Economics 204 Summer/Fall 2016
Lecture 9–Thursday August 4, 2016

Section 3.3. Quotient Vector Spaces[1]

Given a vector space X over a field F and a vector subspace W of X, define an equivalence relation by

$$x \sim y \iff x - y \in W$$

Form a new vector space X/W: the set of elements of X/W is

$$\{[x] : x \in X\}$$

where $[x]$ denotes the equivalence class of x with respect to $\sim$. X/W is read "X mod W". Note that the vectors in X/W are *sets* of vectors in X: for $x \in X$,

$$[x] = \{x + w : w \in W\}$$

We claim that X/W can be viewed as a vector space over F. Define the vector space operations $+, \cdot$ in X/W as follows:

$$\begin{aligned} [x] + [y] &= [x + y] \\ \alpha[x] &= [\alpha x] \end{aligned}$$

The exercise below asks you to verify that these operations are well-defined. Then X/W is a vector space over F with these definitions for $+$ and $\cdot$.

Exercise: Verify that $\sim$ above is an equivalence relation and that vector addition and scalar multiplication are well-defined, i.e.

$$\begin{aligned} [x] = [x'], [y] = [y'] &\Rightarrow [x + y] = [x' + y'] \\ [x] = [x'], \alpha \in F &\Rightarrow [\alpha x] = [\alpha x'] \end{aligned}$$

Example: Let $X = \mathbf{R}^3$ and let $W = \{x \in \mathbf{R}^3 : x_1 = x_2 = 0\}$. Then for $x, y \in \mathbf{R}^3$,

$$\begin{aligned} x \sim y &\iff x - y \in W \\ &\iff x_1 - y_1 = 0, x_2 - y_2 = 0 \\ &\iff x_1 = y_1, x_2 = y_2 \end{aligned}$$

and

$$[x] = \{x + w : w \in W\} = \{(x_1, x_2, z) : z \in \mathbf{R}\}$$

So the equivalence class corresponding to x is the line in $\mathbf{R}^3$ through x parallel to the axis of the third coordinate. See Figure 1. What is X/W? Intuitively this equivalence relation ignores the third coordinate, and we can identify the equivalence class $[x]$ with the vector $(x_1, x_2) \in \mathbf{R}^2$. The next two results show how to formalize this connection.

[1]The first part of this material is not in de la Fuente.

Theorem 1 *If X is a vector space with $\dim X = n$ for some $n \in \mathbf{N}$ and W is a vector subspace of X, then*

$$\dim(X/W) = \dim X - \dim W$$

Proof: (Sketch) Begin with a basis $\{w_1, \ldots, w_c\}$ for W, and a basis $\{[x_1], \ldots, [x_k]\}$ for X/W. Show that

$$\{w_1, \ldots, w_c\} \cup \{x_1, \ldots, x_k\}$$

is a basis for X. ■

Theorem 2 *Let X and Y be vector spaces over the same field F and $T \in L(X, Y)$. Then $\operatorname{Im} T$ is isomorphic to $X/\ker T$.*

Proof: Notice that if X is finite-dimensional, then

$$\begin{aligned}
\dim(X/\ker T) &= \dim X - \dim\ker T \quad \text{(by the previous theorem)} \\
&= \operatorname{Rank} T \quad \text{(by the Rank-Nullity Theorem)} \\
&= \dim \operatorname{Im} T
\end{aligned}$$

so $X/\ker T$ is isomorphic to $\operatorname{Im} T$.

We prove that this is true in general, and that the isomorphism is natural.

Define $\tilde{T} : X/\ker T \to \operatorname{Im} T$ by

$$\tilde{T}([x]) = T(x)$$

We first need to check that this is well-defined, that is, that if $[x] = [x']$ then $\tilde{T}([x]) = \tilde{T}([x'])$.

$$\begin{aligned}
[x] = [x'] &\Rightarrow x \sim x' \\
&\Rightarrow x - x' \in \ker T \\
&\Rightarrow T(x - x') = 0 \\
&\Rightarrow T(x) = T(x')
\end{aligned}$$

so $\tilde{T}$ is well-defined.

