MATH 235 Midterm 1 SOS Review Package

June 4, 2011

Fundamental Subspaces

First we review the four fundamental subspaces of a matrix and the associated subspace of the linear map which represents the matrix mapping.

Definition. Let A be an $m \times n$ matrix with columns $\vec{c}_1, \cdots, \vec{c}_n \in \mathbb{R}^m$ and rows $\vec{r}_1, \cdots \vec{r}_m \in \mathbb{R}^n$ so $A = \begin{bmatrix} \vec{c}_1 & \cdots & \vec{c}_n \end{bmatrix} = \begin{bmatrix} \vec{r}_1^{\,T} \\ \vdots \\ \vec{r}_m^{\,T} \end{bmatrix}$.
The **column space** of A is the set $\text{col}(A) = \text{span}(\{\vec{c}_1, \cdots, \vec{c}_n\})$.
The **null space** of A is the set $\text{null}(A) = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\}$.
The **row space** of A is the set $\text{row}(A) = \text{span}(\{\vec{r}_1, \cdots, \vec{r}_m\})$.
The **left null space** of A is the set $\text{null}(A^T)$.

Definition. Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map.
The **range** of L is the set $\{L(\vec{x}) : \vec{x} \in \mathbb{R}^n\}$.
The **kernel** of L is the set $\{\vec{x} \in \mathbb{R}^n : L(\vec{x}) = \vec{0}\}$.

We recall the following:

1. $\text{null}(A)$, $\text{row}(A)$, $\ker(L)$ are subspaces of $\mathbb{R}^n$
2. $\text{col}(A)$, $\text{null}(A^T)$, $\text{range}(L)$ are subspaces of $\mathbb{R}^m$
3. Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be defined by $L(\vec{x}) = A\vec{x}$. Then $\text{range}(L) = \text{col}(A)$ and $\ker(L) = \text{null}(A)$.

Proof. We prove only part 3 which follows immediately from the fact that

$$a_1\vec{c}_1 + \cdots + a_n\vec{c}_n = \begin{bmatrix} \vec{c}_1 & \cdots & \vec{c}_n \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = A\vec{x}$$

where $\vec{x} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$. □

Definition. Let A be an $m \times n$ matrix. The **rank** of A is the number of leading ones in the reduced row echelon form of A.

Example. If $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$, then $R = \text{RREF}(A) = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ and so $\text{rank}(A) = 2$.

Using the definition of the reduced row echelon form, we obtain the following fact.

Fact. Let A be an $m \times n$ matrix and let $R = \text{RREF}(A)$. Then the columns of R with leading ones form a basis for $\text{col}(R)$ and the non-zero rows of R form a basis for $\text{row}(R)$.

Theorem. Let A be an $m \times n$ matrix. Then

1. $\text{rank}(A) = \dim(\text{col}(A))$
2. $\text{rank}(A) = \dim(\text{row}(A))$.

Proof. We prove part 1. Let $\text{rank}(A) = r$. The idea of the proof is to find a basis for $\text{col}(A)$ with r elements. We consider two cases.
Case 1: $r = 0$. If $r = 0$, then $R = 0$. Let $E_1, \cdots E_k$ be a sequence of elementary row operations such that $R = E_k \cdots E_1 A$ so

$$A = E_1^{-1} \cdots E_k^{-1} R = E_1^{-1} \cdots E_k^{-1} 0 = 0.$$

Therefore $\text{col}(A) = \{\vec{0}\}$ and $\dim(\text{col}(A)) = 0 = r$.
Case 2: $r > 0$. Let $i_1 \cdots i_r$ be the columns where the leading ones occur and let $\mathcal{B} = \{\vec{c}_{i_1}, \cdots, \vec{c}_{i_n}\}$ be the set of corresponding columns. of R. By the fact $\mathcal{B}$ is basis for $\text{col}(R)$. Our plan now is to use $\mathcal{B}$ to find a basis for $\text{col}(A)$. Let $E = E_k \cdots E_1$. Then $R = EA$ and $A = E^{-1}R$. We claim $\mathcal{C} = \{A\vec{e}_{i_1}, \cdots, A\vec{e}_{i_r}\}$ is basis for $\text{col}(A)$. Notice that $A\vec{e}_{i_j} = E^{-1}R\vec{e}_{i_j} = E^{-1}\vec{c}_{i_j}$ and so $\mathcal{C} = \{E^{-1}\vec{c}_{i_1}, \cdots, E^{-1}\vec{c}_{i_r}\}$.
We must verify $\mathcal{C}$ is linearly independent. Suppose $a_1 E^{-1}\vec{c}_{i_1} + \cdots + a_r E^{-1}\vec{c}_{i_r} = \vec{0}$ for some scalars $a_1 \cdots a_r \in \mathbb{R}$. Then $E^{-1}(a_1\vec{c}_{i_1} + \cdots + a_r\vec{c}_{i_r}) = \vec{0}$ and hence $a_1\vec{c}_{i_1} + \cdots a_r\vec{c}_{i_r} \in \text{null}(E^{-1}) = \{\vec{0}\}$ since E is one-to-one.
Hence $a_1\vec{c}_{i_1} + \cdots + a_r\vec{c}_{i_r} = \vec{0} \implies a_1 = \cdots = a_r = 0$ as $\mathcal{B}$ is linearly independent.
We must check $\text{span}(\mathcal{C}) = \text{col}(A)$.
" $\subseteq$ " It is easy to see that $\mathcal{C} \subseteq \text{col}(A)$ and so $(\mathcal{C}) \subseteq \text{col}(A)$.
" $\supseteq$ " Let $\vec{y} \in \text{col}(A)$. Then $\vec{y} = A\vec{x}$ for some $\vec{x} \in \mathbb{R}^n$. But $A = E^{-1}R$ so $\vec{y} = E^{-1}(R\vec{x})$. Since $R\vec{x} \in \text{col}(R)$, we can write $R\vec{x} = a_1\vec{c}_{i_1} \cdots + \vec{c}_{i_r}$ for some scalars $a_1, \cdots, a_r \in \mathbb{R}$.
Therefore $\vec{y} = E^{-1}(R\vec{x}) = a_1 E^{-1}\vec{c}_{i_1} + \cdots + a_r E^{-1}\vec{c}_{i_r} \in \text{span}(\mathcal{C})$.
We have shown $\mathcal{C}$ is a basis for $\text{col}(A)$ and so $\dim(\text{col}(A)) = r = \text{rank}(A)$. □

Corollary. Let A be an $m \times n$ matrix. Then $\text{rank}(A) = \text{rank}(A^T)$.

