Thus D is represented by the same matrix in the ordered bases $\mathcal{B}$ and $\mathcal{B}'$. Of course, one can see this somewhat more directly since

$$\begin{aligned} Dg_1 &= 0 \\ Dg_2 &= g_1 \\ Dg_3 &= 2g_2 \\ Dg_4 &= 3g_3. \end{aligned}$$

This example illustrates a good point. If one knows the matrix of a linear operator in some ordered basis $\mathcal{B}$ and wishes to find the matrix in another ordered basis $\mathcal{B}'$, it is often most convenient to perform the coordinate change using the invertible matrix P; however, it may be a much simpler task to find the representing matrix by a direct appeal to its definition.

Definition. *Let* A *and* B *be* n $\times$ n *(square) matrices over the field* F. *We say that* B **is similar to** A **over** F *if there is an invertible* n $\times$ n *matrix* P *over* F *such that* B = P^{-1}AP.

According to Theorem 14, we have the following: If V is an n-dimensional vector space over F and $\mathcal{B}$ and $\mathcal{B}'$ are two ordered bases for V, then for each linear operator T on V the matrix $B = [T]_{\mathcal{B}'}$ is similar to the matrix $A = [T]_{\mathcal{B}}$. The argument also goes in the other direction. Suppose A and B are $n \times n$ matrices and that B is similar to A. Let V be any n-dimensional space over F and let $\mathcal{B}$ be an ordered basis for V. Let T be the linear operator on V which is represented in the basis $\mathcal{B}$ by A. If $B = P^{-1}AP$, let $\mathcal{B}'$ be the ordered basis for V obtained from $\mathcal{B}$ by P, i.e.,

$$\alpha_j' = \sum_{i=1}^{n} P_{ij}\alpha_i.$$

Then the matrix of T in the ordered basis $\mathcal{B}'$ will be B.

Thus the statement that B is similar to A means that on each n-dimensional space over F the matrices A and B represent the same linear transformation in two (possibly) different ordered bases.

Note that each $n \times n$ matrix A is similar to itself, using $P = I$; if B is similar to A, then A is similar to B, for $B = P^{-1}AP$ implies that $A = (P^{-1})^{-1}BP^{-1}$; if B is similar to A and C is similar to B, then C is similar to A, for $B = P^{-1}AP$ and $C = Q^{-1}BQ$ imply that $C = (PQ)^{-1}A(PQ)$. Thus, similarity is an equivalence relation on the set of $n \times n$ matrices over the field F. Also note that the only matrix similar to the identity matrix I is I itself, and that the only matrix similar to the zero matrix is the zero matrix itself.

Exercises

1. Let T be the linear operator on C^2 defined by $T(x_1, x_2) = (x_1, 0)$. Let $\mathcal{B}$ be the standard ordered basis for C^2 and let $\mathcal{B}' = \{\alpha_1, \alpha_2\}$ be the ordered basis defined by $\alpha_1 = (1, i)$, $\alpha_2 = (-i, 2)$.

(a) What is the matrix of T relative to the pair $\mathcal{B}$, $\mathcal{B}'$?

(b) What is the matrix of T relative to the pair $\mathcal{B}'$, $\mathcal{B}$?

(c) What is the matrix of T in the ordered basis $\mathcal{B}'$?

(d) What is the matrix of T in the ordered basis $\{\alpha_2, \alpha_1\}$?

2. Let T be the linear transformation from R^3 into R^2 defined by

$$T(x_1, x_2, x_3) = (x_1 + x_2, 2x_3 - x_1).$$

(a) If $\mathcal{B}$ is the standard ordered basis for R^3 and $\mathcal{B}'$ is the standard ordered basis for R^2, what is the matrix of T relative to the pair $\mathcal{B}$, $\mathcal{B}'$?

(b) If $\mathcal{B} = \{\alpha_1, \alpha_2, \alpha_3\}$ and $\mathcal{B}' = \{\beta_1, \beta_2\}$, where

$$\alpha_1 = (1, 0, -1), \quad \alpha_2 = (1, 1, 1), \quad \alpha_3 = (1, 0, 0), \quad \beta_1 = (0, 1), \quad \beta_2 = (1, 0)$$

what is the matrix of T relative to the pair $\mathcal{B}$, $\mathcal{B}'$?