Clearly, $\tilde{T} : X/\ker T \to \operatorname{Im} T$. It is easy to check that $\tilde{T}$ is linear, so $\tilde{T} \in L(X/\ker T, \operatorname{Im} T)$. Next we show that $\tilde{T}$ is an isomorphism.

$$\begin{aligned}
\tilde{T}([x]) = \tilde{T}([y]) &\Rightarrow T(x) = T(y) \\
&\Rightarrow T(x - y) = 0 \\
&\Rightarrow x - y \in \ker T \\
&\Rightarrow x \sim y \\
&\Rightarrow [x] = [y]
\end{aligned}$$

so $\tilde{T}$ is one-to-one.

$$\begin{aligned} y \in \operatorname{Im} T &\Rightarrow \exists x \in X \text{ s.t. } T(x) = y \\ &\Rightarrow \tilde{T}([x]) = y \end{aligned}$$

so $\tilde{T}$ is onto, hence $\tilde{T}$ is an isomorphism. ■

Example: Consider $T \in L(\mathbf{R}^3, \mathbf{R}^2)$ defined by

$$T(x_1, x_2, x_3) = (x_1, x_2)$$

Then $\ker T = \{x \in \mathbf{R}^3 : x_1 = x_2 = 0\}$ is the x_3-axis. (Also notice $\ker T = W$ from the previous example.)

Given x, the equivalence class $[(x_1, x_2, x_3)]$ is just the line through x parallel to the x_3-axis. $\tilde{T}([x]) = T(x_1, x_2, x_3) = (x_1, x_2)$.

$$\operatorname{Im} T = \mathbf{R}^2, \quad X/\ker T \cong \mathbf{R}^2 = \operatorname{Im} T$$

as we suggested intuitively above (here the symbol $\cong$ denotes isomorphism, that is, we write $Y \cong Z$ if Y and Z are isomorphic.)

Every real vector space X with dimension n is isomorphic to $\mathbf{R}^n$. What's the isomorphism?

Let X be a finite-dimensional vector space over $\mathbf{R}$ with $\dim X = n$. Fix any Hamel basis $V = \{v_1, \ldots, v_n\}$ of X. Any $x \in X$ has a unique representation

$$x = \sum_{j=1}^{n} \beta_j v_j$$

(here, we allow $\beta_j = 0$). (Generally, vectors are represented as column vectors, not row vectors.) Then given the representation of x above, we write

$$crd_V(x) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \in \mathbf{R}^n$$

That is, $crd_V(x)$ is the vector of coordinates of x with respect to the basis V.

$$crd_V(v_1) = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \quad crd_V(v_2) = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \quad crd_V(v_n) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

crd_V is an isomorphism from X to $\mathbf{R}^n$.

Matrix Representation of a Linear Transformation

Suppose $T \in L(X,Y)$, $\dim X = n$ and $\dim Y = m$. Fix bases

$$\begin{array}{rcl} V & = & \{v_1, \ldots, v_n\} \text{ of } X \\ W & = & \{w_1, \ldots, w_m\} \text{ of } Y \end{array}$$

$T(v_j) \in Y$, so

$$T(v_j) = \sum_{i=1}^{m} \alpha_{ij} w_i$$

Define

$$Mtx_{W,V}(T) = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix}$$

Notice that the columns are the coordinates (expressed with respect to W) of $T(v_1), \ldots, T(v_n)$.

Observe

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{m1} \end{pmatrix}$$

so

$$\begin{array}{rcl} Mtx_{W,V}(T) \cdot crd_V(v_j) & = & crd_W(T(v_j)) \\ Mtx_{W,V}(T) \cdot crd_V(x) & = & crd_W(T(x)) \ \ \forall x \in X \end{array}$$

Multiplying a vector by a matrix does two things:

- Computes the action of T
- Accounts for the change in basis

Example: $X = Y = \mathbf{R}^2$, $V = \{(1,0),(0,1)\}$, $W = \{(1,1),(-1,1)\}$, $T = id$, that is, $T(x) = x$ for all x.

$$Mtx_{W,V}(T) \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$Mtx_{W,V}(T)$ is the matrix that *changes basis* from V to W. How do we compute it?