Proof. We have

$$\text{rank}(A) = \dim(\text{row}(A)) = \dim(\text{col}(A^T)) = \text{rank}(A^T)$$

. ∎

Rank-Nullity Theorem

Definition. Let V and W be vector spaces. $L : \mathsf{V} \to \mathsf{W}$ is **linear** if for each $\vec{x}, \vec{y} \in \mathsf{V}$ and $c \in \mathbb{R}$, we have

$L(\vec{x} + \vec{y}) = L(\vec{x}) + L(\vec{y})$ (L preserves addition)

$L(c\vec{x}) = cL(\vec{x})$ (L preserves scalar multiplication)

Given vector spaces V and W, we have the following important construction of a linear map $L : \mathsf{V} \to \mathsf{W}$.

Lemma. Let V and W be vector spaces. Let $\mathcal{B} = \{\vec{v}_1, \cdots, \vec{v}_n\}$ be a basis for V and let $\mathcal{C} = \{\vec{w}_1, \cdots, \vec{w}_n\} \subset \mathsf{W}$. Then the map $L : \mathsf{V} \to \mathsf{W}$ given by $L(\vec{v}_i) = \vec{w}_i$ is linear.

Proof. Let $\vec{x}, \vec{y} \in \mathsf{V}$ and $c \in \mathbb{R}$ be given. Then $\vec{x} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$ and $\vec{y} = b_1\vec{v}_1 + \cdots + b_n\vec{v}_n$ for some $a_1, \cdots, a_n, b_1, \cdots, b_n \in \mathbb{R}$ and

$$\begin{aligned} L(c\vec{x} + \vec{y}) &= L((ca_1 + b_1)\vec{v}_1 + \cdots + (ca_n + b_n)\vec{v}_n) \\ &= (ca_1 + b_1)\vec{w}_1 + \cdots + (ca_n + b_n)\vec{w}_n) \\ &= c(a_1\vec{w}_1 + \cdots + a_n\vec{w}_n) + (b_1\vec{w}_1 + \cdots + b_n\vec{w}_n) \\ &= cL(\vec{x}) + L(\vec{y}) \end{aligned}$$

∎

We give an example demonstrating the use of this construction.

Example. Let V be a vector space with $\dim(\mathsf{V}) = n$. Let S be a subspace of V. Show that there exists a linear map L with $\ker(L) = \mathsf{S}$.

Solution. Let $\mathcal{B} = \{\vec{v}_1, \cdots, \vec{v}_k\}$ be a basis for S. We can get $\{\vec{v}_{k+1}, \cdots, \vec{v}_n\} \subset \mathsf{V}$ such that $\mathcal{C} = \{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis for V. We define a linear map $L : \mathsf{V} \to \mathsf{V}$ by $L(\vec{v}_i) = \vec{0}$ for $1 \leq i \leq k$ and $L(\vec{v}_i) = \vec{v}_i$ for $k+1 \leq i \leq n$. It remains to verify that $\ker(L) = \mathsf{S}$.

Let $\vec{x} \in \mathsf{S}$. Then $\vec{x} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_k$ for some $a_1, \cdots, a_k \in \mathbb{R}$. Therefore $L(\vec{x}) = \vec{0}$ and thus $\mathsf{S} \subseteq \ker(L)$.

Let $\vec{x} \in \ker(L)$. Let $\vec{x} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$. Then

$$L(\vec{x}) = a_{k+1}\vec{v}_{k+1} + \cdots + a_n\vec{v}_n = \vec{0}$$

and, since $\mathcal{C}$ is linearly independent, $a_{k+1} = \cdots = a_n = 0$. Thus

$$\vec{x} = a_1\vec{v}_1 + \cdots + a_k\vec{v}_k \in \mathsf{S}$$

and $\ker(L) \subseteq \mathsf{S}$. □

Remark. Notice how we first obtained a basis for $\ker(L)$ and then extended it to a basis for V, as opposed to starting with a basis for V. In the latter approach, it is not necessary that any subset of our basis is a basis for $\ker(L)$. Starting with a basis for a subspace and then extending it is an important proof technique. This idea was also used in the proof of the Rank-Nullity theorem.

Definition. Let $L : \mathsf{V} \to \mathsf{W}$ be a linear map from a vector space V to a vector space W. The **nullity of** L is $\text{nullity}(L) = \dim(\ker(L))$ and the **rank of** L is $\text{rank}(L) = \dim(\text{range}(L))$.

The following theorem is the first of two extremely important theorems in this course. The other theorem is the Principal Axis Theorem which will be covered on Midterm 2.

Theorem (Rank-Nullity)**.** Let V and W be vector spaces with $\dim(\mathsf{V})$ finite and let $L : \mathsf{V} \to \mathsf{W}$ be a linear map. Then

$$\text{rank}(L) + \text{nullity}(L) = \dim(\mathsf{V}).$$

We give an example.

Example. Let $L : M_{2\times 2}(\mathbb{R}) \to P_3$ be defined by

$$L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a+b)x^3 + cx.$$

Find a basis for $\text{range}(L)$ and $\ker(L)$ and verify the Rank-Nullity theorem.

Solution. We have

$$\begin{aligned}
\vec{x} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \ker L &\iff (a+b)x^3 + cx = 0 \\
&\iff a = -b \text{ and } c = 0 \\
&\iff \vec{x} = \begin{bmatrix} a & -a \\ 0 & d \end{bmatrix} = a\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\
&\iff \vec{x} \in \text{span}\left(\left\{\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}\right)
\end{aligned}$$

Therefore $\ker(L) = \text{span}\left(\left\{\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}\right)$ and so $\text{nullity}(L) = 2$.

Also, we have

$$\begin{aligned}
\text{range}(L) &= \left\{ L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{R}) \right\} \\
&= \{(a+b)x^3 + cx : a, b, c \in \mathbb{R}\} \\
&= \{ax^3 + bx : a, b \in \mathbb{R}\} \\
&= \text{span}(\{x, x^3\})
\end{aligned}$$

Therefore $\text{rank}(L) = 2$ and so $\text{rank}(L) + \text{nullity}(L) = 4 = \dim(M_{2\times 2}(\mathbb{R}))$. □

Next we give some examples to show how the Rank-Nullity theorem can be used in theoretical problems.

