

Dodson & Poston: Problem I.2.6b

Change of basis matrix

Peter Mao

April 11, 2020

Abstract

If β is a basis for vector space X and $\mathbf{A} : X \rightarrow X$ is an isomorphism, the change of basis matrix $[\mathbf{I}]_{\beta}^{\mathbf{A}\beta}$ is exactly the matrix $([\mathbf{A}]_{\beta}^{\beta})^{\leftarrow}$.

1 Ground rules

For the standard basis of \mathbb{R}^n , I will use the symbol $\mathcal{E} = [e_1, e_2, \dots, e_n]$, where e_j is a vector with 1 in the j^{th} position and 0's everywhere else.

We consider the basis β as an ordered set of vectors in \mathbb{R}^n , which we represent as the columns of the matrix $[\beta] = [b_1, b_2, \dots, b_n]$ (with the choice of basis not yet explicitly stated). Since \mathbf{A} is an isomorphism, the matrix-matrix product $\mathbf{A}\beta$ is also a basis for X with columns determined by the matrix-vector product $\mathbf{A}b_i$, established in problem I.2.6a.

The matrix $[\mathbf{I}]_{\beta}^{\mathbf{A}\beta}$ is the change of basis matrix from the β basis to the $\mathbf{A}\beta$ basis, while the matrix $[\mathbf{A}]_{\beta}^{\beta}$ is the operator \mathbf{A} written in the β basis.

We want to show that $[\mathbf{A}]_{\beta}^{\beta}$ is the change of basis matrix from the new basis $\mathbf{A}\beta$ back to the original β basis, ie,

$$[\mathbf{I}]_{\mathbf{A}\beta}^{\beta} = [\mathbf{A}]_{\beta}^{\beta}$$

or equivalently,

$$[\mathbf{I}]_{\beta}^{\mathbf{A}\beta} = ([\mathbf{A}]_{\beta}^{\beta})^{\leftarrow}.$$

2 A matrix is...

Section I.2.07 of Dodson and Poston essentially demonstrates that a matrix representing a linear transformation consists of the vectors *to which* the standard basis transforms, as it's column vectors. We can see this is so because the e_j , the j^{th} standard basis vector “picks out” the j^{th} column of \mathbf{A} :

$$\mathbf{A}e_j = a_j^i.$$

For the problem at hand, consider the matrix of β 's vectors written in the standard basis:

$$[\beta]_{\mathcal{E}}^{\mathcal{E}} = [[b_1]^{\mathcal{E}} [b_2]^{\mathcal{E}} \cdots [b_n]^{\mathcal{E}}].$$

If we multiply this matrix with one of the standard basis vectors, we get the corresponding basis vector:

$$[\beta]_{\mathcal{E}}^{\mathcal{E}}[e_j]^{\mathcal{E}} = [[b_1]^{\mathcal{E}} [b_2]^{\mathcal{E}} \cdots [b_n]^{\mathcal{E}}] [e_j]^{\mathcal{E}} = [b_j]^{\mathcal{E}}.$$

Another way to think about the matrix is that it is the identity transform from the basis that its columns make up to the standard basis. The first important identity to remember here is that b_1 in the β basis *looks like* e_1 in the standard basis:

$$[b_1]^{\beta} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = [e_1]^{\mathcal{E}}.$$

In this context, it doesn't make sense to multiply a matrix in one basis, say β , with one in a different basis, say \mathcal{E} . Numerically, however, we know can multiply the matrix we are calling $[\beta]_{\mathcal{E}}^{\mathcal{E}}$ with $[b_1]^{\beta}$ and, numerically, the result will be $[b_1]^{\mathcal{E}}$. So, following the discussion in Section I.2.08 in Dodson and Poston, we call this the identity matrix that changes our basis from β to \mathcal{E} :

$$\begin{aligned} [\beta]_{\mathcal{E}}^{\mathcal{E}}[e_j]^{\mathcal{E}} &= [b_j]^{\mathcal{E}} && \text{(Standard method for extracting a column from a matrix)} \\ [\beta]_{\mathcal{E}}^{\mathcal{E}}[b_j]^{\beta} &= [b_j]^{\mathcal{E}} && \text{(Numerically equivalent “nonsense equation” with } e_j \mapsto b_j) \\ [\mathbf{I}]_{\beta}^{\mathcal{E}}[b_j]^{\beta} &= [b_j]^{\mathcal{E}} && \text{(Operator replaced to formalize the change of basis)} \end{aligned}$$

The takeaway from this section is that **a matrix of basis vectors ($[b_1 \cdots b_n]$) written in an arbitrary basis (\mathcal{E}) is numerically equivalent to the identity change of basis matrix from the basis of the matrix (β) to the basis in which it is written (\mathcal{E}).** Succinctly,

$$[\beta]_{\mathcal{E}}^{\mathcal{E}} = [\mathbf{I}]_{\beta}^{\mathcal{E}}, \tag{1}$$

where we interpret the right side as an identity change of basis from β to \mathcal{E} . We almost have the result we seek.

