Change of Bases

Isaiah Lankham, Bruno Nachtergaele, Anne Schilling
(March 8, 2007)

We have seen in previous lectures that linear operators on an n dimensional vector space are in one-to-one correspondence with $n \times n$ matrices. This correspondence depends on the choice of basis for the vector space, however. In this lecture we address the question how the matrix for a linear operator changes if we change from one orthonormal basis to another.

1 Coordinate vectors

Let V be a finite-dimensional inner product space with inner product $\langle \cdot, \cdot \rangle$ and dimension $\dim V = n$. Then V has an orthonormal basis $e = (e_1, \dots, e_n)$. As we have seen in a previous lecture we can write every $v \in V$ as

$$v = \sum_{i=1}^{n} \langle v, e_i \rangle e_i.$$

This induces a map

$$\begin{aligned} [\cdot]_e : V &\to \mathbb{F}^n \\ v &\mapsto \begin{bmatrix} \langle v, e_1 \rangle \\ \vdots \\ \langle v, e_n \rangle \end{bmatrix} \end{aligned}$$

which maps the vector $v \in V$ to the $n \times 1$ column vector of its coordinates with respect to the basis e. The column vector $[v]_e$ is also called the **coordinate vector** of v with respect to the basis e.

Furthermore, the map $[\cdot]_e$ is an isomorphism (meaning that it is an injective and surjective linear map). On $\mathbb{F}^n$ we have the usual inner product defined as

$$\langle x, y \rangle_{\mathbb{F}^n} = \sum_{k=1}^{n} x_k \overline{y}_k.$$

The map $[\cdot]_e$ preserves the inner product, that is,

$$\langle v, w\rangle_V = \langle [v]_e, [w]_e\rangle_{\mathbb{F}^n} \qquad \text{for all } v, w \in V,$$

since

$$\begin{aligned}\langle v, w\rangle_V = \sum_{i,j=1}^n \langle \langle v, e_i\rangle e_i, \langle w, e_j\rangle e_j\rangle &= \sum_{i,j=1}^n \langle v, e_i\rangle \overline{\langle w, e_j\rangle}\langle e_i, e_j\rangle \\ &= \sum_{i,j=1}^n \langle v, e_i\rangle \overline{\langle w, e_j\rangle}\delta_{ij} = \sum_{i=1}^n \langle v, e_i\rangle \overline{\langle w, e_i\rangle} = \langle [v]_e, [w]_e\rangle_{\mathbb{F}^n}.\end{aligned}$$

It is important to remember that the map $[\cdot]_e$ depends on the choice of basis $e = (e_1, \ldots, e_n)$.

2 Change of basis transformation

Recall that we can associate a matrix $A \in \mathbb{F}^{n\times n}$ to every operator $T \in \mathcal{L}(V, V)$. More precisely, the j-th column of the matrix $A = M(T)$ with respect to a basis $e = (e_1, \ldots, e_n)$ is obtained by expanding Te_j in terms of the basis e. If the basis e is orthonormal, the coefficient of e_i is just the inner product of the vector with e_i. Hence

$$M(T) = (\langle Te_j, e_i\rangle)_{1\le i,j\le n}$$

where i is the row index and j is the column index of the matrix.

Conversely, if $A \in \mathbb{F}^{n\times n}$ is a matrix, then we can associate to it a linear operator $T \in \mathcal{L}(V, V)$ by

$$\begin{aligned}Tv = \sum_{j=1}^n \langle v, e_j\rangle Te_j &= \sum_{j=1}^n \sum_{i=1}^n \langle Te_j, e_i\rangle \langle v, e_j\rangle e_i \\ &= \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij}\langle v, e_j\rangle \right) e_i = \sum_{i=1} (A[v]_e)_i e_i\end{aligned}$$

where $(A[v]_e)_i$ denotes the i-th component of the column vector $A[v]_e$. With this construction we have $M(T) = A$. The coefficients of Tv in the basis $(e_1, \ldots, e_n)$ are recorded by the column vector obtained by multiplying the $n \times n$ matrix A with the $n \times 1$ column vector $[v]_e$ with components $v_j = \langle v, e_j\rangle$.