Example. Let $L : \mathbb{R}^n \to \mathbb{R}^m$ and $M : \mathbb{R}^m \to \mathbb{R}^p$ be linear mappings. Show that (i) $\text{rank}(M \circ L) \leq \text{rank}(M)$ and (ii) $\text{rank}(M \circ L) \leq \text{rank}(L)$.

Solution. (i) We see that $\text{range}(M \circ L) = \{M(L(\vec{x})) : \vec{x} \in \mathbb{R}^n\} \subseteq \text{range}(M)$ so it follows that $\text{rank}(M \circ L) \leq \text{rank}(M)$.

(ii) We see that $\ker(L) \subseteq \ker(M \circ L)$ since if $\vec{x} \in \ker(L)$ then we have $(M \circ L)(\vec{x}) = M(\vec{0}) = \vec{0}$. Thus $\text{nullity}(L) \leq \text{nullity}(M \circ L)$. Therefore

$$\begin{aligned}
\text{rank}(M \circ L) &= \dim(\mathbb{R}^n) - \text{nullity}(M \circ L) \\
&= n - \text{nullity}(M \circ L) \\
&\leq n - \text{nullity}(L) \\
&= \text{rank}(L).
\end{aligned}$$

□

Remark. Notice how we proved something about the $\text{nullity}(L)$ then used the Rank-Nullity theorem to say something about $\text{rank}(L)$. This is an important proof technique.

Example. Let A be an $n \times m$ matrix. Show that $A\vec{x} = \vec{b}$ is consistent for all $\vec{b} \in \mathbb{R}^m \iff A^T\vec{y} = \vec{0}$ has only the trivial solution.

Solution. ($\Longrightarrow$) Suppose $A\vec{x} = \vec{b}$ is consistent for all $\vec{b} \in \mathbb{R}^m$. Then $\text{col}(A) = \mathbb{R}^m$ since for each $\vec{b} \in \mathbb{R}^m$ we can find $\vec{x} \in \mathbb{R}^n$ with $A\vec{x} = \vec{b}$ so $\vec{b} \in \text{col}(A)$. Then $\text{rank}(A) = m$. We wish to show $\text{nullity}(A^T) = 0$. We have

$$\begin{aligned}
\text{nullity}(A^T) &= \mathbb{R}^m - \text{rank}(A^T) \text{ by the Rank-Nullity theorem} \\
&= m - \text{rank}(A) \text{ since } \text{rank}(A) = \text{rank}(A^T) \\
&= 0
\end{aligned}$$

Therefore $\text{null}(A^T) = \{\vec{0}\}$ and so $A^T\vec{y} = \vec{0}$ has only the trivial solution.
($\Longleftarrow$) Suppose $A^T\vec{y} = \vec{0}$ has only the trivial solution. Then $\text{nullity}(A^T) = 0$ so $\text{rank}(A) = \text{rank}(A^T) = \mathbb{R}^m - \text{nullity}(A^T) = m$ and $A\vec{x} = \vec{b}$ is consistent for all $\vec{b}$ as $\text{col}(A) = \mathbb{R}^m$. □

Remark. Notice here how we turned the statements "$A\vec{x} = \vec{b}$ is consistent for all $\vec{b} \in \mathbb{R}^m$" and "$A^T\vec{y} = \vec{0}$ has only the trivial solution" into statements about $\mathrm{rank}(A)$ and $\mathrm{nullity}(A)$ so that we could apply the Rank-Nullity Theorem.

Isomorphisms

We recall the following definitions from calculus.

Definition. Let X, Y be non-empty sets and let $f : X \to Y$ be a function.

1. f is **one-to-one** if whenever $f(x_1) = f(x_2)$ for some $x_1, x_2 \in X$, then $x_1 = x_2$
2. f is **onto** if for each $y \in Y$, there exists $x \in X$ such that $f(x) = y$.

The following is easily verified.

Proposition. Let $\mathsf{U}, \mathsf{V}, \mathsf{W}$ be vector spaces and let $L : \mathsf{U} \to \mathsf{V}$ and $M : \mathsf{V} \to$ be linear maps. Then

1. $M \circ L : \mathsf{U} \to \mathsf{W}$ given by $(M \circ L)(\vec{u}) = M(L(\vec{u}))$ is linear.
2. If L and M are one-to-one, then $M \circ L$ is one-to-one.
3. If L and M are onto- then $M \circ L$ is onto.
4. L is one-to-one $\iff$ $\ker(L) = \{\vec{0}\}$.

Proof. This is a good exercise and likely to appear on your midterm! □

Example. Let $\mathsf{U}, \mathsf{V}, \mathsf{W}$ be vector spaces and let $L : \mathsf{U} \to \mathsf{V}$ and $M : \mathsf{V} \to$ be linear maps.

1. If $M \circ L$ is one-to-one, is L one-to-one? Is M one-to-one?
2. If $M \circ L$ is onto, is L onto? Is M onto?

Solution. 1. We see that L is one-to-one. We could verify this using the definition, however, we use the previous proposition. Since $\ker(L) \subseteq \ker(M \circ L) = \{\vec{0}\}$, we have that $\ker(L) = \{\vec{0}\}$ and so L is one-to-one.
M may not be one-to-one. For instance, consider the maps $L : \mathbb{R} \to \mathbb{R}^2$ defined by $L(t) = t\begin{bmatrix}1\\1\end{bmatrix}$ and $M : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $M\left(\begin{bmatrix}x\\y\end{bmatrix}\right) = \begin{bmatrix}x+2y\\2x+4y\end{bmatrix}$. Then $\ker(M) = \mathrm{span}\left(\left\{\begin{bmatrix}2\\-1\end{bmatrix}\right\}\right)$. Then we have

$$\begin{aligned}(M \circ L)(t) = \vec{0} &\iff t\begin{bmatrix}1\\1\end{bmatrix} \in \mathrm{span}\left\{\left(\begin{bmatrix}2\\-1\end{bmatrix}\right)\right\}\\ &\iff t = 0\end{aligned}$$

and thus $\ker(M \circ L) = \{0\}$ and so $M \circ L$ is one-to-one. But M is not one-to-one as $M\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}\right) = \vec{0}$.