$$\begin{array}{rcl} v_1 = (1,0) & = & \alpha_{11}(1,1) + \alpha_{21}(-1,1) \\ \alpha_{11} - \alpha_{21} & = & 1 \\ \alpha_{11} + \alpha_{21} & = & 0 \\ 2\alpha_{11} & = & 1, \alpha_{11} = \dfrac{1}{2} \end{array}$$

$$
\begin{aligned}
\alpha_{21} &= -\frac{1}{2} \\
v_2 = (0,1) &= \alpha_{12}(1,1) + \alpha_{22}(-1,1) \\
\alpha_{12} - \alpha_{22} &= 0 \\
\alpha_{12} + \alpha_{22} &= 1 \\
2\alpha_{12} &= 1, \alpha_{12} = \frac{1}{2} \\
\alpha_{22} &= \frac{1}{2}
\end{aligned}
$$

$$
Mtx_{W,V}(id) = \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix}
$$

Theorem 3 (Thm. 3.5') *Let X and Y be vector spaces over the same field F, with $\dim X = n$, $\dim Y = m$. Then $L(X,Y)$, the space of linear transformations from X to Y, is isomorphic to $F_{m\times n}$, the vector space of $m \times n$ matrices over F. If $V = \{v_1, \ldots, v_n\}$ is a basis for X and $W = \{w_1, \ldots, w_m\}$ is a basis for Y, then*

$$
Mtx_{W,V} \in L(L(X,Y), F_{m\times n})
$$

and $Mtx_{W,V}$ is an isomorphism from $L(X,Y)$ to $F_{m\times n}$.

Theorem 4 (From Handout) *Let X, Y, Z be finite-dimensional vector spaces over the same field F with bases U, V, W respectively. Let $S \in L(X,Y)$ and $T \in L(Y,Z)$. Then*

$$
Mtx_{W,V}(T) \cdot Mtx_{V,U}(S) = Mtx_{W,U}(T \circ S)
$$

i.e. matrix multiplication corresponds via the matrix representation isomorphism to composition of linear transformations.

Proof: See handout. ■

Note that $Mtx_{W,V}$ is a function from $L(X,Y)$ to the space $F_{m\times n}$ of $m \times n$ matrices, while $Mtx_{W,V}(T)$ is an $m \times n$ matrix.

The theorem can be summarized by the following "Commutative Diagram:"

$$
\begin{tikzcd}
X \arrow[r, "S"] \arrow[d, leftrightarrow, "crd_U"'] & Y \arrow[r, "T"] \arrow[d, leftrightarrow, "crd_V"] & Z \arrow[d, leftrightarrow, "crd_W"] \\
\mathbf{R}^n \arrow[r, "Mtx_{V,U}(S)"'] & \mathbf{R}^m \arrow[r, "Mtx_{W,V}(T)"'] & \mathbf{R}^r
\end{tikzcd}
$$

We say the diagram commutes because you get the same answer any way you go around the diagram (in directions allowed by the arrows). The *crd* arrows go in both directions because *crd* is an isomorphism.

Section 3.5. Change of Basis and Similarity

Let X be a finite-dimensional vector space with basis V. If $T \in L(X, X)$ it is customary to use the same basis in the domain and range. In this case,

$$Mtx_V(T) \text{ denotes } Mtx_{V,V}(T)$$

Question: If W is another basis for X, how are $Mtx_V(T)$ and $Mtx_W(T)$ related?

$$\begin{aligned} Mtx_{V,W}(id) \cdot Mtx_W(T) \cdot Mtx_{W,V}(id) &= Mtx_{V,W}(id) \cdot Mtx_{W,V}(T \circ id) \\ &= Mtx_{V,V}(id \circ T \circ id) \\ &= Mtx_V(T) \end{aligned}$$

and

$$\begin{aligned} Mtx_{V,W}(id) \cdot Mtx_{W,V}(id) &= Mtx_{V,V}(id) \\ &= \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \end{aligned}$$

So this says that

$$Mtx_V(T) = P^{-1} Mtx_W(T) P$$

for the invertible matrix

$$P = Mtx_{W,V}(id)$$

that is the change of basis matrix. On the other hand, if P is any invertible matrix, then P is also a change of basis matrix for appropriate corresponding bases (see handout).

Definition 5 Square matrices A and B are *similar* if

$$A = P^{-1} B P$$

for some invertible matrix P.

Theorem 6 *Suppose that X is a finite-dimensional vector space.*

1. *If $T \in L(X, X)$ then any two matrix representations of T are similar. That is, if U, W are any two bases of X, then $Mtx_W(T)$ and $Mtx_U(T)$ are similar.*
2. *Conversely, two similar matrices represent the same linear transformation T, relative to suitable bases. That is, given similar matrices A, B with $A = P^{-1}BP$ and any basis U, there is a basis W and $T \in L(X, X)$ such that*

$$\begin{aligned} B &= Mtx_U(T) \\ A &= Mtx_W(T) \\ P &= Mtx_{U,W}(id) \\ P^{-1} &= Mtx_{W,U}(id) \end{aligned}$$

Proof: See Handout on Diagonalization and Quadratic Forms. ■

Section 3.6. Eigenvalues and Eigenvectors

Here, we define eigenvalues and eigenvectors of a linear transformation and show that λ is an eigenvalue of T if and only if λ is an eigenvalue for some matrix representation of T if and only if λ is an eigenvalue for every matrix representation of T.

