Math 333 Linear Algebra
Supplementary Lecture Notes

Mark Pernarowski

September 12, 2005

Contents

1 Vector Spaces

Definition 1 *Let V be a nonempty set on which the operations of addition $+$ and scalar multiplication have been defined:*

(i)	$\mathbf{u}+\mathbf{v}$	*is defined* $\forall \mathbf{u}, \mathbf{v} \in V$
(ii)	$c\mathbf{u}$	*is defined* $\forall \mathbf{u} \in V, \forall c \in \mathbb{R}$.

The set V is called a vector space if additionally, $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $\forall b, c \in \mathbb{R}$ the following axioms hold:

(A1)	$\mathbf{u}+\mathbf{v} \in V$	*V closed under addition*
(A2)	$\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$	*addition is commutative*
(A3)	$\mathbf{u}+(\mathbf{v}+\mathbf{w}) = (\mathbf{u}+\mathbf{v})+\mathbf{w}$	*addition is associative*
(A4)	$\exists \mathbf{0} \in V \, such\ that\ \mathbf{u}+\mathbf{0} = \mathbf{u}$	*existence of a zero vector*
(A5)	$\exists -\mathbf{u} \in V\ such\ that\ \mathbf{u}+(-\mathbf{u}) = \mathbf{0}$	*existence of a negative element*
(A6)	$c\mathbf{u} \in V$	*closed under scalar multiplication*
(A7)	$c(\mathbf{u}+\mathbf{v}) = c\mathbf{u}+c\mathbf{v}$	*distributive property I*
(A8)	$(b+c)\mathbf{u} = b\mathbf{u}+c\mathbf{u}$	*distributive property II*
(A9)	$c(\beta\mathbf{u}) = (c\beta)\mathbf{u}$	*commutativity of scalar multiplication*
(A10)	$1\mathbf{u} = \mathbf{u}$	*scalar multiplication identity element*

Sometimes the symbols $\oplus$ and $\odot$ will be used to denote vector addition and scalar multiplication, respectively. This will be done when vector addition and scalar multiplication is defined in a nontraditional way such as in the example below. Then $\mathbf{u}+\mathbf{v}$ will be written $\mathbf{u} \oplus \mathbf{v}$ and $c\mathbf{u}$ will be written $c \odot \mathbf{u}$.

Example 1 *: Let*

$$V = \{\mathbf{u} : \mathbf{u} = (u_1, u_2) \in \mathbb{R}^2\}$$

and

$$\begin{aligned} \mathbf{u} \oplus \mathbf{v} &\equiv (u_1+v_1+1, u_2+v_2+1) \\ c \odot \mathbf{u} &\equiv c\mathbf{u} = (cu_1, cu_2) \end{aligned}$$

It is easy to show axioms (A1)-(A3) are satisfied. For instance

$$\mathbf{u} \oplus (\mathbf{v} \oplus \mathbf{w}) = (\mathbf{u} \oplus \mathbf{v}) \oplus \mathbf{w} = (u_1+v_1+w_1+2, u_2+v_2+w_2+2)$$

At first glance it might seem (A4) is not satisfied. However, if one defines $\mathbf{0} = (-1,-1)$, we have $\mathbf{0} \oplus \mathbf{u} = \mathbf{u}$ for all $\mathbf{u}$ including $\mathbf{u} = \mathbf{0}$. For this definition of $\mathbf{0}$ assumption (A5) is then satisfied if one defines

$$-\mathbf{u} \equiv (-u_1-2, -u_2-2)$$

since then

$$\mathbf{u} \oplus (-\mathbf{u}) = (-1,-1) \quad , \quad \forall \mathbf{u} \, .$$

Assumption (A6) is trivially satisified and it is easy to verify (A9)-(A10) are as well. However, (A7) is not satisfied since

$$\begin{aligned} c(\mathbf{u} \oplus \mathbf{v}) &= (cu_1+cv_1+c, cu_2+cv_2+c) \\ c\mathbf{u} \oplus c\mathbf{v} &= (cu_1+cv_1+1, cu_2+cv_2+1) \end{aligned}$$

implies $c(\mathbf{u} \oplus \mathbf{v}) \neq c\mathbf{u} \oplus c\mathbf{v}$ *for all* c*. They are equal for* $c = 1$ *but not* *for all* c*.*

Similarly, (A8) is not satisfied since

$$\begin{aligned} (b+c)\mathbf{u} &= (bu_1 + cu_1, bu_2 + cu_2) \\ b\mathbf{u} \oplus c\mathbf{u} &= (bu_1 + cu_1 + 1, bu_2 + cu_2 + 1) \end{aligned}$$