3. Let T be a linear operator on F^n, let A be the matrix of T in the standard ordered basis for F^n, and let W be the subspace of F^n spanned by the column vectors of A. What does W have to do with T?

4. Let V be a two-dimensional vector space over the field F, and let $\mathcal{B}$ be an ordered basis for V. If T is a linear operator on V and

$$[T]_{\mathcal{B}} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

prove that $T^2 - (a + d)T + (ad - bc)I = 0$.

5. Let T be the linear operator on R^3, the matrix of which in the standard ordered basis is

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 4 \end{bmatrix}.$$

Find a basis for the range of T and a basis for the null space of T.

6. Let T be the linear operator on R^2 defined by

$$T(x_1, x_2) = (-x_2, x_1).$$

(a) What is the matrix of T in the standard ordered basis for R^2?

(b) What is the matrix of T in the ordered basis $\mathcal{B} = \{\alpha_1, \alpha_2\}$, where $\alpha_1 = (1, 2)$ and $\alpha_2 = (1, -1)$?

(c) Prove that for every real number c the operator $(T - cI)$ is invertible.

(d) Prove that if $\mathcal{B}$ is any ordered basis for R^2 and $[T]_{\mathcal{B}} = A$, then $A_{12}A_{21} \neq 0$.

7. Let T be the linear operator on R^3 defined by

$$T(x_1, x_2, x_3) = (3x_1 + x_3, -2x_1 + x_2, -x_1 + 2x_2 + 4x_3).$$

(a) What is the matrix of T in the standard ordered basis for R^3?

(b) What is the matrix of T in the ordered basis

$$\{\alpha_1, \alpha_2, \alpha_3\}$$

where $\alpha_1 = (1, 0, 1)$, $\alpha_2 = (-1, 2, 1)$, and $\alpha_3 = (2, 1, 1)$?

(c) Prove that T is invertible and give a rule for T^{-1} like the one which defines T.

8. Let θ be a real number. Prove that the following two matrices are similar over the field of complex numbers:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \qquad \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}$$

(*Hint:* Let T be the linear operator on C^2 which is represented by the first matrix in the standard ordered basis. Then find vectors α_1 and α_2 such that $T\alpha_1 = e^{i\theta}\alpha_1$, $T\alpha_2 = e^{-i\theta}\alpha_2$, and $\{\alpha_1, \alpha_2\}$ is a basis.)

9. Let V be a finite-dimensional vector space over the field F and let S and T be linear operators on V. We ask: When do there exist ordered bases $\mathcal{B}$ and $\mathcal{B}'$ for V such that $[S]_{\mathcal{B}} = [T]_{\mathcal{B}'}$? Prove that such bases exist if and only if there is an invertible linear operator U on V such that $T = USU^{-1}$. (*Outline of proof:* If $[S]_{\mathcal{B}} = [T]_{\mathcal{B}'}$, let U be the operator which carries $\mathcal{B}$ onto $\mathcal{B}'$ and show that $S = UTU^{-1}$. Conversely, if $T = USU^{-1}$ for some invertible U, let $\mathcal{B}$ be any ordered basis for V and let $\mathcal{B}'$ be its image under U. Then show that $[S]_{\mathcal{B}} = [T]_{\mathcal{B}'}$.)

10. We have seen that the linear operator T on R^2 defined by $T(x_1, x_2) = (x_1, 0)$ is represented in the standard ordered basis by the matrix

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

This operator satisfies $T^2 = T$. Prove that if S is a linear operator on R^2 such that $S^2 = S$, then $S = 0$, or $S = I$, or there is an ordered basis $\mathcal{B}$ for R^2 such that $[S]_{\mathcal{B}} = A$ (above).

11. Let W be the space of all $n \times 1$ column matrices over a field F. If A is an $n \times n$ matrix over F, then A defines a linear operator L_A on W through left multiplication: $L_A(X) = AX$. Prove that every linear operator on W is left multiplication by some $n \times n$ matrix, i.e., is L_A for some A.