2. We see that M is onto. Let $\vec{w} \in \mathsf{W}$ be given. Since $M \circ L$ is onto, there exists $\vec{u} \in \mathsf{U}$ such that $(M \circ L)(\vec{u}) = \vec{w}$. Setting $\vec{v} = L(\vec{u})$ we have $M(\vec{v}) = \vec{w}$. Thus we have M is onto.

L may not be onto. Consider the maps $L : \mathbb{R} \to \mathbb{R}^2$ given by $L(t) = \begin{bmatrix} t \\ 0 \end{bmatrix}$ and $M : \mathbb{R}^2 \to \mathbb{R}$ given by $M\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = x$. Then $M \circ L$ is onto as for each $x \in \mathbb{R}$, $(M \circ L)(x) = x$. But L is not onto because $\begin{bmatrix} 0 \\ 1 \end{bmatrix} \notin \text{range}(L)$. □

Definition. Let V and W be vector spaces. We say a map $L : \mathsf{V} \to \mathsf{W}$ is an **isomorphism** if L is linear, one-to-one, and onto and we say V and W are **isomorphic**.

Example. Determine if the following vector spaces are isomorphic:

(a) $M_{2\times 2}(\mathbb{R})$ and $\mathbb{R}^4$

(b) $\mathbb{R}$ and $\mathbb{R}^3$

Solution.

(a) $M_{2\times 2}(\mathbb{R})$ and $\mathbb{R}^4$ are isomorphic. Consider the following map:

$$L : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}^4 \text{ defined by } L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}.$$

We leave it as an exercise to check that L is an isomorphism.

(b) $\mathbb{R}$ and $\mathbb{R}^3$ are not isomorphic. To show this, we must show that there does not exist an isomorphism from $\mathbb{R}$ to $\mathbb{R}^3$. Suppose that $L : \mathbb{R} \to \mathbb{R}^3$ is an isomorphism. Then, we must have $\ker(L) = \{\vec{0}\}$ and $\text{range}(L) = \mathbb{R}^3$ so $\text{nullity}(L) = 0$ and $\text{rank}(L) = 3$. But then

$$\text{rank}(L) + \text{nullity}(L) = 3 \neq 1 = \dim(\mathbb{R})$$

which contradicts the Rank-Nullity theorem.

□

The next lemma shows that in some sense isomorphic vector spaces have the same bases.

Lemma. Let V and W be vector spaces and let $L : \mathsf{V} \to \mathsf{W}$ be an isomorphism. Then

1. If $\mathcal{B} = \{\vec{v}_1, \cdots \vec{v}_n\}$ is a basis for V then $\mathcal{C} = \{L(\vec{v}_1), \cdots, L(\vec{v}_n)\}$ is a basis for W.
2. If $\mathcal{B} = \{\vec{w}_1, \cdots, \vec{w}_n\}$ is a basis for W, then $\mathcal{C} = \{L^{-1}(\vec{w}_1), \cdots L^{-1}(\vec{w}_n)\}$ is a basis for V.

Proof. Notice that part 2 of the lemma is part with applied to the isomorphism L^{-1} so it is enough to prove part 1. We must check that $\mathcal{C}$ is linearly independent and spans W.
$\mathcal{C}$ is linearly independent.
Suppose $c_1 L(\vec{v}_1) + \cdots + c_n L(\vec{v}_n) = \vec{0}$ for some scalars $c_1, \cdots, c_n \in \mathbb{R}$.
Then $L(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) = \vec{0} \implies c_1\vec{v}_1 + \cdots + c_n\vec{v}_n \in \ker(L) = \{\vec{0}\}$ as L is one-to-one.
Then $c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0}$ so $c_1 = \cdots = c_n = 0$ as $\mathcal{B}$ is linearly independent so $\mathcal{C}$ is linearly independent.
$\mathcal{C}$ spans W.
Let $\vec{w} \in \mathsf{W}$ be given. Then since L is an isomorphism, we can find $\vec{v} \in \mathsf{V}$ with $L(\vec{v}) = \vec{w}$. Since $\mathcal{B}$ is a basis for V we may write $\vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$ for some scalars $a_1, \cdots, a_n \in \mathbb{R}$. Then $\vec{w} = a_1 L(\vec{v}_1) + \cdots + c_n L(\vec{v}_n) \in \operatorname{span}(\mathcal{C})$. □

Theorem. Let V and W be finite-dimensional vector spaces. Then V and W are isomorphic if and only if $\dim(\mathsf{V}) = \dim(\mathsf{W})$.

Proof. ($\Longrightarrow$) follows from the previous lemma and ($\Longleftarrow$) follows from our linear map construction. □

Two Important Isomorphisms

The "Taking Coordinates" Map

This section is a review of some MATH 136 material. We recall the following definition.

Definition. Let V be a vector space and let $\mathcal{B} = \{\vec{v}_1, \cdots, \vec{v}_n\}$ be a basis for V. Let $\vec{v} \in \mathsf{V}$. We define the **$\mathcal{B}$-coordinates of** $\vec{v}$, denoted $[\vec{v}]_{\mathcal{B}}$,by

$$[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

where $a_1, \cdots, a_n \in \mathbb{R}$ are such that $\vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$.

We also have the *taking $\mathcal{B}$-coordinates map* $[\cdot]_{\mathcal{B}} : \mathsf{V} \to \mathbb{R}^n$ defined by mapping $\vec{v}$ to $[\vec{v}]_{\mathcal{B}}$. The next theorem, which we saw in MATH 136, says that this map is linear.

Theorem. Let V be a vector space and let $\mathcal{B} = \{\vec{v}_1, \cdots, \vec{v}_n\}$ be a basis for V. Then, for any $\vec{v}, \vec{w} \in \mathsf{V}$ and $c \in \mathbb{R}$, we have

$$[c\vec{v} + \vec{w}]_{\mathcal{B}} = c[\vec{v}]_{\mathcal{B}} + [\vec{w}]_{\mathcal{B}}.$$

Proof. We may write $\vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$ and $\vec{w} = b_1\vec{v}_1 + \cdots + b_n\vec{v}_n$ for some unique scalars $a_1, \cdots, a_n, b_1, \cdots, b_n \in \mathbb{R}$. Then we have

$$c\vec{v} + \vec{w} = (ca_1 + b_1)\vec{v}_1 + \cdots + (ca_n + b_n)\vec{v}_n$$

and therefore

$$[c\vec{v} + \vec{w}]_{\mathcal{B}} = \begin{bmatrix} ca_1 + b_1 \\ \vdots \\ ca_n + b_n \end{bmatrix} = c\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = c[\vec{v}]_{\mathcal{B}} + [\vec{w}]_{\mathcal{B}}.$$

□

We give an example to review working with $\mathcal{B}$-coordinates.