So in summary, all but axioms (A7) and (A8) are satisfied and one concludes that V *is not a vector space. Of course with the usual (Euclidean) definitions of addition and scalar multiplication* $\mathbb{R}^2$ *is a vector space.*

Example 2 *Let* V *be the set of (possibly empty) sets of real numbers. Thus, an "element" of* V *is a set of real numbers. Then we define vector addition to be set union:*

$$\mathbf{u} \oplus \mathbf{v} \equiv \mathbf{u} \cup \mathbf{v}$$

For example, if $\mathbf{u} = \{1, 2, 3\}$ *and* $\mathbf{v} = \{3, 4\}$ *we have* $\mathbf{u} \oplus \mathbf{v} = \{1, 2, 3, 4\}$*. Next we define scalar multiplication as the identity operation. That is, for any real number* c *and vector* $\mathbf{u} \in V$

$$c \odot \mathbf{u} \equiv \mathbf{u} \, .$$

Then (A1) is satisified. With the stated definitions the question one must ask is if a union of two sets of real numbers $\mathbf{u}$ *and* $\mathbf{v}$ *is a set of real numbers. Clearly it is so (A1) is satisfied. Furthermore, set union is commutative so (A2) and (A3) are also satisfied. Oddly enough, (A4) is satisfied if one identifies the zero vector* $\mathbf{0}$ *as the empty set, i.e.,* $\mathbf{0} = \{ \}$*. Because of how scalar multiplication was defined (A6)-(A10) are also all satisfied. The only axiom which fails is (A5). There is no set "*$-\mathbf{u}$*" which when unioned with* $\mathbf{u}$ *yields the empty set* $\mathbf{0}$*. Note that in this case, "real numbers" could have been replaced with anything, i.e. bank accounts, gliorps, cats...*

Some Common Vector spaces:

$\mathbb{R}^n$	the set of all ordered n-tuples of real numbers
$M_{mn} = \mathbb{R}^{m \times n}$	the set of all real m by n matrices
P_n	the set of all n-th degree polynomials
$C(\mathbb{R})$	the set of all continuous functions on $\mathbb{R}$
$C^n(\mathbb{R})$	the set of all functions on $\mathbb{R}$ with n continous derivatives
$C^\infty(\mathbb{R})$	the set of all functions on $\mathbb{R}$ with continuous derivatives of all orders
$F(\mathbb{R})$	the set of all function defined on $\mathbb{R}$

Note that the function spaces are subsets:

$$P_n \subset C(\mathbb{R}) \subset C^1(\mathbb{R}) \subset C^2(\mathbb{R}) \subset \cdots C^\infty(\mathbb{R}) \subset F(\mathbb{R})$$

2 Basic Definitions:

In all of the following V is a vector space:

Definition 2 *W is a subspace of V if*

a)	$W \subset V$	*(subset)*
b)	$\mathbf{u}, \mathbf{v} \in W \Rightarrow \mathbf{u} + \mathbf{v} \in W$	*(closure under addition)*
c)	$\mathbf{u} \in W, c \in \mathbb{R} \Rightarrow c\mathbf{u} \in W$	*(closure under scalar addition)*

This theorem implies W is also a vector space (see text).

Definition 3 *$\mathbf{w} \in V$ is a linear combination of $\mathbf{v}_1, \ldots \mathbf{v}_n \in V$ if $\exists c_k \in \mathbb{R}$ such that*

$$\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots c_n\mathbf{v}_n$$

Definition 4 *Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n\} \subset V$.*

$$span(S) \equiv \left\{ \mathbf{w} \in V : \mathbf{w} = \sum_{k=1}^{n} c_k \mathbf{v}_k \quad for\ some\ c_k \in \mathbb{R} \right\}$$

In words, $W = span(S)$ is the set of all linear combinations of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n$. Note that W is a subspace of V.

Definition 5 *A set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n\} \subset V$ is linearly independent if*

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots c_n\mathbf{v}_n = 0 \quad \Rightarrow \quad c_k = 0 \quad , \quad \forall k = 1, \ldots n.$$

If S is not linearly independent S is said to be linearly dependent.

If S is (linearly) dependent then at least one vector $\mathbf{v} \in S$ is a linear combination of the remaining vectors.

