Math 350 Fall 2011 Worksheet 2

Name:

The goal of this worksheet is to help you understand why every linear map between two finite dimensional vector spaces can be represented by a matrix. All vector spaces in this worksheet will be finite dimensional.

To begin with, suppose V is a vector space, and $\{v_1, \ldots, v_n\}$ is a basis of V. We are going to call v_1 the *first* basis vector in this basis, v_2 the *second* vector in this basis, and so on, making $\{v_1, \ldots, v_n\}$ an *ordered basis*[1] of V (i.e. a basis of V with an ordering of the basis vectors). An ordered basis is usually denoted by Greek letters (like α, β, γ) in the book, and we adopt the same convention.

Example 1. *Let $V = \mathbb{R}^2$ and $\alpha = \{(1,0),(0,1)\}$. Then α is an ordered basis of V, where $(1,0)$ is the first basis vector, $(0,1)$ is the second basis vector.*

Example 2. *Let $V = \mathbb{R}^2$ and $\beta = \{(1,1),(0,1)\}$. Then β is another ordered basis of V, which is different from α.*

Example 3. *Let $V = \mathbb{R}^2$ and $\gamma = \{(0,1),(1,0)\}$. Then γ is also an ordered basis of V, which is also different from α. Reason: The first basis vector in γ is $(0,1)$, which is not the same as the first basis vector in α (which was $(1,0)$).*

Now suppose V is a vector space with an ordered basis $\alpha = \{v_1, \ldots, v_n\}$. Then every vector in V can be written uniquely as $v = a_1 v_1 + \cdots + a_n v_n$ for some scalars $a_1, \ldots, a_n$. We define the coordinate vector of v with respect to α to be

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

This is usually denoted by $[v]_\alpha$.

Question 1. *Suppose $V = \mathbb{R}^2$ and $v = (3,7) \in V$. Then if α, β, γ are the ordered bases as in the previous examples, check that*

$$[v]_\alpha = \begin{pmatrix} 3 \\ 7 \end{pmatrix}, \quad [v]_\gamma = \begin{pmatrix} 7 \\ 3 \end{pmatrix}, \quad \textit{and} \quad [v]_\beta = \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

[1] Technically, to indicate that there is an ordering of the elements in say α above, it is better to write $\alpha = ((1,0),(0,1))$ with a pair of paratheses ( ) rather than a pair of braces { }. But since this is the notation adopted in the book, we will follow this notation.

In fact, if you weren't told the answer that $[v]_\beta = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ in the above example, one way of figuring this out is the following. Suppose $[v]_\beta = \begin{pmatrix} a \\ b \end{pmatrix}$ for some real numbers a and b that you want to determine. Then

$$(3,7) = a(1,1) + b(0,1)$$

(why?), so

$$(3,7) = (a, a+b),$$

i.e.

$$\begin{cases} a = 3 \\ a + b = 7 \end{cases}.$$

From here you can solve for a and b and show that $a = 3$, $b = 4$.

Question 2. *Suppose $V = \mathbb{R}^2$ and $v = (7, -1) \in V$. Let α, β and γ be the ordered basis of V as above. Find $[v]_\alpha$, $[v]_\beta$ and $[v]_\gamma$.*

Question 3. *Suppose V is the vector space of all polynomials on $\mathbb{R}$ of degree at most 3. Then $\alpha := \{1, x, x^2, x^3\}$ is an ordered basis of V. Let $p(x) = 3x^3 + 4x^2 - 1$ and $q(x) = (x+1)^2$ so that $p(x), q(x) \in V$. What is $[p(x)]_\alpha$? What is $[q(x)]_\alpha$?*

Question 4. *Suppose V and α are as in the previous question, and $r(x) \in V$ is such that $[r(x)]_\alpha = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 3 \end{pmatrix}$. What is $r(x)$?*

Now we move on to representing linear maps by matrices. Let's begin with an explicit example.