Example. Let $\mathcal{B} = \left\{ \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ and $\mathcal{C} = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ and define $L : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$[L(\vec{x})]_{\mathcal{C}} = [\vec{x}]_{\mathcal{B}}.$$

a) Find $L\left(\begin{bmatrix} 3 \\ 5 \end{bmatrix}\right)$.

b) Find $L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)$.

Solution.

a) We have $\left[L\left(\begin{bmatrix} 3 \\ 5 \end{bmatrix}\right)\right]_{\mathcal{C}} = \left[\begin{bmatrix} 3 \\ 5 \end{bmatrix}\right]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ since $\begin{bmatrix} 3 \\ 5 \end{bmatrix} = 1\begin{bmatrix} 2 \\ 3 \end{bmatrix} + 1\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and so $L\left(\begin{bmatrix} 3 \\ 5 \end{bmatrix}\right) = 1\begin{bmatrix} 2 \\ 1 \end{bmatrix} + 1\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$.

b) Similarly, we have $\left[L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)\right]_{\mathcal{C}} = \left[\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right]_{\mathcal{B}}$ so we must find $a_1, a_2 \in \mathbb{R}$ such that $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = a_1\begin{bmatrix} 2 \\ 3 \end{bmatrix} + a_2\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$.

We get $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 - x_2 \\ -3x_1 + 2x_2 \end{bmatrix}$.

Thus $\left[L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) \right]_{\mathcal{C}} = \begin{bmatrix} 2x_1 - x_2 \\ -3x_1 + 2x_2 \end{bmatrix}$

Therefore $L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = (2x_1 - x_2) \begin{bmatrix} 2 \\ 1 \end{bmatrix} + (-3x_1 + 2x_2) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ 4x_1 - 4x_2 \end{bmatrix}$.

□

Next we show $[\cdot]_{\mathcal{B}}$ is one-to-one and onto and hence is an isomorphism between the vector spaces V and $\mathbb{R}^n$. This gives an alternate proof of the result that two n-dimensional vector spaces are isomorphic.

Theorem. Let V be a vector space and let $\mathcal{B} = \{\vec{v}_1, \cdots, \vec{v}_n\}$ be a basis for V. Then $[\cdot]_{\mathcal{B}} : \mathsf{V} \to \mathbb{R}^n$ is an isomorphism.

Proof. First we show $[\cdot]_{\mathcal{B}}$ is onto. Let $\vec{x} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n$ be given and set $\vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}$. Then $[\vec{v}]_{\mathcal{B}} = \vec{x}$ which shows $[\cdot]_{\mathcal{B}}$ is onto. Since $\dim(\mathsf{V}) = n = \dim(\mathbb{R}^n)$, we conclude $[\cdot]_{\mathcal{B}}$ is also one-to-one and hence an isomorphism. Another way to show this is to appeal to the Unique Representation Theorem from MATH 136. □

We also recall the change of basis map.
Let $\mathcal{B} = \{\vec{v}_1, \cdots, \vec{v}_n\}$ and $\mathcal{C} = \{\vec{w}_1, \cdots, \vec{w}_n\}$ be bases for a vector space V. The **change of basis map from $\mathcal{B}$-coordinates to $\mathcal{C}$-coordinates**, which we denote ${}_{\mathcal{C}}[I]^{\mathcal{B}}$ is given by

$$ {}_{\mathcal{C}}[I]^{\mathcal{B}} = \begin{bmatrix} [\vec{v}_1]_{\mathcal{C}} & \cdots & [\vec{v}_n]_{\mathcal{C}} \end{bmatrix}. $$

${}_{\mathcal{C}}[I]^{\mathcal{B}}$ has the following properties.

1. ${}_{\mathcal{C}}[I]^{\mathcal{B}}[\vec{v}]_{\mathcal{B}} = [\vec{v}]_{\mathcal{C}}$
2. ${}_{\mathcal{C}}[I]^{\mathcal{B}}{}_{\mathcal{B}}[I]^{\mathcal{C}} = {}_{\mathcal{B}}[I]^{\mathcal{C}}{}_{\mathcal{C}}[I]^{\mathcal{B}} = I$ so in particular ${}_{\mathcal{C}}[I]^{\mathcal{B}}$ is invertible.

Proof. Good midterm review for you. □

Matrix Representation of a Linear Map

Definition. Let V and W be vector spaces and let $L : \mathsf{V} \to \mathsf{W}$ be linear. Let $\mathcal{B} = \{\vec{v}_1, \cdots, \vec{v}_m\}$ be a basis for V and let $\mathcal{C} = \{\vec{w}_1, \cdots, \vec{w}_n\}$ be a basis for W. The **matrix of L with respect to $\mathcal{B}$ and $\mathcal{C}$** is

$$_{\mathcal{C}}\left[L\right]^{\mathcal{B}} = \begin{bmatrix}[L(\vec{v}_1)]_{\mathcal{C}} & \cdots & [L(\vec{v})]_{\mathcal{C}}\end{bmatrix}$$

Proposition. $[L(\vec{x})]_{\mathcal{C}} = {}_{\mathcal{C}}[L]^{\mathcal{B}}[\vec{x}]_{\mathcal{B}}$

Proof. Suppose $\vec{x} = a_1\vec{v}_1 + \cdots + a_m\vec{v}_m$ for some $a_1, \cdots, a_m \in \mathbb{R}$ so $[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$.