Definition 6 *A set $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n\} \subset V$ is basis for V if*

a) E *is linearly independent*
b) $V = span(E)$

By a theorem, if $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n\}$ is a basis for V then for every $\mathbf{v} \in V$ there are unique scalars $c_1, \ldots c_n$ such that

$$\mathbf{v} = c_1\mathbf{e}_1 + \cdots c_n\mathbf{e}_n$$

Moreover, if

$$\mathbf{w} = b_1\mathbf{e}_1 + \cdots b_n\mathbf{e}_n$$

then

$$\mathbf{v} \neq \mathbf{w} \Leftrightarrow (c_1, \ldots, c_n) \neq (b_1, \ldots, b_n)$$

This permitts the following definition.

Definition 7 *The* *coordinate* $(\mathbf{v})_E$ *of* $\mathbf{v} \in V$ *relative to the basis* $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n\}$ *is that unique* $\mathbf{c} = (c_1, \ldots c_n) \in \mathbb{R}^n$ *such that* $\mathbf{v} = c_1\mathbf{e}_1 + \cdots c_n\mathbf{e}_n$, *i.e.,*

$$\mathbf{c} = (\mathbf{v})_E \quad \Rightarrow \quad \mathbf{v} = c_1\mathbf{e}_1 + \cdots c_n\mathbf{e}_n$$

Definition 8 *If* $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n\}$ *is a basis for* V *and* $1 \leq n < \infty$ *then* V *is said to be* *finite dimensional* *with* *dimension*

$$dim(V) = n$$

If $V = \{0\}$ *then* $dim(V) = 0$.

3 Basic Theorems for spanning, dependence and bases:

Theorem 1 *Let V be a vector space with $dim(V) = n < \infty$, having basis*

$$E = \{\mathbf{e}_1, \ldots \mathbf{e}_n\},$$

W be any subspace of V and let

$$S = \{\mathbf{v}_1, \ldots \mathbf{v}_k\} \subset V$$

be a finite collection of k vectors. Further define the set of coordinate vectors:

$$S_E = \{(\mathbf{v}_1)_E, \ldots (\mathbf{v}_k)_E\} \subset \mathbb{R}^n\ .$$

Then,

S dependent	$\Leftrightarrow$	$\exists \mathbf{v} \in S$ *such that* $\mathbf{v} \in span(S - \{\mathbf{v}\})$.
$k > n$	$\Rightarrow$	*S dependent*
$k < n$	$\Rightarrow$	*S does not span V*
$\mathbf{v} \notin span(S)$ *and S independent*	$\Rightarrow$	$S^+ \equiv S \cup \{\mathbf{v}\}$ *independent*
$\mathbf{v} \in span(S^-) \equiv span(S - \{\mathbf{v}\})$	$\Rightarrow$	$span(S) = span(S^-)$
$V = span(S)$	$\Rightarrow$	$\exists S^- \subset S$, S^- *a basis for V*
$V = span(S)$ *and* $k = n$	$\Rightarrow$	*S a basis for V*
S independent and $k = n$	$\Rightarrow$	*S a basis for V*
$dim(W) \leq dim(V)$		
$dim(W) = dim(V)$	$\Rightarrow$	$V = W$
S independent in V	$\Leftrightarrow$	*S_E independent in $\mathbb{R}^n$*
$V = span(S)$ *and* $k = n$	$\Leftrightarrow$	$\mathbb{R}^n = span(S_E)$

4 Matrices and their Subspaces:

In the following A, $B \in \mathbb{R}^{m\times n}$ are matrices, $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y}, \mathbf{b} \in \mathbb{R}^m$. We shall define $\mathbf{r}_i$ to be the row vectors of A and $\mathbf{c}_j$ to be the column vectors so that

$$A = [a_{ij}] = \begin{bmatrix} \cdots\cdots\mathbf{r}_1\cdots\cdots \\ \cdots\cdots\mathbf{r}_2\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\mathbf{r}_m\cdots\cdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ \mathbf{c}_1 & \mathbf{c}_2 & \vdots & \mathbf{c}_n \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

For any matrix, its transpose A^T is defined by

$$A^T = [a_{ji}]$$

Important properties of the transpose are

$$\begin{aligned} (A+B)^T &= A^T + B^T \\ (AB)^T &= B^T A^T \end{aligned}$$

For square matrices $A, B \in \mathbb{R}^{n\times n}$ having inverses A^{-1} and B^{-1}, respectively,

$$\begin{aligned} (AB)^{-1} &= B^{-1}A{-}1 \\ (A^{-1})^T &= (A^T)^{-1} \end{aligned}$$