Question 5. *Suppose V is the vector space of all real polynomials of degree at most 3 over $\mathbb{R}$, and $T\colon V \to V$ is the linear map*

$$T(p(x)) = p'(x)$$

for all $p(x) \in V$. Let $\beta = \{x^3, x^2, x, 1\}$ be an ordered basis of V.

(a) Suppose $q(x) \in V$ and $[q(x)]_\beta = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 3 \end{pmatrix}$.

(i) What is $q(x)$?

(ii) Compute $T(q(x))$.

(iii) What is $[T(q(x))]_\beta$?

(b) More generally, suppose $p(x) \in V$ is such that $[p(x)]_\beta = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$. Can you compute $[T(p(x))]_\beta$?

From the above question, we see that for any $p(x) \in V$, if $[p(x)]_\beta = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$, then we have

$$[T(p(x))]_\beta = a\begin{pmatrix} 0 \\ 3 \\ 0 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \end{pmatrix} + c\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + d\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$$

It follows that for any $p(x) \in V$, we have

$$[T(p(x))]_\beta = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} [p(x)]_\beta.$$

We therefore say that the linear map $T\colon V \to V$ is represented, with respect to the basis β of V, by the matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

In general, we have the following theorem:

Theorem 1. *Suppose V and W are vector spaces over some field F, with* $\dim V = n$ *and* $\dim W = m$. *Suppose $T\colon V \to W$ is a linear map. If β is an ordered basis for V and γ is an ordered basis for W, then there exists a unique $m \times n$ matrix A such that*

$$[T(v)]_\gamma = A[v]_\beta \quad \textit{for all } v \in V.$$

The matrix A is usually written $[T]_\beta^\gamma$, and we call $[T]_\beta^\gamma$ a matrix representation of T (with respect to the ordered bases β and γ).

Question 6. *Let P_n be the vector space of all real polynomials of degree at most n over $\mathbb{R}$. Let $V = P_3$ and $W = P_2$. Let $T\colon V \to W$ be the linear map*

$$T(p(x)) = 2p'(x) + p''(x).$$

Let $\beta = \{1, x, x^2, x^3\}$ be an ordered basis of V, and $\gamma = \{1, x, x^2\}$ be an ordered basis of W. Find $[T]_\beta^\gamma$.

Conversely, if V and W are as in the theorem above, and β, γ are ordered bases of V and W respectively, then for any $m \times n$ matrix A, then one can *define* a linear map $T: V \to W$ such that

$$[T(v)]_\gamma = A[v]_\beta \quad \text{for all } v \in V.$$

In other words, there is a linear operator $T: V \to W$ with $[T]_\beta^\gamma = A$.

Question 7. *Let P_n be as in the previous question. Suppose $V = P_2$ and $W = P_3$. Let $\beta := \{1, x, x^2\}$ and $\gamma := \{1, x, x^2, x^3\}$ be ordered bases of V and W respectively. Let A be the matrix*

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{pmatrix}.$$

Let $T: V \to W$ be a map defined by $[T(v)]_\gamma = A[v]_\beta$ for all $v \in V$. Show that

$$T(p(x)) = \int_0^x p(t)dt$$

for all $p(x) \in V$, and that T is linear.

From the discussion above we see that there is a correspondence between linear maps and matrices. Now if you are given two linear maps (with the correct domains and targets), you can *compose* them, and you still get a linear map; if you are given two matrices (of the correct sizes), you can *multiply* them. It turns out even this composition and multiplication corresponds to each other.

Theorem 2. *Suppose $T_1 \colon U \to V$ and $T_2 \colon V \to W$ are linear maps between vector spaces. Let α be an ordered basis of U, β be an ordered basis of V and γ be an ordered basis of W. Then*

$$[T_2 \circ T_1]_\alpha^\gamma = [T_2]_\beta^\gamma [T_1]_\alpha^\beta.$$

In other words, the matrix representation of the linear map $T_2 \circ T_1 \colon U \to W$ is the product of the matrix representations of T_2 and T_1 (taken with respect to the respective bases).