Then we have

$$\begin{aligned}
[L(\vec{x})]_{\mathcal{C}} &= [a_1 L(\vec{v_1}) + \cdots + a_m L(\vec{v}_m)]_{\mathcal{C}} \\
&= a_1[L(\vec{v}_1)]_{\mathcal{C}} + \cdots + a_m[L(\vec{v}_m)]_{\mathcal{C}} \\
&= \begin{bmatrix}[L(\vec{v}_1)]_{\mathcal{C}} & \cdots & [L(\vec{v}_m)]_{\mathcal{C}}\end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \\
&= {}_{\mathcal{C}}\left[L\right]^{\mathcal{B}} [\vec{x}]_{\mathcal{B}}
\end{aligned}$$

□

Proposition. Let $\mathsf{U}, \mathsf{V}, \mathsf{W}$ be vector spaces and let β, γ, δ be bases for $\mathsf{U}, \mathsf{V}, \mathsf{W}$ respectively. Let $L : \mathsf{U} \to \mathsf{V}, M : \mathsf{U} \to \mathsf{V}, T : \mathsf{V} \to \mathsf{W}$ be linear maps. Then

1. ${}_{\gamma}[cL + M]^{\beta} = c_{\gamma}[L]^{\beta} +_{\gamma} [M]^{\beta}$
2. ${}_{\delta}[T \circ L]^{\beta} = {}_{\delta}[T]^{\gamma}{}_{\gamma}[L]^{\beta}$

Proof. 1. Let $\vec{v} \in \mathsf{V}$. Then we have

$$[(cL + M)(\vec{v})]_{\gamma} = {}_{\gamma}[cL + M]^{\beta}[\vec{v}]_{\beta}$$

and

$$\begin{aligned}
[(cL + M)(\vec{v})]_{\gamma} &= c[L(\vec{v})]_{\gamma} + [M(\vec{v})]_{\gamma} \\
&= c_{\gamma}[L]^{\beta}[\vec{v}]_{\beta} + {}_{\gamma}[M]^{\beta}[\vec{v}]_{\beta} \\
&= (c_{\gamma}[L]^{\beta} + {}_{\gamma}[M]^{\beta})[\vec{v}]_{\beta}
\end{aligned}$$

Therefore for each $\vec{v} \in \mathsf{V}$ we have ${}_{\gamma}[cL+M]^{\beta}[\vec{v}]_{\beta} = (c_{\gamma}[L]^{\beta} + {}_{\gamma}[M]^{\beta})[\vec{v}]_{\beta}$ and so we conclude that $[cL + M]^{\beta} = c_{\gamma}[L]^{\beta} +_{\gamma} [M]^{\beta}$.
2. Let $\vec{v} \in \mathsf{V}$. Then, by the above proposition,

$$[(M \circ L)(\vec{v})]_{\delta} = {}_{\delta}[M \circ L]^{\beta}[\vec{v}]_{\beta}.$$

But we also have

$$[(M \circ L)(\vec{v})]_\delta = [M(L(\vec{v}))]_\delta = {}_\delta[M]^\gamma [L(\vec{v})]_\gamma = {}_\delta[M]^\gamma{}_\gamma[L]^\beta[\vec{v}]_\beta$$

and hence for all $\vec{v} \in \mathsf{V}$

$${}_\delta[M \circ L]^\beta[\vec{v}]_\beta = {}_\delta[M]^\gamma{}_\gamma[L]^\beta[\vec{v}]_\beta.$$

Therefore ${}_\delta[M \circ L]^\beta = {}_\delta[M]^\gamma{}_\gamma[L]^\beta$. □

We now also have an explicit isomorphism between $\mathcal{L}(\mathsf{V}, \mathsf{V})$, the set of linear operators from V to V and $M_{n\times n}(\mathbb{R})$.

Fix a basis $\mathcal{B} = \{\vec{v}_1, \cdots, \vec{v}_n\}$ for V and define $T : \mathcal{L}(\mathsf{V}, \mathsf{V}) \to M_{n\times n}(\mathbb{R})$ by $T(L) = {}^\mathcal{B}[L]_\mathcal{B}$.

Proposition. T is an isomorphism.

Proof.

1. T is linear. Suppose $L, M \in \mathcal{L}(\mathsf{V}, \mathsf{V})$ and $c \in \mathbb{R}$. Then we have
$$T(cL + M) = {}_\mathcal{B}[cL + M]^\mathcal{B} = c{}_\mathcal{B}[L]^\mathcal{B} + {}_\mathcal{B}[M]^\mathcal{B} = cT(L) + T(M).$$
2. T is one-to-one. We show $\ker(T) = \{\mathbf{0}\}$ where $\mathbf{0} : \mathsf{V} \to \mathsf{V}$ is defined by $\mathbf{0}(\vec{v}) = \vec{0}$. Let $L \in \ker(T)$. Then ${}_\mathcal{B}[L]^\mathcal{B} = 0_{n\times n}$ and so for each $\vec{v} \in \mathsf{V}$, $[L(\vec{v})]_\mathcal{B} = {}_\mathcal{B}[L]^\mathcal{B}[\vec{v}]_\mathcal{B} = \vec{0} \implies L(\vec{v}) = \vec{0}$. Hence $L = \mathbf{0}$ and $\ker(T) = \{\mathbf{0}\}$.
3. T is onto. Let $M = \begin{bmatrix} \vec{c}_1 & \cdots & \vec{c}_n \end{bmatrix}$ be given. Choose $\vec{w}_i$ so that $[\vec{w}_i]_\mathcal{B} = \vec{c}_i$ and define $L : \mathsf{V} \to \mathsf{V}$ by $L(\vec{v}_i) = \vec{w}_i$. Then
$$T(L) = {}_\mathcal{B}[L]^\mathcal{B} = \begin{bmatrix} [L(\vec{v}_1)]_\mathcal{B} & \cdots & [L(\vec{v}_n)]_\mathcal{B} \end{bmatrix} = M$$

□

Example. Find a basis for $\mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$.

Solution. By our work above, we know $\mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ is isomorphic to $M_{2\times 2}(\mathbb{R})$. A basis for $M_{2\times 2}(\mathbb{R})$ is the set

$$\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

and so by the theory we know

$$\{T^{-1}(\vec{v}_1), T^{-1}(\vec{v}_2), T^{-1}(\vec{v}_3), T^{-1}(\vec{v}_4)\}$$

is a basis for $\mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$, where

$$T^{-1}(\vec{v}_1) : \mathbb{R}^2 \to \mathbb{R}^2 \text{ is defined by } T(x, y) = (x, 0)$$
$$T^{-1}(\vec{v}_2) : \mathbb{R}^2 \to \mathbb{R}^2 \text{ is defined by } T(x, y) = (0, x)$$
$$T^{-1}(\vec{v}_3) : \mathbb{R}^2 \to \mathbb{R}^2 \text{ is defined by } T(x, y) = (y, 0)$$
$$T^{-1}(\vec{v}_4) : \mathbb{R}^2 \to \mathbb{R}^2 \text{ is defined by } T(x, y) = (0, y)$$

□

Example. Let π be the plane with basis $\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$. Exhibit an isomorphism between π and P_1.

