Lecture 7: The change of Basis Formula for the Coordinates of a Vector

This lecture does not come from the text. I am using the notation of a linear algebra text by David Pode, Ch 6.

Let V be an n-dimensional vector space with basis $\mathscr{B} = \{b_1, \ldots, b_n\}$. The basis $\mathscr{B}$ allows us to associate to each vector $v \in V$ an element

$$[v]_{\mathscr{B}} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n.$$

In the computations that follow we will usually write $[v]_{\mathscr{B}}$ as a column vector

$$[v]_{\mathscr{B}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

So we use column vectors for computations.

The rule for going from v to $(x_1, x_2, \ldots, x_n)$ is

$$v = \sum_{i=1}^{n} x_i b_i$$

or

$$v = (b_1, \ldots, b_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

You should think of $(x_1, \ldots, x_n)$ as the coordinates of v relative to $\mathscr{B}$.

Remark: The notation $[v]_{\mathscr{B}}$ looks very strange but it is the key to remembering the first change of basis formula.

In the proof of the First Change of Basis Formula it will be important to think of $v \longrightarrow [v]_{\mathscr{B}}$ as a map from V to $\mathbb{R}^n$. This map is linear.

Lemma (1)

(i) $[v_1 + v_2]_{\mathscr{B}} = [v_1]_{\mathscr{B}} + [v_2]_{\mathscr{B}}$.

(ii) $[cv]_{\mathscr{B}} = c\,[v]_{\mathscr{B}}$.

Proof. The proof is left to you.

Note: (i) says "coordinates add".

The Change of Basis Matrix from $\mathscr{B}$ to $\mathscr{C}$

Let $\mathscr{B} = \{b_1, \ldots, b_n\}$ and $\mathscr{C} = \{c_1, \ldots, c_n\}$ be bases of V. Suppose $v \in V$. So we have the coordinates $[v]_{\mathscr{B}} = (x_1, x_2, \ldots, x_n)$ of v relative to $\mathscr{B}$ and $[v]_{\mathscr{C}} = (y_1, y_2, \ldots, y_n)$ of v relative to $\mathscr{C}$.

So how are $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ related?

To answer this question we need the "change of basis matrix from $\mathscr{B}$ to $\mathscr{C}$".

Now we have one of the most important definitions of the course.

Definition

The change of basis matrix from $\mathscr{B}$ to $\mathscr{C}$ written $P_{\mathscr{C} \longleftarrow \mathscr{B}}$ is the matrix whose columns are the "old basis vectors", that is, the vectors in $\mathscr{B}$ written out in terms of the "new basis" $\mathscr{C}$. So, if $\mathscr{B} = \{b_1, \ldots, b_n\}$ and $\mathscr{C} = \{c_1, \ldots, c_n\}$, we have

$$P_{\mathscr{C} \longleftarrow \mathscr{B}} = \begin{pmatrix} [b_1]_{\mathscr{C}} & [b_2]_{\mathscr{C}} & \ldots & [b_n]_{\mathscr{C}} \\ & & & \\ & & & \end{pmatrix} = \begin{matrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{matrix} \begin{pmatrix} b_1 & b_2 & \ldots & b_n \\ \downarrow & \downarrow & & \downarrow \\ & & & \end{pmatrix}$$

Example: Suppose $\mathscr{B} = \{(1, 1), (1, -1)\}$ and $\mathscr{C} = \mathscr{E} =$ the standard basis $\{e_1, e_2\}$. Then

$$P_{\mathscr{C} \longleftarrow \mathscr{B}} = \begin{matrix} e_1 \\ e_2 \end{matrix} \begin{pmatrix} (1, 1 & (1, -1) \\ \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Think of $\mathscr{C} \longleftarrow \mathscr{B}$ as "$\mathscr{B}$ in terms of $\mathscr{C}$".

Although our main goal here is to learn how to make change of basis computations we need the following result to prove some key formulas. If you understand the definition of the matrix ${}_{\mathscr{C}}[T]_{\mathscr{B}}$ of a linear transformation in terms of the input basis $\mathscr{B}$ and the output basis $\mathscr{C}$ then the following result is obvious.

Proposition (1)

Let $I_V : V \longrightarrow V$ be the identity. Then

$$P_{\mathscr{C} \longleftarrow \mathscr{B}} = {}_{\mathscr{C}}[I_V]_{\mathscr{B}}\,.$$

Proof. Let $\mathscr{B} = \{b_1, \ldots, b_n\}$ and $\mathscr{C} = \{c_1, \ldots, c_n\}$. Let $M = {}_{\mathscr{C}}[I_V]_{\mathscr{B}}$ be the matrix of the identity linear transformation relative to the two bases so

$$M = \begin{matrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{matrix} \overset{\begin{matrix} [I_V(b_1)]_{\mathscr{C}} & [I_V(b_2)]_{\mathscr{C}} & \ldots & [I_V(b_n)]_{\mathscr{C}} \end{matrix}}{\begin{pmatrix} & & & \\ & & & \\ \downarrow & \downarrow & \ldots & \downarrow \\ & & & \end{pmatrix}}$$

But this is just $P_{\mathscr{C} \longleftarrow \mathscr{B}}$. $\square$

The first consequence of Proposition (1) is

Proposition (2)

$P_{\mathscr{B} \longleftarrow \mathscr{C}} = (P_{\mathscr{C} \longleftarrow \mathscr{B}})^{-1}$ so the change of basis matrix from $\mathscr{C}$ to $\mathscr{B}$ is the inverse of the change of basis matrix from $\mathscr{B}$ to $\mathscr{C}$.