A simple proof of the latter can be seen from the calculations:

$$\begin{aligned} \mathbf{x} &= A^{-1}\mathbf{b} \\ \mathbf{x}^T &= \mathbf{b}^T(A^{-1})^T \\ \mathbf{x}^T &= \mathbf{x}^T A^T (A^{-1})^T \quad , \quad \forall \mathbf{x} \\ I &= A^T(A^{-1})^T \quad . \end{aligned}$$

Also, for any matrix one can define the four fundamental subspaces:

Definition 9 *The four fundamantal subspaces of A are*

$$\begin{aligned} row(A) &\equiv span\{\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_m\} \subset \mathbb{R}^n \\ col(A) &\equiv span\{\mathbf{c}_1, \mathbf{c}_2, \ldots \mathbf{c}_m\} \subset \mathbb{R}^m \\ N(A) &\equiv \{\mathbf{x} : A\mathbf{x} = 0\} \subset \mathbb{R}^n \\ N(A^T) &\equiv \{\mathbf{y} : A^T\mathbf{y} = 0\} \subset \mathbb{R}^m \end{aligned}$$

Note that $row(A^T)$ and $col(A^T)$ have not been included since for every $A \in \mathbb{R}^{m\times n}$,

$$col(A) = row(A^T) \ .$$

Bases for $row(A)$, $col(A)$ and $N(A)$ can all be found by row reducing A to its upper echelon form U.

Definition 10 *Two matrices $A, B \in \mathbb{R}^{m\times n}$ are said to be* <u>row equivalent</u> *if a finite number of row operations (addition, multiplication and permutation) convert A to B. When such matrices are row equivalent we write*

$$A \sim B \ .$$

Theorem 2

$$\begin{aligned} A \sim B \quad &\Rightarrow \quad row(A) = row(B) \\ A \sim B \quad &\Rightarrow \quad N(A) = N(B) \end{aligned}$$

Row operations do not preserve the column space. For instance

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \sim B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

by a simple permutation of rows but clearly $col(A) \neq col(B)$.

Definition 11 *Let $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. A system $A\mathbf{x} = \mathbf{b}$ is* <u>consistent</u> *if it has a solution.*

Theorem 3 *(General Solutions) Let $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$,*

$$A\mathbf{x}_0 \quad = \quad \mathbf{b} \ .$$

Then,

$$A\mathbf{x} = \mathbf{b} \quad \Rightarrow \quad \exists \mathbf{v} \in N(A) \ such \ that \ \mathbf{x} = \mathbf{x}_0 + \mathbf{v} \ .$$

Here $\mathbf{x}_0$ is called a <u>particular solution</u> and $\mathbf{v}$ is the <u>homogeneous solution</u>. Written another way, if $\mathbf{x}_0$ is "a" solution and $\mathbf{x}$ is any other solution then there exists constants $c_1, \ldots c_k$ such that

$$\mathbf{x} = \mathbf{x}_0 c_1 \mathbf{v}_1 + \cdots c_k \mathbf{v}_k$$

where

$$E = \{\mathbf{v}_1, \ldots \mathbf{v}_k\}$$

is a basis for $N(A)$. Also, conversely, if $A\mathbf{x}_0 = \mathbf{b}$, $\mathbf{v} \in N(A)$ and $\mathbf{x} = \mathbf{x}_0 + \mathbf{v}$ then $A\mathbf{x} = \mathbf{b}$.

Next we describe one method for finding bases for $row(A)$, $N(A)$ and $col(A)$. Suppose that after row reduction one reduces A to U having the form:

$$A \sim U = \begin{bmatrix} 1 & * & * & * & * & * \\ 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \ldots & \mathbf{u}_1 & \ldots \\ \ldots & \mathbf{u}_2 & \ldots \\ \ldots & \mathbf{u}_3 & \ldots \\ \ldots & \mathbf{u}_4 & \ldots \\ \ldots & 0 & \ldots \end{bmatrix}$$

In this example, there are 4 pivots (leading ones in rows). A basis $E(row(A))$ for $row(A)$ is the four non-zero row vectors of U, i.e.,

$$E(row(A)) = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$$

from which we know $dim(row(A)) = 4$. Also, the 4 pivots in U occur in columns 1,3,4 and 6. A basis $E(col(A))$ for $col(A)$ is the $1^{rst}, 3^{rd}, 4^{rth}$ and 6^{th} columns of A, i.e.,

$$E(col(A)) = \{\mathbf{c}_1, \mathbf{c}_3, \mathbf{c}_4, \mathbf{c}_6\}$$

The columns of U which contain no pivots correspond to *free variables*. There are 2 free variables x_2 and x_5 since columns 2 and 5 contain no pivots. This means that by backsolving $U\mathbf{x} = 0$, the remaining variables can be written in terms of x_2 and x_5. This procedure implies that any solution of $U\mathbf{x} = 0$ can be written in the form

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2$$

where the vectors $\mathbf{v}_1, \mathbf{v}_2$ form a basis $E(N(A))$ for $N(A)$, i.e.,

$$E(N(A)) = \{\mathbf{v}_1, \mathbf{v}_2\}$$

A basis for $N(A^T)$ is found by row reducing A^T and applying a similar procedure.