Solution. First we establish isomorphisms between π and $\mathbb{R}^2$ and P_1 and $\mathbb{R}^2$. It is easy to see that the map $L : P_1 \to \mathbb{R}^2$ given by $L(ax+b) = \begin{bmatrix} a \\ b \end{bmatrix}$ is an isomorphism. The map $M : \pi \to \mathbb{R}^2$ defined by $M\left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $M\left(\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is also an isomorphism.

More explicitly, if $P : a \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + b \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$, then $M(P) = \begin{bmatrix} a \\ b \end{bmatrix}$.

Therefore $L^{-1} \circ M : \pi \to P_1$ is an isomorphism, being the composition of two isomorphisms, which sends P to the polynomial $ax + b$. □

Inner Product Spaces

Definition. Let V be a vector space. An **inner product on** V is a function $\langle , \rangle : \mathsf{V} \times \mathsf{V} \to \mathbb{R}$ satisfying for all $\vec{u}, \vec{v}, \vec{w} \in \mathsf{V}$ and $c \in \mathbb{R}$

1. $\langle \vec{v}, \vec{v} \rangle \geq 0$ and $\langle \vec{v}, \vec{v} \rangle = 0 \iff \vec{v} = \vec{0}$
2. $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$
3. $\langle c\vec{u} + \vec{v}, \vec{w} \rangle = c\langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$

The pair $(\mathsf{V}, \langle , \rangle)$ is called an **inner product space**.

Example. For the vector space $\mathsf{V} = \mathbb{R}^2$, determine if each $\langle , \rangle$ is an inner product on V.

1. $\left\langle \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} a \\ b \end{bmatrix} \right\rangle = xa + 6yb + xb + 2ya$
2. $\left\langle \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} a \\ b \end{bmatrix} \right\rangle = xa + 2yb + \frac{1}{2}xb + \frac{1}{2}ya$

Solution.

1. This is not an inner product on V. We have $\left\langle \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right\rangle = 67$ but $\left\langle \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\rangle = 65$.

2. This is not an inner product on V. We have $\left\langle \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\rangle = -3 < 0$.

$\square$

Lemma. In an inner product space $(\mathsf{V}, \langle , \rangle)$, for each $\vec{v} \in \mathsf{V}$, $\langle \vec{v}, \vec{0} \rangle = 0$.

Definition. Let $(\mathsf{V}, \langle , \rangle)$ be an inner product space. We define the **norm or length** of $\vec{v} \in \mathsf{V}$ by

$$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$$

Proposition (Properties of the Norm)**.** Let $(\mathsf{V}, \langle , \rangle)$ be an inner product space. Then for each $\vec{v}, \vec{w} \in \mathsf{V}$ and $c \in \mathbb{R}$

1. $\|\vec{v}\| \geq 0$ and $\|\vec{v}\| = 0 \iff \vec{v} = \vec{0}$
2. $\|c\vec{v}\| = |c| \, \|\vec{v}\|$
3. $|\langle \vec{v}, \vec{w} \rangle| \leq \|\vec{v}\| \, \|\vec{w}\|$ [Cauchy-Schwarz Inequality]
4. $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$

We also have a generalization of the Pythagorean Theorem.

Theorem (Pythaogrean Theorem)**.** Let $(\mathsf{V}, \langle , \rangle)$ be an inner product space. Then if $\vec{v}, \vec{w} \in \mathsf{V}$ are orthogonal, then

$$\|\vec{v} + \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2 .$$

Proof. We have

$$\begin{aligned} \|\vec{v} + \vec{w}\|^2 &= \langle \vec{v} + \vec{w}, \vec{v} + \vec{w} \rangle \\ &= \langle \vec{v}, \vec{v} \rangle + 2\langle \vec{v}, \vec{w} \rangle + \langle \vec{w}, \vec{w} \rangle \\ &= \|\vec{v}\|^2 + \|\vec{w}\|^2 \end{aligned}$$

$\square$

Definition. Let $(\mathsf{V}, \langle , \rangle)$ be an inner product space.

1. $\vec{v}, \vec{w} \in \mathsf{V}$ are **orthogonal** if $\langle \vec{v}, \vec{w} \rangle = 0$.
2. $\vec{v}$ is a **unit vector** if $\|\vec{v}\| = 1$.
3. $\mathcal{S} \subset \mathsf{V}$ is **orthogonal** if for each $\vec{v} \neq \vec{w} \in \mathcal{S}$, $\langle \vec{v}, \vec{w} \rangle = 0$.
4. $\mathcal{S} \subset \mathsf{V}$ is **orthonormal** if $\mathcal{S}$ is orthogonal and if for each $\vec{v} \in \mathcal{S}$ $\|\vec{v}\| = 1$.
5. $\mathcal{B} = \{\vec{v}_1, \cdots, \vec{v}_n$ is an **orthogonal basis for** V if $\mathcal{B}$ is a basis for V and $\mathcal{B}$ is orthogonal. $\mathcal{B}$ is an **orthonormal basis for** $\mathcal{B}$ if $\mathcal{B}$ is a basis for V and $\mathcal{B}$ is orthonormal.

Proposition. Let $(\mathsf{V}, \langle, \rangle)$ be an inner product space. If $\mathcal{B} = \{\vec{v}_1, \cdots, \vec{v}_n\}$ is orthonormal, then $\mathcal{B}$ is linearly independent.

Proof. Suppose $a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = \vec{0}$ for some $a_1, \cdots, a_n \in \mathbb{R}$. Then for each $1 \leq i \leq n$ we have

$$\begin{aligned}\langle a_1\vec{v}_1 + \cdots + a_n\vec{v}_n, \vec{v}_i\rangle &= a_1\langle \vec{v}_1, \vec{v}_i\rangle + \cdots + a_n\langle \vec{v}_n, \vec{v}_i\rangle \\ &= a_i\langle \vec{v}_i, \vec{v}_i\rangle \\ &= a_i\end{aligned}$$

On the other hand,

$$\langle a_1\vec{v}_1 + \cdots + a_n\vec{v}_n, \vec{v}_i\rangle = \langle \vec{0}, \vec{v}_i\rangle = 0$$

and so $a_i = 0$. Thus $\mathcal{B}$ is linearly independent. □

Proposition. Let $(\mathsf{V}, \langle, \rangle)$ be an inner product space. If $\mathcal{B} = \{\vec{v}_1, \cdots, \vec{v}_n\}$ is an orthonormal basis, then for each $\vec{v} \in \mathsf{V}$, $\vec{v} = \langle \vec{v}, \vec{v}_1\rangle\vec{v}_1 + \cdots + \langle \vec{v}, \vec{v}_n\rangle\vec{v}_n$.