Proof. We will prove the matrix equation

$$P_{\mathscr{C} \longleftarrow \mathscr{B}} \bullet P_{\mathscr{B} \longleftarrow \mathscr{C}} = I = \text{ the identity matrix.}$$

We use Proposition 1:

$$\text{LHS} = {}_{\mathscr{C}}[I_V]_{\mathscr{B}} \bullet {}_{\mathscr{B}}[I_V]_{\mathscr{C}} .$$

Here $\bullet$ is matrix multiplication.

Now the formula relating matrix products and composition of linear transformations say

$$ {}_{\mathscr{C}}[S]_{\mathscr{B}} \bullet {}_{\mathscr{B}}[T]_{\mathscr{C}} = {}_{\mathscr{C}}[S \circ T]_{\mathscr{C}} \, , $$

where $\circ$ is compostion of linear transformations.
So we get

$$ \begin{aligned} \text{LHS} &= {}_{\mathscr{C}}[I_V \circ I_V]_{\mathscr{C}} \\ &= {}_{\mathscr{C}}[I_V]_{\mathscr{C}} \\ &= I. \end{aligned} $$

□

The First Change of Basis Formula

Theorem (The first change of basis formula)

Let $\mathscr{B}$ and $\mathscr{C}$ be bases of V. Then for any $v \in V$, we have

$$[v]_{\mathscr{C}} = P_{\mathscr{C} \longleftarrow \mathscr{B}} [v]_{\mathscr{B}} .$$

(Mnemonic–keep the $\mathscr{B}$'s together.

Proof. Let $[v]_{\mathscr{B}} = (x_1, \ldots, x_n)$ so

$$v = x_1 b_1 + \ldots + x_n b_n.$$

Then

$$[v]_{\mathscr{C}} = [x_1 b_1 + \ldots + x_n b_n]_{\mathscr{C}} .$$

By Lemma (1), we get

$$[v]_{\mathscr{C}} = x_1 [b_1]_{\mathscr{C}} + \ldots + x_n [b_n]_{\mathscr{C}} \qquad (*)$$

So we have to prove that the RHS of $(*)$ is equal to $P_{\mathscr{C} \longleftarrow \mathscr{B}} [v]_{\mathscr{B}}$.

We need to recall a fact about matrix multiplication.

$$A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1(\ 1^{st} \text{ column of A}) + x_2(2^{nd} \text{ column of A}) + \dots + x_n(n^{th} \text{ column of A}).$$

Hence

$$P_{\mathscr{C}\longleftarrow\mathscr{B}}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1(\ 1^{st} \text{ column of } P_{\mathscr{C}\longleftarrow\mathscr{B}}) + x_2(\ 2^{nd} \text{ column of } P_{\mathscr{C}\longleftarrow\mathscr{B}}) + \dots + x_n(\ n^{th} \text{ column of } P_{\mathscr{C}\longleftarrow\mathscr{B}}).$$

But the 1^{st} column of $P_{\mathscr{C} \longleftarrow \mathscr{B}}$ is $[b_1]_{\mathscr{C}}$, ..., the n^{th} column of $P_{\mathscr{C} \longleftarrow \mathscr{B}}$ is $[b_n]_{\mathscr{C}}$.
Hence

$$P_{\mathscr{C} \longleftarrow \mathscr{B}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 [b_1]_{\mathscr{C}} + \ldots + x_n [b_n]_{\mathscr{C}} .$$

□

Example. Let $V = \mathbb{R}^2$. Suppose $\mathscr{B} = \mathscr{E} = \{e_1, e_2\}$ = the standard basis and $\mathscr{C} = \{c_1, c_2\}$ with

$$\begin{aligned} c_1 &= \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \\ c_2 &= \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right). \end{aligned}$$

So $\{c_1, c_2\}$ is the standard basis rotated by $45°$.
What are the coordinates (a, b) of e_2 relative to the basis $\mathscr{C}$? So we want $[e_2]_{\mathscr{C}}$.

Geometrically, a should be the length of the projection of e_2 onto the line through c_1 and b should be the projection of e_2 on the line through c_2 so it looks like $a = b$. This is because c_1 and c_2 are perpendicular.
Let's check. By Theorem (1) we have

$$\begin{aligned} [e_2]_{\mathscr{C}} &= P_{\mathscr{C} \longleftarrow \mathscr{B}} [e_2]_{\mathscr{B}} \\ &= P_{\mathscr{C} \longleftarrow \mathscr{B}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= 2^{nd} \text{ column of } P_{\mathscr{C} \longleftarrow \mathscr{B}}. \end{aligned}$$

Unfortunately, the matrix that is easy to get is $P_{\mathscr{B}\longleftarrow\mathscr{C}}$

$$P_{\mathscr{C}\longleftarrow\mathscr{B}} = \begin{array}{c} [c_1]_{\mathscr{B}}\,[c_2]_{\mathscr{B}} \\ \left(\begin{array}{cc} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{array}\right) \end{array}$$

By Proposition (2) we have

$$\begin{aligned} P_{\mathscr{B}\longleftarrow\mathscr{C}} = \left(P_{\mathscr{C}\longleftarrow\mathscr{B}}\right)^{-1} &= \left(\begin{array}{cc} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{array}\right)^{-1} \\ &= \left(\begin{array}{cc} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{array}\right) \end{aligned}$$

Hence $a = b = \dfrac{1}{\sqrt{2}}$, so $a = b$ as we guessed and we should have seen the length of the projection was $\dfrac{1}{\sqrt{2}}$.