Note that an alternate method for finding a basis for $col(A)$ uses the fact that $col(A) = row(A^T)$. Thus, by finding a basis for $row(A^T)$ thru row reduction of A^T, one is actually finding a basis for $col(A)$.

Knowing these methods for finding bases we have the following definitions and Theorem.

Definition 12

$$\begin{aligned} rank(A) &\equiv dim(row(A)) \\ nullity(A) &\equiv dim(N(A)) \end{aligned}$$

Theorem 4 *Let* $r = rank(A)$ *and* $A \in \mathbb{R}^{m\times n}$.

$$\begin{aligned} dim(row(A)) &= r \\ dim(col(A)) &= r \\ dim(N(A)) &= n - r \\ dim(N(A^T)) &= m - r \end{aligned}$$

5 Linear Transformations on $\mathbb{R}^n$

Definition 13 *A* <u>*linear transformation*</u> T *on* $\mathbb{R}^n$ *is a function* $T : \mathbb{R}^n \to \mathbb{R}^m$ *such that*

$$T(\mathbf{x}) = A\mathbf{x}$$

for some matrix $A \in \mathbb{R}^{m\times n}$*. The matrix* A *is called the* <u>*standard matrix*</u> *associated with* T *which we notationally denote*

$$[T] = A$$

so that $T(\mathbf{x}) = [T]\mathbf{x}$.

This definition implies certain algebraic properties about linear transformations on $\mathbb{R}^n$:

Theorem 5 $T : \mathbb{R}^n \to \mathbb{R}^m$ *is a linear transformation if and only if*

$$(a)\ T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}) \quad , \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \tag{1}$$

$$(b)\ T(k\mathbf{x}) = kT(\mathbf{x}) \quad , \quad \forall \mathbf{x} \in \mathbb{R}^n, \forall k \in \mathbb{R} \tag{2}$$

This equivalence mean that properties a)-b) of the Theorem could be used to define linear transformations on $\mathbb{R}^n$. Later, this will be the definition for linear transformations on abstract vector spaces V.

Definition 14 *Let* f *be a function from* X *into* Y*, i.e.,* $f : X \to Y$*. The* <u>*domain*</u> $D(f)$ *of* f *is defined by:*

$$D(f) = \{\mathbf{x} \in X : f(\mathbf{x})\ is\ defined\}$$

The <u>*range*</u> $R(f)$ *of* f *is defined by:*

$$R(f) = \{\mathbf{y} \in Y : \mathbf{y} = f(\mathbf{x})\ for\ some\ \mathbf{x} \in X\}$$

In this setting Y *is called the* <u>*codomain*</u> *of* f*. Also, if* $\mathbf{y} = f(\mathbf{x})$ *for some* $\mathbf{x} \in D(f)$*, then* $\mathbf{y}$ *is the* <u>*image*</u> *of* $\mathbf{x}$ *under* f*.*

Note that if T is a linear transformation on $\mathbb{R}^n$, $D(T) = \mathbb{R}^n$. In general, however, $R(T) \subset \mathbb{R}^m$.

Definition 15 *The function* $f : X \to Y$ *is* $1-1$ *on* $D(f)$ *if*

$$\forall \mathbf{x}_1, \mathbf{x}_2 \in D(f), \quad , \quad f(\mathbf{x}_1) = f(\mathbf{x}_2) \quad \Rightarrow \mathbf{x}_1 = \mathbf{x}_2$$

Definition 16 *If* $f : X \to Y$ *is 1-1 on* $D(f)$ *then* f *has an* <u>*inverse*</u> $f^{-1} : Y \to X$ *where* $D(f^{-1}) = R(f)$ *and*

$$f^{-1}(f(\mathbf{x})) = f(f^{-1}(\mathbf{x})) = \mathbf{x}, \quad , \quad \forall \mathbf{x} \in D(f)$$