Proof. Since $\mathcal{B}$ is a basis for V, we may write $\vec{v} = a_1\vec{v}_n + \cdots + a_n\vec{v}_n$ for some $c_1, \cdots, c_n \in \mathbb{R}$. It remains to show that $c_i = \langle \vec{v}, \vec{v}_i\rangle$. We have

$$\begin{aligned}\langle \vec{v}, \vec{v}_i\rangle &= \langle a_1\vec{v}_n + \cdots + a_n\vec{v}_n, \vec{v}_i\rangle \\ &= a_1\langle \vec{v}_1, \vec{v}_i\rangle + \cdots + a_n\langle \vec{v}_n, \vec{v}_i\rangle \\ &= a_i\langle \vec{v}_i, \vec{v}_i\rangle \\ &= a_i\end{aligned}$$

as claimed. □

Definition. Let $U \in M_{n\times n}(\mathbb{R})$. U is **orthogonal** if $UU^T = U^TU = I$. That is $U^{-1} = U^T$.

Proposition. The following are equivalent.

1. U is orthogonal.
2. The columns of U form an orthonormal basis for $\mathbb{R}^n$.
3. The rows of U form an orthonormal basis for $\mathbb{R}^n$.

Proof. Let $U = \begin{bmatrix}\vec{v}_1 & \cdots & \vec{v}_n\end{bmatrix}$.
1 $\implies$ 2. Suppose U is orthogonal. Then we have

$$I = U^TU = \begin{bmatrix}\vec{v}_1{}^T \\ \vdots \\ \vec{v}_n{}^T\end{bmatrix}\begin{bmatrix}\vec{v}_1 & \cdots & \vec{v}_n\end{bmatrix}$$

and so $[I]_{ij} = (U^T U)_{ij} = \vec{v}_i^{\,T} \vec{v}_j = \vec{v}_i \cdot \vec{v}_j$. Therefore $\vec{v}_i \cdot \vec{v}_i = 1$ and $\vec{v}_i \cdot \vec{v}_j = 0$ for $i \neq j$. Hence the columns of U are an orthonormal set and hence linearly independent. Therefore, they form a basis, being n linearly independent vectors in an n-dimensional space.
$2 \implies 1$. Suppose the columns of U are an orthonormal basis for $\mathbb{R}^n$. By the same calculation above, we see

$$[U^T U]_{ij} = \vec{v}_i \cdot \vec{v}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

and so $U^T U = I$. By a theorem from MATH 136, $U^{-1} = U^T$.
$2 \iff 3$. We have

$$\begin{aligned} &\text{the columns of } U \text{ form an orthonormal basis} \\ &\iff U \text{ is orthonormal} \\ &\iff U^T \text{ is orthonormal} \\ &\iff \text{the columns of } U^T \text{ form an orthonormal basis} \\ &\iff \text{the rows of } U \text{ form an orthonormal basis} \end{aligned}$$

□

Proposition. Let U be orthogonal. Then for each $\vec{x}, \vec{y} \in \mathbb{R}^n$

1. $\|U\vec{x}\| = \|\vec{x}\|$
2. $U\vec{x} \cdot U\vec{y} = \vec{x} \cdot \vec{y}$

Proof. We have

$$\begin{aligned} \|U\vec{x}\|^2 &= U\vec{x} \cdot U\vec{x} \\ &= (U\vec{x})^T U\vec{x} \\ &= \vec{x}^{\,T} U^T U \vec{x} \\ &= \vec{x}^{\,T} \vec{x} \\ &= \vec{x} \cdot \vec{x} \\ &= \|\vec{x}\|^2 \end{aligned}$$

$\implies \|U\vec{x}\| = \|\vec{x}\|$
2 is similar and omitted. □

Example. Let V be finite-dimensional and let $L : \mathsf{V} \to \mathsf{V}$ be a linear operator. Prove that if $\mathcal{B} = \{\vec{v}_1, \cdots, \vec{v}_n\}$ is an orthonormal basis for V, then the ij-entry of ${}_{\mathcal{B}}[L]^{\mathcal{B}}$ is $\langle L(\vec{v}_j), \vec{v}_i \rangle$.

Solution. Recall that ${}_{\mathcal{B}}[L]^{\mathcal{B}} = \begin{bmatrix} [L(\vec{v})_1]_{\mathcal{B}} & \cdots & [L(\vec{v}_n)_{\mathcal{B}}] \end{bmatrix}$. So ij-entry is the ith element of $[L(\vec{v}_j)]_{\mathcal{B}}$. But by our work above since $\mathcal{B}$ is orthonormal

$$[L(\vec{v}_j)]_{\mathcal{B}} = \begin{bmatrix} \langle L(\vec{v}_j), \vec{v}_1 \rangle \\ \vdots \\ \langle L(\vec{v}_j), \vec{v}_n \rangle \end{bmatrix}.$$

□

Example. Let $T : \mathsf{V} \to \mathsf{V}$ be a linear operator on the vector space V such that for all $\vec{x}, \vec{y} \in \mathsf{V}$, $\langle T(\vec{x}), T(\vec{y}) \rangle = \langle \vec{x}, \vec{y} \rangle$ (T is called an **isometry**). Show that T is an isomorphism.

Solution. Suppose $T(\vec{x}) = T(\vec{y})$. Then we have

$$\begin{aligned} \langle \vec{x} - \vec{y}, \vec{x} - \vec{y} \rangle &= \langle T(\vec{x} - \vec{y}), T(\vec{x} - \vec{y}) \rangle \\ &= \langle \vec{0}, \vec{0} \rangle \\ &= 0 \end{aligned}$$

and so $\vec{x} - \vec{y} = \vec{0}$, which shows that T is one-to-one. Since T is a linear operator, T is one-to-one $\iff$ T is onto. Therefore T is an isomorphism. □