Now suppose V is an n-dimensional vector space over the field F, and let $\mathcal{B}$ be an ordered basis for V. For each α in V, define $U\alpha = [\alpha]_{\mathcal{B}}$. Prove that U is an isomorphism of V onto W. If T is a linear operator on V, then UTU^{-1} is a linear operator on W. Accordingly, UTU^{-1} is left multiplication by some $n \times n$ matrix A. What is A?

12. Let V be an n-dimensional vector space over the field F, and let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ be an ordered basis for V.

(a) According to Theorem 1, there is a unique linear operator T on V such that

$$T\alpha_j = \alpha_{j+1}, \qquad j = 1, \ldots, n-1, \qquad T\alpha_n = 0.$$

What is the matrix A of T in the ordered basis $\mathcal{B}$?

(b) Prove that $T^n = 0$ but $T^{n-1} \neq 0$.

(c) Let S be any linear operator on V such that $S^n = 0$ but $S^{n-1} \neq 0$. Prove that there is an ordered basis $\mathcal{B}'$ for V such that the matrix of S in the ordered basis $\mathcal{B}'$ is the matrix A of part (a).

(d) Prove that if M and N are $n \times n$ matrices over F such that $M^n = N^n = 0$ but $M^{n-1} \neq 0 \neq N^{n-1}$, then M and N are similar.

13. Let V and W be finite-dimensional vector spaces over the field F and let T be a linear transformation from V into W. If

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\} \quad \text{and} \quad \mathcal{B}' = \{\beta_1, \ldots, \beta_m\}$$

are ordered bases for V and W, respectively, define the linear transformations $E^{p,q}$ as in the proof of Theorem 5: $E^{p,q}(\alpha_i) = \delta_{iq}\beta_p$. Then the $E^{p,q}$, $1 \leq p \leq m$, $1 \leq q \leq n$, form a basis for $L(V, W)$, and so

$$T = \sum_{p=1}^{m} \sum_{q=1}^{n} A_{pq}E^{p,q}$$

for certain scalars A_{pq} (the coordinates of T in this basis for $L(V, W)$). Show that the matrix A with entries $A(p, q) = A_{pq}$ is precisely the matrix of T relative to the pair $\mathcal{B}$, $\mathcal{B}'$.

3.5. Linear Functionals

If V is a vector space over the field F, a linear transformation f from V into the scalar field F is also called a **linear functional** on V. If we start from scratch, this means that f is a function from V into F such that

$$f(c\alpha + \beta) = cf(\alpha) + f(\beta)$$

for all vectors α and β in V and all scalars c in F. The concept of linear functional is important in the study of finite-dimensional spaces because it helps to organize and clarify the discussion of subspaces, linear equations, and coordinates.

EXAMPLE 18. Let F be a field and let $a_1, \ldots, a_n$ be scalars in F. Define a function f on F^n by

$$f(x_1, \ldots, x_n) = a_1x_1 + \cdots + a_nx_n.$$

Then f is a linear functional on F^n. It is the linear functional which is represented by the matrix $[a_1 \cdots a_n]$ relative to the standard ordered basis for F^n and the basis $\{1\}$ for F:

$$a_j = f(\epsilon_j), \qquad j = 1, \ldots, n.$$

Every linear functional on F^n is of this form, for some scalars $a_1, \ldots, a_n$. That is immediate from the definition of linear functional because we define $a_j = f(\epsilon_j)$ and use the linearity

$$\begin{aligned} f(x_1, \ldots, x_n) &= f\left(\sum_j x_j\epsilon_j\right) \\ &= \sum_j x_jf(\epsilon_j) \\ &= \sum_j a_jx_j. \end{aligned}$$