For linear transformations T on $\mathbb{R}^n$ that are 1-1, the inverse of T is denoted T^{-1} and

$$[T^{-1}] = [T]^{-1}\ .$$

Theorem 6 *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ and $T(\mathbf{x}) = [T]\mathbf{x} = A\mathbf{x}$. Then, the following are equivalent:*

a) T *is 1-1*

b) A *is invertible*

c) $N(A) = \{0\}$

d) $A\mathbf{x} = \mathbf{b}$ *is consistent* $\forall \mathbf{b} \in \mathbb{R}^n$.

e) $det(A) \neq 0$

f) $R(T) = col(A) = row(A) = \mathbb{R}^n$

g) $rank(A) = n$

h) $nullity(A) = 0$

If the standard basis vectors for $\mathbb{R}^n$ are $\mathbf{e}_1, \ldots \mathbf{e}_n$ then we have the following useful Theorem for determining the standard matrix $[T]$ of a linear transformation T:

Theorem 7 *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then,*

$$[T] = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \vdots & T(\mathbf{e}_n) \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

6 Inner Products

Definition 17 *Let V be a vector space. By an* <u>*inner product*</u> *on V we mean a real valued function $< u, v >$ on $V \times V$ which satisfies the following axioms:*

a) $< \mathbf{u}, \mathbf{v} >=< \mathbf{v}, \mathbf{u} > \quad , \quad \forall \mathbf{u}, \mathbf{v} \in V$

b) $< \mathbf{u} + \mathbf{w}, \mathbf{v} >=< \mathbf{u}, \mathbf{v} > + < \mathbf{w}, \mathbf{v} > \quad , \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$

c) $< k\mathbf{u}, \mathbf{v} >= k < \mathbf{u}, \mathbf{v} > \quad , \quad \forall \mathbf{u}, \mathbf{v} \in V, k \in \mathbb{R}$

d) $< \mathbf{u}, \mathbf{u} > \geq 0 \quad , \quad \forall \mathbf{u} \in V$

e) $< \mathbf{u}, \mathbf{u} >= 0 \Leftrightarrow u = 0$

If V has an inner product defined on it then V is said to be an inner product space.

In the definition above since $< \mathbf{u}, \mathbf{v} >$ and k are real, V is sometimes said to be an inner product space over the real field. In this case, if $f(\mathbf{u}, \mathbf{v}) \equiv< \mathbf{u}, \mathbf{v} >$ then $f : V \times V \to \mathbb{R}$. However, if $< \mathbf{u}, \mathbf{v} >$ and k are complex numbers, V is an inner product space over the complex field where a) and c) are replaced by

a') $< \mathbf{u}, \mathbf{v} >= \overline{< \mathbf{v}, \mathbf{u} >} \quad , \quad \forall \mathbf{u}, \mathbf{v} \in V$

c') $< k\mathbf{u}, \mathbf{v} >= \bar{k} < \mathbf{u}, \mathbf{v} > \quad , \quad \forall \mathbf{u}, \mathbf{v} \in V, k \in \mathbb{C}$

and $\bar{(\ )}$ denotes complex conjugate.

Below we give examples of several inner product spaces. In these examples, note that V may have many different inner products.

Example 3 *Scalar multiplication on $V = \mathbb{R}$:*

$$< u, v >= uv$$

Example 4 *Euclidean inner product on $V = \mathbb{R}^n$:*

$$< u, v >= u_1 v_1 + \ldots u_n v_n = \sum_{i=1}^{n} u_i v_i$$

This is also known as the dot *product and notationally written*

$$< \mathbf{u}, \mathbf{v} >= \mathbf{u} \cdot \mathbf{v}$$

Considering $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{n \times 1}$ as matrices, this can equivalently be written

$$< \mathbf{u}, \mathbf{v} >= \mathbf{u}^T \mathbf{v}$$

Example 5 *Weighted Euclidean inner product on $V = \mathbb{R}^n$. Let $\omega_i > 0, \forall i$.*

$$< \mathbf{u}, \mathbf{v} >= \omega_1 u_1 v_1 + \ldots \omega_n u_n v_n = \sum_{i=1}^{n} \omega_i u_i v_i$$

Example 6 *Matrix induced inner product on* $V = \mathbb{R}^n$*: Let* $A \in \mathbb{R}^{n\times n}$ *have an inverse.*

$$< \mathbf{u}, \mathbf{v} >= (A\mathbf{u}) \cdot (A\mathbf{v}) = (A\mathbf{u})^T(A\mathbf{v})$$