3 Solution via the change of basis equation

In order to save you from flipping pages, the important result of the last section (aka, Equation 1) is:

$$[\beta]_{\mathcal{E}}^{\mathcal{E}} = [\mathbf{I}]_{\beta}^{\mathcal{E}}.$$

To change the basis of operator \mathbf{A} , we conjugate it with the identity “change of basis” matrix of section I.2.08, middle of page 29:

$$[\mathbf{A}]_{\beta}^{\beta} = [\mathbf{I}]_{\mathcal{E}}^{\beta} [\mathbf{A}]_{\mathcal{E}}^{\mathcal{E}} [\mathbf{I}]_{\beta}^{\mathcal{E}}.$$

I prefer to start from the standard basis, because I am weak-minded, and if I see a matrix, it is hard for me to imagine it as anything other than a matrix in the standard basis representation. The terms on the right side of the above equation should all be familiar: $[\mathbf{A}]_{\mathcal{E}}^{\mathcal{E}}$ is the standard basis representation of \mathbf{A} , and $[\mathbf{I}]_{\beta}^{\mathcal{E}}$ is $[\beta]_{\mathcal{E}}^{\mathcal{E}}$ (Equation 1), the standard basis representation of β . We can therefore rewrite the above equation as:

$$\begin{aligned} [\mathbf{A}]_{\beta}^{\beta} &= [\mathbf{I}]_{\mathcal{E}}^{\beta} [\mathbf{A}]_{\mathcal{E}}^{\mathcal{E}} [\mathbf{I}]_{\beta}^{\mathcal{E}} && \text{(change basis of } \mathbf{A} \text{ from } \mathcal{E} \text{ to } \beta) \\ &= [\mathbf{I}]_{\mathcal{E}}^{\beta} [\mathbf{A}]_{\mathcal{E}}^{\mathcal{E}} [\beta]_{\mathcal{E}}^{\mathcal{E}} && \text{(using Equation 1)} \\ &= [\mathbf{I}]_{\mathcal{E}}^{\beta} [\mathbf{A}\beta]_{\mathcal{E}}^{\mathcal{E}} && \text{(last 2 terms are in the same basis)} \\ &= [\mathbf{I}]_{\mathcal{E}}^{\beta} [\mathbf{I}]_{\mathbf{A}\beta}^{\mathcal{E}} && \text{(using Equation 1 with } \beta \mapsto \mathbf{A}\beta) \\ [\mathbf{A}]_{\beta}^{\beta} &= [\mathbf{I}]_{\mathbf{A}\beta}^{\beta} && \text{(the total change of basis is from } \mathbf{A}\beta \text{ to } \beta). \end{aligned}$$

Voilà! This is the result we seek. The importance of this result is that it allows us to explicitly calculate an operator in any basis from standard “knowns.”

This identity is used in the next I.2.6c to establish a “physical” meaning for operators that are similar by conjugation. For comparison with group theory, see section 1.9 of Alekseev.

4 A more concrete example

This section's approach falls far short of a mathematical proof, but it may be useful as a way to pull out the stated identity and, perhaps, elucidate its geometric meaning.

We are asked to show that $[\mathbf{I}]_{\mathbf{A}\beta}^\beta = [\mathbf{A}]_\beta^\beta$. In order to get away from the standard basis, let's use the normalized eigenvectors of \mathbf{A} as the basis β . Then $[\mathbf{A}]_\beta^\beta$ is diagonal with the corresponding eigenvalues on the diagonal, and the normalized eigenvectors in their own basis look exactly like standard basis vectors for \mathbb{R}^n . The $\mathbf{A}\beta$ basis is the β basis scaled by the eigenvalues,¹ so the vector $(1, 0, \dots, 0)$ in the $\mathbf{A}\beta$ basis is $(\lambda_1, 0, \dots, 0)$ in the β basis. The identity follows.

¹since each basis vector in β is an eigenvector of \mathbf{A} .