Definition 7 Let X be a vector space and $T \in L(X, X)$. We say that λ is an *eigenvalue* of T and $v \neq 0$ is an *eigenvector corresponding to* λ if $T(v) = \lambda v$.

Theorem 8 (Theorem 4 in Handout) *Let X be a finite-dimensional vector space, and U a basis. Then λ is an eigenvalue of T if and only if λ is an eigenvalue of $Mtx_U(T)$. v is an eigenvector of T corresponding to λ if and only if $crd_U(v)$ is an eigenvector of $Mtx_U(T)$ corresponding to λ.*

Proof: By the Commutative Diagram Theorem,

$$\begin{aligned} T(v) = \lambda v &\Leftrightarrow crd_U(T(v)) = crd_U(\lambda v) \\ &\Leftrightarrow Mtx_U(T)(crd_U(v)) = \lambda(crd_U(v)) \end{aligned}$$

■

Computing eigenvalues and eigenvectors:

Suppose $\dim X = n$; let I be the $n \times n$ identity matrix. Given $T \in L(X, X)$, fix a basis U and let

$$A = Mtx_U(T)$$

Find the eigenvalues of T by computing the eigenvalues of A:

$$\begin{aligned} Av = \lambda v &\iff (A - \lambda I)v = 0 \\ &\iff (A - \lambda I) \text{ is not invertible} \\ &\iff \det(A - \lambda I) = 0 \end{aligned}$$

We have the following facts:

- If $A \in \mathbf{R}_{n\times n}$,
$$f(\lambda) = \det(A - \lambda I)$$
is an n^{th} degree polynomial in λ with real coefficients; it is called the *characteristic polynomial* of A.

- f has n roots in $\mathbf{C}$, counting multiplicity:
$$f(\lambda) = (\lambda - c_1)(\lambda - c_2)\cdots(\lambda - c_n)$$
where $c_1, \ldots, c_n \in \mathbf{C}$ are the eigenvalues; the c_j's are not necessarily distinct. Notice that $f(\lambda) = 0$ if and only if $\lambda \in \{c_1, \ldots, c_n\}$, so the roots are the solutions of the equation $f(\lambda) = 0$.

- the roots that are not real come in conjugate pairs:
$$f(a + bi) = 0 \Leftrightarrow f(a - bi) = 0$$

- if $\lambda = c_j \in \mathbf{R}$, there is a corresponding eigenvector in $\mathbf{R}^n$.

- if $\lambda = c_j \notin \mathbf{R}$, the corresponding eigenvectors are in $\mathbf{C}^n \setminus \mathbf{R}^n$.

Diagonalization

Definition 9 Suppose X is a finite-dimensional vector space with basis U. Given a linear transformation $T \in L(X, X)$, let
$$A = Mtx_U(T)$$
We say that A *can be diagonalized* (or is *diagonalizable*) if there is a basis W for X such that $Mtx_W(T)$ is diagonal, i.e.
$$Mtx_W(T) = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

Notice that the eigenvectors of $Mtx_W(T)$ are exactly the standard basis vectors of $\mathbf{R}^n$. But w_j is an eigenvector of T corresponding to λ_j if and only if $crd_W(w_j)$ is an eigenvector of $Mtx_W(T)$, and $crd_W(w_j)$ is the j^{th} standard basis vector of $\mathbf{R}^n$, so $W = \{w_1, \ldots, w_n\}$ where w_j is an eigenvector corresponding to λ_j.

Then the action of T is clear: it stretches each basis element w_i by the factor λ_i.