Note that one has to keep the order of the matrix multiplication on the right hand side here, since matrix multiplications do not commute.

In fact matrix multiplication was *defined* so that this theorem holds; this is the real reason why matrices have to be multiplied in the peculiar way that you have learned.

The proof of this theorem is very easy: if $u \in U$, then

$$[(T_2 \circ T_1)(u)]_\gamma = [T_2(T_1(u))]_\gamma = [T_2]_\beta^\gamma [T_1(u)]_\beta = [T_2]_\beta^\gamma [T_1]_\alpha^\beta [u]_\alpha$$

so by the uniqueness assertion in Theorem 1, one has

$$[T_2 \circ T_1]_\alpha^\gamma = [T_2]_\beta^\gamma [T_1]_\alpha^\beta.$$

If $T \colon V \to W$ is a linear map between two vector spaces and β, γ are ordered bases of V and W respectively, then the matrix $[T]_\beta^\gamma$ obviously depends on β and γ.

Question 8. *For example, consider the identity map $Id \colon \mathbb{R}^2 \to \mathbb{R}^2$. Let $\beta = \{(1,3),(0,2)\}$ be an ordered basis of $\mathbb{R}^2$, and $\gamma = \{(1,0),(0,1)\}$ be the standard ordered basis of $\mathbb{R}^2$.*

(i) Can you compute $[Id]_\gamma^\gamma$?

(Remark: In fact, if V is any vector space and γ is any ordered basis of V, then $[Id]_\gamma^\gamma$ is the identity matrix of the appropriate size.)

(ii) What about $[Id]_\beta^\gamma$?

(iii) What about $[Id]_\gamma^\beta$*?*

In general, if V is a vector space, and β, $\tilde{\beta}$ are two ordered bases of V, then the matrix $[Id_V]_\beta^{\tilde{\beta}}$ is called the *change of coordinate matrix* from β to $\tilde{\beta}$, where $Id_V : V \to V$ is the identity map on V.

Question 9. *Let* V *be the vector space of all real polynomials of degree at most* 3 *over* $\mathbb{R}$*. Let* $\beta = \{1, x, x^2, x^3\}$ *and* $\tilde{\beta} = \{1 + x, x + x^2, x^2 + x^3, 1 + x + x^2 + x^3\}$ *be two ordered basis of* V*. What is the change of coordinate matrix from* β *to* $\tilde{\beta}$*?*

The important theorem about the change of coordinate matrices is the following:

Theorem 3. *Suppose $T\colon V \to W$ is a linear map. Suppose β, $\tilde{\beta}$ are two ordered bases of V, and γ, $\tilde{\gamma}$ are two ordered bases of W. Then*

$$[T]_{\tilde{\beta}}^{\tilde{\gamma}} = [Id_W]_{\gamma}^{\tilde{\gamma}}[T]_{\beta}^{\gamma}[Id_V]_{\tilde{\beta}}^{\beta}.$$

This is in fact just a direct application of Theorem 2, once one notes that

$$T = Id_W \circ T \circ Id_V.$$

Let's see an example of an application of this theorem. Let $V = P_3$, $W = P_2$ and $T\colon V \to W$ be given by $T(p(x)) = 2p'(x) + p''(x)$ as in Question 6. Let $\beta = \{1, x, x^2, x^3\}$ be an ordered basis of V, and $\gamma = \{1, x, x^2\}$ be an ordered basis of W. We have computed $[T]_{\beta}^{\gamma}$ in Question 6. If now we want to compute $[T]_{\tilde{\beta}}^{\gamma}$ instead, where $\tilde{\beta}$ is the ordered basis of V in Question 9, then one just needs to apply the previous theorem; in particular,

$$[T]_{\tilde{\beta}}^{\gamma} = (\text{answer of Q.6})(\text{answer of Q.9}).$$

Next, we have the following theorem:

Theorem 4. *If V is a vector space, with two ordered basis β and $\tilde{\beta}$, and if B is the change of coordinate matrix from β to $\tilde{\beta}$, then B is invertible, and the change of coordinate matrix from $\tilde{\beta}$ to β is given by B^{-1}.*

Indeed, suppose the change of coordinate matrix from $\tilde{\beta}$ to β is a matrix C. Then

$$BC = [Id_V]_{\beta}^{\tilde{\beta}}[Id_V]_{\tilde{\beta}}^{\beta} = [Id_V]_{\tilde{\beta}}^{\tilde{\beta}} = I$$

where I is the identity matrix of the appropriate size. The second equality is an application of Theorem 2, and the last equality follows from the remark after Question 8(i). Similarly, one can show that $CB = I$. Thus

$$BC = CB = I,$$

which says B is invertible and $C = B^{-1}$, as desired.

A piece of convention: If $T\colon V \to V$ is a linear map mapping a vector space V into itself, and if β is an ordered basis of V, then $[T]_{\beta}^{\beta}$ is usually abbreviated as $[T]_{\beta}$.

From the above discussion, we see that

Theorem 5. *Suppose $T\colon V \to V$ is a linear map, and β, $\tilde{\beta}$ are two ordered bases of V. If B is the change of coordinate matrix from β to $\tilde{\beta}$, then*

$$[T]_{\tilde{\beta}} = B[T]_{\beta}B^{-1}.$$

Two $n \times n$ matrices M and N are said to be similar, if there is an invertible matrix B such that $M = BNB^{-1}$. Thus if $T\colon V \to V$ is a linear map of V into itself, and if β, $\tilde{\beta}$ are two ordered bases of V, then $[T]_{\beta}$ and $[T]_{\tilde{\beta}}$ are similar to each other.

Finally, we have the following theorem:

Theorem 6. *Suppose V and W are vector spaces over some field F of dimensions n and m respectively. The set of all linear maps from V to W is a vector space of dimension mn, and is in fact isomorphic to the vector space of all $m \times n$ matrices with entries in F via the isomorphism*

$$T \mapsto [T]_\beta^\gamma,$$

where β, γ are any ordered bases of V and W respectively.

See Theorem 2.20 of the book and its Corollary.

We close by mentioning the following. We have been studying linear maps in this course, and we have seen a correspondence between matrices and linear maps in the above discussion. It turns out that this correspondence can be used, in conjunction with our understanding of linear maps, to understand matrices. For instance, if M is a $m \times n$ real matrix, then there is an associated linear map $T_M \colon \mathbb{R}^n \to \mathbb{R}^m$ given by

$$T_M(v) = Mv \quad \text{for all } v \in \mathbb{R}^n.$$

The rank of the matrix M is then *defined* to be the rank of this linear map T_M; the nullity of the matrix M is then *defined* to be the nullity of this linear map T_M. It follows that the rank of a matrix is the dimension of its column space, and the nullity of a matrix is the dimension of its nullspace. The rank-nullity theorem for linear maps then says for any $m \times n$ matrix M, one has

$$\text{rank}(M) + \text{nullity}(M) = n. \tag{1}$$

In particular, if $m = n$, i.e. if M is an $n \times n$ matrix, then the following are equivalent:

(a) The columns of M are linearly independent vectors in $\mathbb{R}^n$;
(b) The system of equations $Mx = 0$ has only the trivial solution $x = 0$;
(c) The nullity of M is 0;
(d) The rank of M is n;
(e) The columns of M span $\mathbb{R}^n$;
(f) The map $x \in \mathbb{R}^n \mapsto Mx \in \mathbb{R}^n$ is bijective;
(g) The matrix M is invertible.

In fact, it is easy to see, from definitions, that (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (c), and that (d) $\Leftrightarrow$ (e). The equivalence of (c) and (d) follows from (1). Thus (a) through (e) are equivalent; then their equivalence with (f) and (g) follows. We have seen a proof of the equivalence of (a) and (e) already in Chapter 1; the above provides a second proof of this equivalence, via the rank-nullity theorem.