Example 7 *An inner product space on* $V = M_{nn}$*,* $n \geq 1$*.*

$$< \mathbf{u}, \mathbf{v} >= Tr(\mathbf{u}^T\mathbf{v})$$

where if $A \in \mathbb{R}^{n\times n} = [a_{ij}]$*, the trace* $Tr(A)$ *is the sum of its diagonal elements, i.e.,*

$$Tr(A) = a_{11} + \ldots a_{nn} = \sum_{i=1}^{n} a_{ii}$$

Example 8 *Other inner products on* $V = M_{nn}$*,* $n \geq 1$*. For every element* $\mathbf{v} \in V$ *one can define a unique element* $\hat{\mathbf{v}} \in \mathbb{R}^{n^2}$ *as follows:*

$$\mathbf{v} = [v_{ij}] \quad \Rightarrow \quad \hat{\mathbf{v}} = \begin{pmatrix} v_{11} \\ \vdots \\ v_{1n} \\ v_{21} \\ \vdots \\ v_{nn} \end{pmatrix}$$

Then if we let $< \hat{\mathbf{u}}, \hat{\mathbf{v}} >_{\mathbb{R}^n}$ *be any inner product on* $\mathbb{R}^n$ *we define the inner product on* V *as follows:*

$$< \mathbf{u}, \mathbf{v} >=< \hat{\mathbf{u}}, \hat{\mathbf{v}} >_{\mathbb{R}^n}$$

If one chooses $< \hat{\mathbf{u}}, \hat{\mathbf{v}} >_{\mathbb{R}^n}$ *to be the Euclidean inner product on* $\mathbb{R}^n$*, the definition above yields the same inner product described in Example 7, i.e.,*

$$< \mathbf{u}, \mathbf{v} >=< \hat{\mathbf{u}}, \hat{\mathbf{v}} >_{\mathbb{R}^n}= Tr(\mathbf{u}^T\mathbf{v})$$

Example 9 L^2 *inner product on the function space* $V = C[a, b]$*:*

$$< \mathbf{u}, \mathbf{v} >= \int_a^b u(x)v(x)dx$$

Example 10 *Weighted* L^2 *inner product on the function space* $V = C[a, b]$*. Let* $\omega(x) > 0, \omega \in C[a, b]$*, then*

$$< \mathbf{u}, \mathbf{v} >= \int_a^b \omega(x)u(x)v(x)dx$$

We now make an observation that if $V = \mathbb{R}^n$ then for each fixed $\mathbf{v}$

$$T_{\mathbf{v}}(\mathbf{u}) \equiv< \mathbf{u}, \mathbf{v} >$$

is a linear transformation from $\mathbb{R}^n$ into $\mathbb{R}$, i.e., $T_{\mathbf{v}} : \mathbb{R}^n \to \mathbb{R}$. This fact follows from b) and c) in the definition of the inner product.

7 Norms induced by Inner products

A norm on any vector space is defined by:

Definition 18 *We say* $\| u \|$ *is a norm on a vector space* V *if* $\forall u, v \in V$ *and* $\alpha \in \mathbb{R}$*,*

a) $\| \alpha u \| = |\alpha| \, \| u \|$

b) $\| u \| \geq 0$

c) $\| u \| = 0 \Leftrightarrow u = 0$

d) $\| u + v \| \leq \| u \| + \| v \|$

If V is an inner product space then

$$\| u \| \equiv \sqrt{< u, u >}$$

is the inner product induced norm for V. That this norm satisifies a)-c) in the above definition is easy to see. Showing the triangle inequality d) is satisfied requires the Cauchy-Schwartz inequality, however.

Theorem 8 *Let* V *be an inner product space and assume* $\| u \|$ *is the inner product induced norm. Then*

$$| < u, v > | \leq \| u \| \| v \| \quad , \quad \forall u, v \in V$$

Proof: If $u = 0$ *equality is attained so the statement is true. Thus, assume* $u \neq 0$ *and define* $P(t) = \| tu + v \|^2$*. By properties of inner products we have*

$$P(t) = at^2 + 2bt + c = \| u \|^2 \, t^2 + 2 < u, v > t + \| v \|^2$$

Since $P(t) \geq 0$ *and is quadratic in* t *it has either one root or no roots. In either case*

$$b^2 - ac \leq 0$$

Written another way,

$$< u, v >^2 \leq \| u \|^2 \| v \|^2$$

from which the result follows.