EXAMPLE 19. Here is an important example of a linear functional. Let n be a positive integer and F a field. If A is an $n \times n$ matrix with entries in F, the **trace** of A is the scalar

$$\operatorname{tr} A = A_{11} + A_{22} + \cdots + A_{nn}.$$

The trace function is a linear functional on the matrix space $F^{n\times n}$ because

$$\begin{aligned} \operatorname{tr}(cA + B) &= \sum_{i=1}^{n} (cA_{ii} + B_{ii}) \\ &= c \sum_{i=1}^{n} A_{ii} + \sum_{i=1}^{n} B_{ii} \\ &= c \operatorname{tr} A + \operatorname{tr} B. \end{aligned}$$

EXAMPLE 20. Let V be the space of all polynomial functions from the field F into itself. Let t be an element of F. If we define

$$L_t(p) = p(t)$$

then L_t is a linear functional on V. One usually describes this by saying that, for each t, 'evaluation at t' is a linear functional on the space of polynomial functions. Perhaps we should remark that the fact that the functions are polynomials plays no role in this example. Evaluation at t is a linear functional on the space of all functions from F into F.

EXAMPLE 21. This may be the most important linear functional in mathematics. Let $[a, b]$ be a closed interval on the real line and let $C([a, b])$ be the space of continuous real-valued functions on $[a, b]$. Then

$$L(g) = \int_a^b g(t)\, dt$$

defines a linear functional L on $C([a, b])$.

If V is a vector space, the collection of all linear functionals on V forms a vector space in a natural way. It is the space $L(V, F)$. We denote this space by V^* and call it the **dual space** of V:

$$V^* = L(V, F).$$

If V is finite-dimensional, we can obtain a rather explicit description of the dual space V^*. From Theorem 5 we know something about the space V^*, namely that

$$\dim V^* = \dim V.$$

Let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ be a basis for V. According to Theorem 1, there is (for each i) a unique linear functional f_i on V such that

$$f_i(\alpha_j) = \delta_{ij}. \tag{3-11}$$

In this way we obtain from $\mathcal{B}$ a set of n distinct linear functionals $f_1, \ldots, f_n$ on V. These functionals are also linearly independent. For, suppose

$$f = \sum_{i=1}^{n} c_i f_i. \tag{3-12}$$

Then

$$\begin{aligned} f(\alpha_j) &= \sum_{i=1}^{n} c_i f_i(\alpha_j) \\ &= \sum_{i=1}^{n} c_i \delta_{ij} \\ &= c_j. \end{aligned}$$

In particular, if f is the zero functional, $f(\alpha_j) = 0$ for each j and hence the scalars c_j are all 0. Now $f_1, \ldots, f_n$ are n linearly independent functionals, and since we know that V^* has dimension n, it must be that $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ is a basis for V^*. This basis is called the **dual basis** of $\mathcal{B}$.

Theorem 15. *Let* V *be a finite-dimensional vector space over the field* F, *and let* $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ *be a basis for* V. *Then there is a unique dual basis* $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ *for* V^* *such that* $f_i(\alpha_j) = \delta_{ij}$. *For each linear functional* f *on* V *we have*

$$f = \sum_{i=1}^{n} f(\alpha_i) f_i \tag{3-13}$$

and for each vector α *in* V *we have*

$$\alpha = \sum_{i=1}^{n} f_i(\alpha)\alpha_i. \tag{3-14}$$

Proof. We have shown above that there is a unique basis which is 'dual' to $\mathcal{B}$. If f is a linear functional on V, then f is some linear combination (3-12) of the f_i, and as we observed after (3-12) the scalars c_j must be given by $c_j = f(\alpha_j)$. Similarly, if

$$\alpha = \sum_{i=1}^{n} x_i \alpha_i$$

is a vector in V, then

$$\begin{aligned} f_j(\alpha) &= \sum_{i=1}^{n} x_i f_j(\alpha_i) \\ &= \sum_{i=1}^{n} x_i \delta_{ij} \\ &= x_j \end{aligned}$$

so that the unique expression for α as a linear combination of the α_i is

$$\alpha = \sum_{i=1}^{n} f_i(\alpha)\alpha_i. \quad \blacksquare$$

Equation (3-14) provides us with a nice way of describing what the dual basis is. It says, if $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ is an ordered basis for V and