Example 1. Let

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

With respect to the canonical basis we can define $T \in \mathcal{L}(V, V)$.

$$T \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} z_1 - iz_2 \\ iz_1 + z_2 \end{bmatrix}.$$

Suppose that we want to use another orthonormal basis $f = (f_1, \ldots, f_n)$ of V. Then as before we have $v = \sum_{i=1}^n \langle v, f_i \rangle f_i$. Comparing this with $v = \sum_{j=1}^n \langle v, e_j \rangle e_j$ we find

$$v = \sum_{i,j=1}^n \langle \langle v, e_j \rangle e_j, f_i \rangle f_i = \sum_{i=1}^n \left(\sum_{j=1}^n \langle e_j, f_i \rangle \langle v, e_j \rangle \right) f_i.$$

Hence

$$[v]_f = S[v]_e,$$

where

$$S = (s_{ij})_{i,j=1}^n \qquad \text{with } s_{ij} = \langle e_j, f_i \rangle.$$

The j-th column of S is given by the coefficients of the expansion of e_j in terms of the basis $f = (f_1, \ldots, f_n)$. The matrix S describes a linear map in $\mathcal{L}(\mathbb{F}^n)$ which is called the **change of basis transformation**.

We may also interchange the role of the bases e and f. In this case we obtain the matrix $R = (r_{ij})_{i,j=1}^n$ where

$$r_{ij} = \langle f_j, e_i \rangle.$$

Then by the uniqueness of the expansion in a basis we obtain

$$[v]_e = R[v]_f$$

so that

$$RS[v]_e = [v]_e \qquad \text{for all } v \in V.$$

Since this equation is true for all $[v]_e \in \mathbb{F}^n$ it follows that $RS = I$ is the identity or $R = S^{-1}$. In particular, S and R are invertible. We can also check this explicitly by using the properties of orthonormal bases. Namely

$$\begin{aligned} (RS)_{ij} &= \sum_{k=1}^n r_{ik} s_{kj} = \sum_{k=1}^n \langle f_k, e_i \rangle \langle e_j, f_k \rangle \\ &= \sum_{k=1}^n \langle e_j, f_k \rangle \overline{\langle e_i, f_k \rangle} = \langle [e_j]_f, [e_i]_f \rangle_{\mathbb{F}^n} = \delta_{ij}. \end{aligned}$$

The matrices S (and also R) have the interesting property that their columns (and rows) are orthonormal, since they are the coordinates of orthonormal vectors in another orthonormal basis.

Example 2. Let $V = \mathbb{C}^2$ and choose the orthonormal bases $e = (e_1, e_2)$ and $f = (f_1, f_2)$ with

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad e_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
$$f_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad f_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Then

$$S = \begin{bmatrix} \langle e_1, f_1 \rangle & \langle e_2, f_1 \rangle \\ \langle e_1, f_2 \rangle & \langle e_2, f_2 \rangle \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

and

$$R = \begin{bmatrix} \langle f_1, e_1 \rangle & \langle f_2, e_1 \rangle \\ \langle f_1, e_2 \rangle & \langle f_2, e_2 \rangle \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

One can check explicitly that indeed

$$RS = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

So far we have discussed how the coordinate vectors of a given vector $v \in V$ change under the change of basis from e to f. The next question we can ask is how the matrix $M(T)$ of an operator $T \in \mathcal{L}(V)$ changes if we change the basis. Let A be the matrix of T with respect to the basis $e = (e_1, \ldots, e_n)$ and let B be the matrix for T with respect to the basis $f = (f_1, \ldots, f_n)$. Can we determine B from A? Note that

$$[Tv]_e = A[v]_e$$

so that

$$[Tv]_f = S[Tv]_e = SA[v]_e = SAR[v]_f = SAS^{-1}[v]_f.$$

This implies that

$$B = SAS^{-1}.$$

Example 3. Continuing Example 2, let

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

be the matrix of a linear operator with respect to the basis e. Then the matrix B with respect to the basis f is given by

$$B = SAS^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$