finite subset of S is linearly independent. Thus the countable set { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_k$,...} is linearly independent if Sn = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_k$} is linearly independent for all n∈**N**.

<u>DEFINITION #3</u>. Let the countable set S = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_k$,...} ⊆ W ⊆ V where W is a subspace of a vector space V. S is a Hamel spanning set for W if for all $\vec{x} \in W$, there exists n∈**N** and c_1, c_2, ..., c_n such that $\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n$. If V is a topological vector space (e.g. a normed linear space), then S is a Schauder spanning set for W if for all $\vec{x} \in W$, there exist c_1, c_2, ..., c_n,... such that $\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n + \cdots = \sum_{n=1}^{\infty} c_n\vec{x}_n$.

DEFINITION #3. Let B = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_n$ } ⊆ W ⊆ V where W is a subspace of a vector space V. B is a Hamel basis for W if it is linearly independent and a Hamel spanning set for W. If V is a topological vector space (e.g. a normed linear space), then B is a Schauder basis for W if it is linearly independent and a Schauder spanning set for W..

EXAMPLE. 1) Let B = {1, x, x^2, x^3, ...}. Then B is a Hamel basis for the set of all polynomials P (R,R) and a Scauder basis for the set of all analytic functions with an infinite radius of convergence about x = 0. Note that both of these spaces are infinite dimensional.
2) Let B = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_n$,...} be an infinite linearly independent set in a real topological vector space. Then B is a Hamel basis for the subspace W_1 = { $\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n$: c_1, c_2, ,...$c_n \in \mathbf{R}$} of V and B is a Schauder basis for the subspace W_2 = { $\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n +$: c_1, c_2, ,...c_n,...$\in \mathbf{R}$ where the series converges} of V. Again, note that both of these spaces are infinite dimensional.

If B = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_n$,...} ⊆V is an orthogonal basis for the Hilbert space H, then the coordinates for $\vec{x} \in H$ are particularly easy to compute. Let

$$\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n + \cdots \in H \quad . \tag{4}$$

To find c_i, take the inner product of both sides with $\vec{x}_j$.

$$(\vec{x}_j,\vec{x}) = (\vec{x}_j, c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n + \cdots) \, .$$
$$(\vec{x}_j,\vec{x}) = (\vec{x}_j, c_1\vec{x}_1) + (\vec{x}_j, c_2\vec{x}_2) + \cdots + (\vec{x}_j, c_n\vec{x}_n) + \cdots$$
$$(\vec{x}_j,\vec{x}) = c_1(\vec{x}_j,\vec{x}_1) + c_2(\vec{x}_j,\vec{x}_2) + \cdots + c_n(\vec{x}_j,\vec{x}_n) + \cdots$$
$$(\vec{x}_j,\vec{x}) = c_j(\vec{x}_j,\vec{x}_j) \tag{5}$$

$$c_j = \frac{(\vec{x}_j,\vec{x})}{(\vec{x}_j,\vec{x}_j)} \, . \tag{6}$$

Note that this is the same formula as for the finite dimensional case.