One-to-one and Onto Linear Transformations

One-to-one (injective) and onto (surjective) linear transformations are defined in a corresponding manner to functions: a linear transformation is one-to-one if it maps distinct vectors to distinct vectors, and onto if it covers the whole range. Formally, a mapping $T : V \to W$ is one-to-one if, for all $u, v \in V$,

$$u \neq v \implies T(u) \neq T(v),$$

or equivalently,

$$T(u) = T(v) \implies u = v.$$

T is onto if, for all $w \in W$, $\exists v \in V$ such that $T(v) = W$.
A linear mapping is called an isomorphism if it is one-to-one and onto, *i.e.* if it is a bijection. Two vector spaces that have an isomorphism between them are called isomorphic; this is denoted by $V \cong W$.

Coordinates in a Vector Space

Given a vector space V with basis $B = \{x_1, x_2, \cdots, x_n\}$, any vector $v \in V$ can be expressed as $c_1v_1 + c_2v_2 + \cdots + c_nv_n$ for some collection of c_i. These c_i are called the coordinates of v, and the column vector consisting of the c_i is called the coordinate vector of v with respect to B, denoted by

$$[v]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

Note that the same vector has different coordinate vectors with respect to different bases.

Properties of Coordinate Vectors

1. Let $B = \{v_1, v_2, \cdots, v_n\}$ be a basis of a vector space V. If $u, v \in V$ and c is a scalar, then

(i) $[v + u]_B = [v]_B + [u]_B$, and

(ii) $[cu]_B = c[u]_B$.

Proof. Let $v = x_1v_1 + x_2v_2 + \cdots + x_nv_n$ and $u = y_1v_1 + y_2v_2 + \cdots + y_nv_n$. Then $v + u = (x_1 + y_1)v_1 + (x_2 + y_2)v_2 + \cdots + (x_n + y_n)v_n$.
Then, $[v + u]_B = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n) = (x_1, x_2, \cdots, x_n) + (y_1, y_2, \cdots, y_n) = [v]_B + [u]_B$. This proves property (i).
Further, $cu = cy_1v_1 + cy_2v_2 + \cdots + cy_nv_n$; therefore $[cu]_B = (cy_1, cy_2, \cdots, cy_n) = c(y_1, y_2, \cdots, y_n) = c[u]_B$. This proves property (ii).

2. Let B and V be as above. If $\{u_1, u_2, \cdots, u_k\} \subseteq V$ is a linearly independent set of vectors, then $\{[u_1]_B, [u_2]_B, \cdots, [u_n]_B\}$ is also linearly independent.
Proof. To prove this, suppose

$$u_1 = x_{11}v_1 + x_{12}v_2 + \cdots + x_{1n}v_n,$$

$$u_2 = x_{21}v_1 + x_{22}v_2 + \cdots + x_{2n}v_n,$$

$$\vdots$$

$$u_n = x_{n1}v_1 + x_{n2}v_2 + \cdots + x_{nn}v_n.$$

This means that

$$[u_1]_B = (x_{11}, x_{12}, \cdots, x_{1n}),$$

$$[u_2]_B = (x_{21}, x_{22}, \cdots, x_{2n}),$$

$$\vdots$$

$$[u_n]_B = (x_{n1}, x_{n2}, \cdots, x_{nn}).$$

If the set of coordinate vectors is linearly dependent, there exists scalars c_i such that $c_1[u_1]_B + c_2[u_2]_B + \cdots + c_n[u_n]_B = 0$. This means that

$$c_1 \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{bmatrix} + c_2 \begin{bmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2n} \end{bmatrix} + \cdots + c_n \begin{bmatrix} x_{n1} \\ x_{n2} \\ \vdots \\ x_{nn} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Now, consider the sum $c_1u_1 + c_2u_2 + \cdots + c_nu_n$. Taking the coefficients of each basis vector common, we get this to be equal to $(c_1x_{11} + c_2v_{21} + \cdots + c_nv_{n1})v_1 + (c_1x_{12} + c_2x_{22} + \cdots + c_nv_{n2})v_2 + \cdots + (c_1x_{1n} + c_2x_{2n} + \cdots + c_nx_{nn})v_n$. From the equation in the preceding paragraph, we find the coefficient of each of the v_i to be 0, making the sum 0. Therefore, the u_i are linearly dependent, which is a contradiction. Therefore the $[u_i]_B$ must also be linearly independent.

Note: This proof was not given in class. (No proof was.) If anyone knows a shorter or more elegant proof, please feel free to update it.

Matrix Representations of Linear Transformations

If V and W are finite dimensional vector spaces with bases B and C respectively, where $B = \{v_1, v_2, \cdots, v_n\}$. If $T : V \to W$ is a linear mapping, its matrix representation is the $m \times n$ matrix A defined by $A = \{[T(v_1)]_C, [T(v_2)]_C, \cdots, T[v_n]_C\}$. The significance of this matrix is that for all $v \in V$,

$$[T(v)]_C = A[v]_B.$$

Proof. Suppose v can be expressed as $v = x_1v_1 + x_2v_2 + \cdots + x_nv_n$. Then $T(v) = x_1T(v_1) + x_2T(v_2) + \cdots + x_nT(v_n)$. Further, the right hand side of the given equation becomes

$$\begin{bmatrix}[T(v_1)]_C & [T(v_2)]_C & \cdots & [T(v_n)]_C\end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}.$$

Consider the first column of A. It is equal to $T[v_1]_C$. By the definition of matrix multiplication, its first component is multiplied by x_1 and added to row 1 of the LHS; its second component is multiplied by x_1 and added to row 2 of the LHS, and so on. Therefore the first component of the LHS column matrix is the first component of the vector $x_1[T(v_1)]_C + x_1[T(v_2)]_C + \cdots + x_1[T(v_n)]_C$, which we have seen above is nothing but $[T(v)]_C$. Similarly, all other rows of the LHS are the other components of $[T(v)]_C$, QED.

Inner Product

An inner product over a vector space $V(\mathbb{R})$ is an operation that assigns to every pair of vectors $u, v \in V$ a real number (u, v) such that:
(i) $(u, v) = (v, u)$
(ii) $(u, v + w) = (u, v) + (u, w)$
(iii) $(cu, v) = c(u, v)$
(iv) $(u, v) \geq 0$ and $(u, u) = 0 \iff u = 0$.
Note: Strictly, this is the definition of the inner product restricted to the special case of vector spaces over the real numbers. A general definition exists for arbitrary vector spaces; this will be taught in the next class.

Other properties of the inner product (derivable from the above) are:
(v) $(u + v, w) = (u, w) + (v, w)$
(vi) $(u, cv) = c(u, v)$
(vii) $(u, 0) = (0, u) = 0$.
Proof. The proof is trivial, so we will not go into details. (v) and (vi) can be proved by switching u and v in (i) and switching u and $v + w$ in (ii). (vii) can be proved by substituting $v = w = 0$ in (ii).

Other Functions on Vectors

Some functions on and properties of vectors defined in terms of the inner product are:
(a) The length or norm of a vector v, denoted by $\|v\|$, defined as $\|v\| = \sqrt{(v, v)}$.
(b) The distance between two vectors, denoted as $d(u, v)$, defined as $d(u, v) = \|u - v\|$.
(c) Orthogonality; two vectors u and v are said to be orthogonal if $(u, v) = 0$.