Coordinate Vectors

References are to Anton–Rorres, 7th Edition

In order to compute in a general vector space V, we usually need to install a coordinate system on V. In effect, we refer everything back to the standard vector space $\mathbf{R}^n$, with its standard basis $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$. It is not enough to say that V looks like $\mathbf{R}^n$; it is necessary to choose a specific linear isomorphism.

We consistently identify vectors $\mathbf{x} \in \mathbf{R}^n$ with $n \times 1$ *column* vectors. We know all about linear transformations between the spaces $\mathbf{R}^n$.

Theorem 1 (=Thm. 4.3.3)

(a) *Every linear transformation $T\colon \mathbf{R}^n \to \mathbf{R}^m$ has the form*

$$T(\mathbf{x}) = A\mathbf{x} \qquad \textit{for all } \mathbf{x} \in \mathbf{R}^n \tag{2}$$

for a unique $m \times n$ matrix A. Explicitly, A is given by

$$A = [T(\mathbf{e}_1) | T(\mathbf{e}_2) | \ldots | T(\mathbf{e}_n)]\,. \tag{3}$$

(b) *Assume $m = n$. Then T is an invertible linear transformation if and only if A is an invertible matrix, and if so, the matrix of T^{-1} is A^{-1}.*

We call A the *matrix of the linear transformation T*.

Coordinate vectors The commonest way to establish an invertible linear transformation (i. e. *linear isomorphism*) between $\mathbf{R}^n$ and a general n-dimensional vector space V is to choose a *basis* $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ of V. Then we define the linear transformation $L_B\colon \mathbf{R}^n \to V$ by

$$L_B(k_1, k_2, \ldots, k_n) = k_1\mathbf{b}_1 + k_2\mathbf{b}_2 + \ldots + k_n\mathbf{b}_n \qquad \text{in } V. \tag{4}$$

So $L_B(\mathbf{e}_i) = \mathbf{b}_i$ for each i.

The purpose of requiring B to be a basis of V is to ensure that L_B is invertible. To get $R(L_B) = V$, we need B to span V; and to get L_B to be 1–1, we need B to be linearly independent. If we already know $\dim(V) = n$, Thm. 5.4.5 shows that it is enough to verify *either* of these conditions; then the other will follow. However, the basis B is only a means to define the linear isomorphism L_B, which is what we are *really* after. It allows us to pass back and forth between V and $\mathbf{R}^n$. Sometimes, it is more convenient to specify L_B or L_B^{-1} directly and forget about the basis.

For the inverse linear transformation $L_B^{-1}\colon V \to \mathbf{R}^n$, we clearly have $L_B^{-1}(\mathbf{b}_i) = \mathbf{e}_i$. More generally, we convert any vector in V to an n-tuple of numbers.

Definition 5 Given a vector $\mathbf{v} \in V$, its *coordinate vector with respect to the basis* B is the vector

$$[\mathbf{v}]_B = L_B^{-1}(\mathbf{v}) \qquad \text{in } \mathbf{R}^n. \tag{6}$$

Example If $V = \mathbf{R}^n$ and we choose the standard basis E, L_E is the identity natural transformation and we have $[\mathbf{x}]_E = \mathbf{x}$. (But if we choose a different basis B, $[\mathbf{x}]_B \neq \mathbf{x}$ in general.

Change of basis Of course the coordinate vector $[\mathbf{v}]_B$ depends on the choice of basis B. The first basis chosen may not be the best. Let $C = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ be another basis of V. Then $L_B^{-1} \circ L_C$ is a linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^n$, and by Theorem 1(a), it has a matrix.

DEFINITION 7 The *transition matrix from C to B* is the matrix P of the linear transformation $L_B^{-1} \circ L_C \colon \mathbf{R}^n \to \mathbf{R}^n$.

To find P explicitly, we use equations (2) and (6) to compute

$$P\mathbf{e}_i = L_B^{-1}(L_C(\mathbf{e}_i)) = L_B^{-1}(\mathbf{c}_i) = [\mathbf{c}_i]_B.$$

Then equation (3) yields immediately

$$P = \Big[[\mathbf{c}_1]_B \Big| [\mathbf{c}_2]_B \Big| \ldots \Big| [\mathbf{c}_n]_B\Big]. \tag{8}$$

Thus the columns of P express the *new* basis vectors in terms of the *old* basis; note the direction.

THEOREM 9 *Let B and C be bases of V, and P be the transition matrix from C to B. Then*

$$[\mathbf{v}]_B = P[\mathbf{v}]_C \qquad \text{in } \mathbf{R}^n. \tag{10}$$

Proof This equation translates the trivial statement

$$L_B^{-1}(\mathbf{v}) = L_B^{-1}(L_C(L_C^{-1}(\mathbf{v}))) = (L_B^{-1} \circ L_C)(L_C^{-1}(\mathbf{v})) \qquad \text{in } \mathbf{R}^n$$

using equations (2) and (6). □

So P is the matrix we need to transform *new* coordinates to *old* ones. We often need the reverse direction, to convert old coordinates to new coordinates.

LEMMA 11 *If P is the transition matrix from the basis C to the basis B, then the transition matrix from B to C is its inverse, P^{-1}.*

Proof Theorem 1(b) shows that P^{-1} is the matrix of the linear transformation

$$(L_B^{-1} \circ L_C)^{-1} = L_C^{-1} \circ (L_B^{-1})^{-1} = L_C^{-1} \circ L_B \colon \mathbf{R}^n \to \mathbf{R}^n.$$

By Definition 7, this is just the matrix we want.

Alternatively, we can simply multiply equation (10) on the left by P^{-1}. □

Inner product spaces If V is an inner product space, we want to take advantage of the extra structure and choose not just any basis.

THEOREM 12 (=Thm. 6.3.2) *If V is an inner product space and B is a basis of V, then L_B preserves the inner product structure, $\langle L_B(\mathbf{x}), L_B(\mathbf{y})\rangle = \mathbf{x}.\mathbf{y}$ (and hence the norm, $\|L_B(\mathbf{x})\| = \|\mathbf{x}\|$), if and only if the basis B is orthonormal.*

Similarly, the relevant transition matrices take a special form.

THEOREM 13 *Suppose B and C are orthonormal bases of V. Then the transition matrix P from C to B is an orthogonal matrix.*

Proof By Theorem 12, L_B and L_C preserve the inner-product structure, and therefore so does $L_B^{-1} \circ L_C$. By Theorem 6.5.3 (or 6.5.1), its matrix P is orthogonal. □