With this we now state

Theorem 9 *Let* V *be an inner product space and let*

$$\| u \| \equiv \sqrt{< u, u >}$$

Then $\| u \|$ *defines a norm on* V*.*

Proof: We only verify d) since a)-c) are trivial. Let $u, v \in V$. Then

$$\begin{aligned} \| u+v \|^2 &= \| u \|^2 + 2 < u, v > + \| v \|^2 \\ &\leq \| u \|^2 + 2| < u, v > | + \| v \|^2 \\ &\leq \| u \|^2 + 2 \| u \| \| v \| + \| v \|^2 \\ &\leq (\| u \| + \| v \|)^2 \end{aligned}$$

from which the result follows.

Example 11 *Euclidean norm on $V = \mathbb{R}^n$.*

$$\| u \| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$

Example 12 *L^2 norm on $V = C[a, b]$.*

$$\| u \| = \sqrt{\int_a^b u(x)^2 dx}$$

Example 13 *Norm on $V = M_{nn}$.*

$$\| u \| = \sqrt{Tr(u^T u)}$$

Given every inner product space has a norm, every inner product space is also a metric space with metric (or "distance")

$$d(u, v) = \| u - v \|$$

8 Orthogonality

Definition 19 *Let V be an inner product space. $u, v \in V$ are said to be orthorgonal if*

$$< u, v > = 0$$

For any subspace W of V, one can define the space of vectors which are orthogonal to every element of W:

Definition 20 *Let V be an inner product space and W be a subspace of V. Then, the orthogonal complement $W^\perp$ of W is*

$$W^\perp = \{v \in V : < v, w > = 0 \; , \; \forall w \in W\}$$

The following Theorem (withour proof) summarizes several important facts about orthogonal complements:

Theorem 10 *Let V be a finite dimensional inner product space and X, Y, W be subspaces of V. Then*

a) $\{0\}^\perp = V$

b) $W^\perp$ *is a subspace of* V.

c) $W \cap W^\perp = \{0\}$

d) $(W^\perp)^\perp = W$.

e) $X \subset Y \Rightarrow Y^\perp \subset X^\perp$

A very important Theorem in linear algebra relates to the four fundamental matrix subspaces.

Theorem 11 *(Orthogonality of Matrix Subspaces) Let $A \in \mathbb{R}^{m\times n}$ and let orthogonal complements be defined using the Euclidean inner product. Then,*

a) $row(A) = N(A)^\perp$

b) $col(A) = N(A^T)^\perp$

From this arises the Fredholm Alternative [1] on $\mathbb{R}^n$:

Theorem 12 *Let $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$. Then*

$$Ax = b \; has \; a \; solution \; x \qquad \Leftrightarrow \qquad < v, b >= 0 \; , \; \forall v \in N(A^T)$$

A further large result is that W and $W^\perp$ can be used to "decompose" a finite dimensional space into two parts. To make this precise we first make the following definitions:

Definition 21 *Let $X, Y \subset V$ where V is a vector space. Then, the set $X + Y$ is defined as all possible sums of elements in X and Y:*

$$X + Y = \{x + y : x \in X, y \in Y\}$$

Definition 22 *Let V be a vector space and suppose*

i) X, Y *are subspaces of* V.

ii) $X \cap Y = \{0\}$

iii) $V = X + Y$

then $X + Y$ is called a direct sum of X and Y and we write

$$V = X \oplus Y$$

Now we state the decomposition Theorem:

[1] technically this is only part of the "alternative"

Theorem 13 *Let W be a subspace of a finite dimensional inner product space V. Then,*

$$V = W \oplus W^{\perp}$$

Moreover, for every $v \in V$ there exist unique $w \in W$ and $w^{\perp} \in W^{\perp}$ such that

$$v = w + w^{\perp}$$

Here the unique $w \in W$ is called the projection of v onto W and is denoted:

$$w = proj_W v$$

When applied to the fundamental matrix subspaces, this Theorem implies for any matrix $A \in \mathbb{R}^{m \times n}$

$$\begin{aligned} \mathbb{R}^n &= row(A) \oplus N(A) \\ \mathbb{R}^m &= col(A) \oplus N(A^T) \end{aligned}$$

9 Appendix on Symbol Notations

$=$	equals
$\equiv$	is defined as
$\Rightarrow$	implies
$\Leftrightarrow$	is equivalent to
$\exists$	there exists
$\forall$	for all
$\in$	is an element of
$\cup$	union
$\cap$	intersect
$\subset$	subset or proper subset
$\subseteq$	subset
$+$	vector addition
$\oplus$	vector addition or direct sum
$\odot$	scalar multiplication
$\cdot$	dot product or scalar multiplication
$\parallel u \parallel$	norm of u
Σ	sum
$\Sigma_{i=0}^{n} u_i$	$u_1 + u_2 + \ldots u_n$
$d(u, v)$	distance between u and v