Suppose you have $\mathbb{R}^n$ with the standard basis

$\underline{e} = \{e_1, \dots, e_n\}$ and another basis $\underline{w} = \{w_1, \dots, w_n\}$

There are two types of "change of basis" matrices

#1: The change of coordinate matrix from $\underline{w}$ to $\underline{e}$

$$P_{\underline{e} \leftarrow \underline{w}} [x]_{\underline{w}} = [x]_{\underline{e}} \quad , \forall \text{ vector } x \in \mathbb{R}^n$$

(e.g when $x = v_i$, we have $P_{\underline{e} \leftarrow \underline{w}} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = [w_i]_e$, so $P_{\underline{e} \leftarrow \underline{w}} = (w_1, \dots, w_n)$)

#2: The matrix which takes the $\underline{w}$ vectors to the $\underline{e}$ vectors

$$P_{\underline{w} \leftarrow \underline{e}} (w_i) = e_i$$

These two matrices are inverses of each other!

The reason for the choice of notation / terminology is that

I believe #1 is more important that #2 in the big picture

Now suppose you have a linear function $f: \mathbb{R}^n \to \mathbb{R}^m$ and you wish to represent it with respect to two types of bases: either ($\underline{e}$ of $\mathbb{R}^n$ and $\underline{e}$ of $\mathbb{R}^m$) or ($\underline{w}$ of $\mathbb{R}^n$ and $\underline{v}$ of $\mathbb{R}^m$)

$$[f(z)]_{\underline{e}} = B[z]_{\underline{e}} \qquad (*) \qquad [f(z)]_{\underline{v}} = A[z]_{\underline{w}}$$

The relation between A and B is $A = P_{\underline{v} \leftarrow \underline{e}} B P_{\underline{e} \leftarrow \underline{w}}$

(Thm 16.1, change of basis for matrices)

This is all in line with #1 above; in terms of #2, we have $P_{\underline{e} \leftarrow \underline{w}} e_i = w_i$

$$\Downarrow$$

$$B P_{\underline{e} \leftarrow \underline{w}} e_i = B w_i = \sum_{j=1}^{m} a_{ji} v_j$$

$$\Downarrow$$

$$A e_i = P_{\underline{v} \leftarrow \underline{e}} B P_{\underline{e} \leftarrow \underline{w}} e_i = \sum_{j=1}^{m} a_{ji} e_j$$

Entries of A in the $(\underline{e}, \underline{e})$ basis match entries of B in $(\underline{w}, \underline{v})$ basis

(as we saw in Lecture 16, intro)

Though this might seem "the other way around" from

(*) above, it is actually equivalent to it, and in fact the same kind of pseudo-paradox as the distinction between #1 and #2