Theorem 10 (Thm. 6.7') *Let X be an n-dimensional vector space, $T \in L(X, X)$, U any basis of X, and $A = Mtx_U(T)$. Then the following are equivalent:*

1. *A can be diagonalized*
2. *there is a basis W for X consisting of eigenvectors of T*
3. *there is a basis V for $\mathbf{R}^n$ consisting of eigenvectors of A*

Proof: Follows from Theorem 6.7 in de la Fuente and Theorem 4 from the Handout. ∎

Theorem 11 (Thm. 6.8') *Let X be a vector space and $T \in L(X, X)$.*

1. *If $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T with corresponding eigenvectors $v_1, \ldots, v_m$, then $\{v_1, \ldots, v_m\}$ is linearly independent.*
2. *If $\dim X = n$ and T has n distinct eigenvalues, then X has a basis consisting of eigenvectors of T; consequently, if U is any basis of X, then $Mtx_U(T)$ is diagonalizable.*

Proof: This is an adaptation of the proof of Theorem 6.8 in de la Fuente. ∎

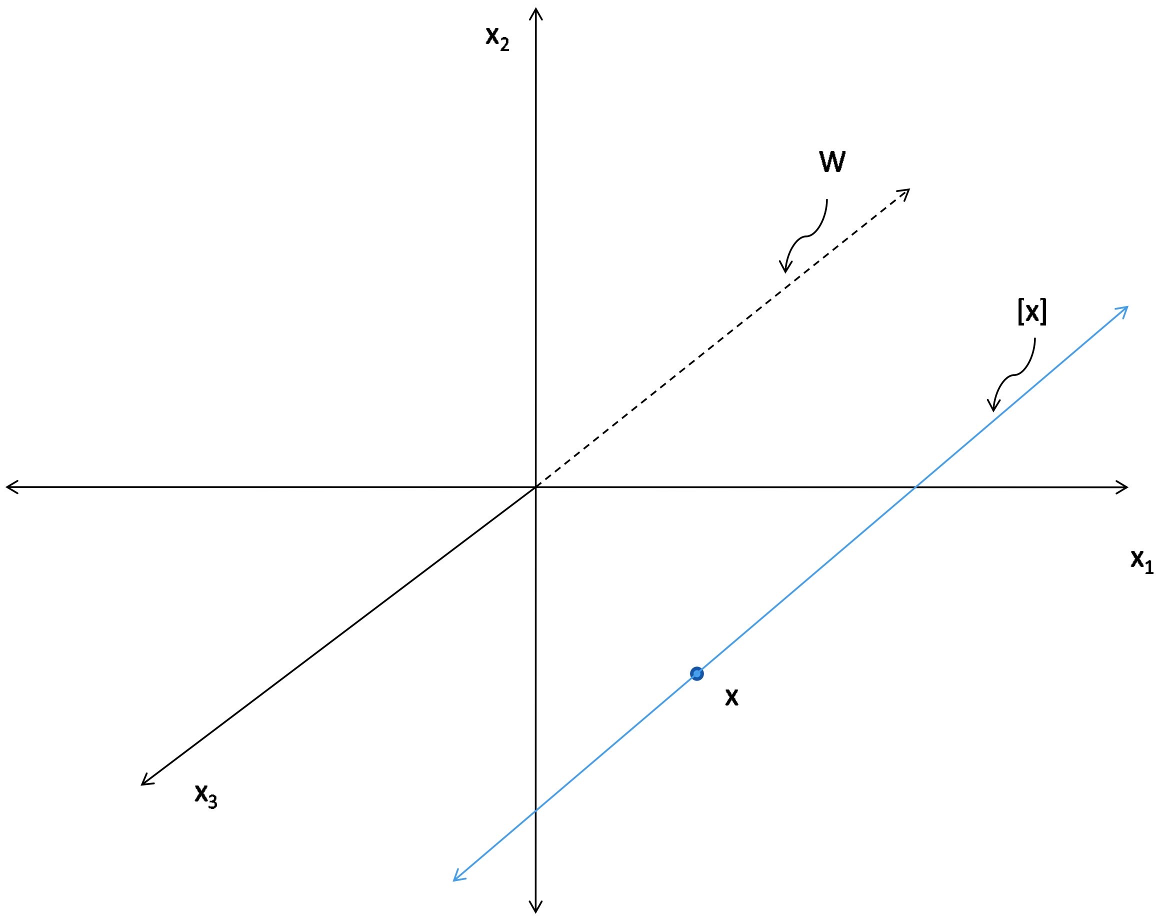

Figure 1: An illustration of X/W where $X = \mathbf{R}^3$ and $W = \{x \in \mathbf{R}^3 : x_1 = x_2 = 0\}$. Here $[x] = \{(x_1, x_2, z) : z \in \mathbf{R}\}$ is the line through x parallel to the axis of the third coordinate.