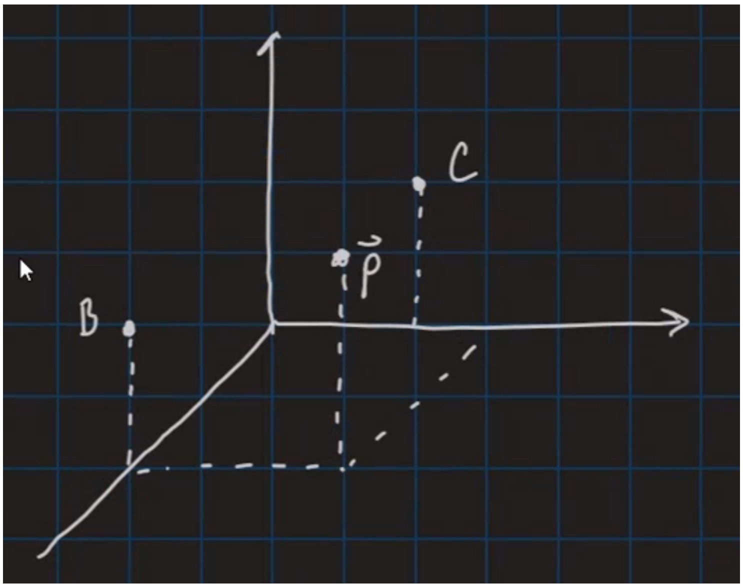
1 Change of Basis

Change of basis is useful for changing coordinate systems using direct computations.

Introductory Example

Consider a point in $\mathbb{R}^3$ that is viewed by 2 different observers: B and C.



We have $\vec{p}$, $[\vec{p}]_B$, $[\vec{p}]_C$. This reads the vector p, then vector p with respect to B, and the vector p with respect to C. There's nothing else here, I don't know why he did this example.

Example

Let $B = \{\vec{b_1}, \vec{b_2}\}$ and $C = \{\vec{c_1}, \vec{c_2}\}$ be bases for a vector spec V, and suppose that $\vec{b_1} = -\vec{c_1} + 4\vec{c_2}$ and $\vec{b_2} = 5\vec{c_1} - 3\vec{c_2}$. Suppose that $\vec{x} = 5\vec{b_1} + 3\vec{b_2}$.

Find $[\vec{x}]_C$.

We can write $\vec{x}$ with respect to B. $\vec{x} = 5\vec{b_1} + 3\vec{b_2}$ can also be written as $[\vec{x}]_B = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$ (as coordinates).

$$[\vec{x}]_C = [5\vec{b_1} + 3\vec{b_2}]_C = 5[\vec{b_1}]_C + 3[\vec{b_2}]_C$$

Or we can write this as a matrix

$$[\vec{x}]_C = \begin{bmatrix} [\vec{b_1}]_C & [\vec{b_2}]_C \end{bmatrix} \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

$$[\vec{b_1}]_C = \begin{bmatrix} -1 \\ 4 \end{bmatrix} \qquad [\vec{b_2}]_C = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$

$$[\vec{x}]_C = \underbrace{\begin{bmatrix} -1 & 5 \\ 4 & -3 \end{bmatrix}} \begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 11 \end{bmatrix}$$

This is our change of coordinates matrix from B to C. We would write this to denote a change of basis matrix from B to C.

$$\underset{C \leftarrow B}{P} = \begin{bmatrix} -1 & 5 \\ 4 & -3 \end{bmatrix}$$

Theorem 1 *Let* $B = \left\{\vec{b_1}, \cdots, \vec{b_n}\right\}$ *and* $C = \{\vec{c_1}, \cdots, \vec{c_n}\}$ *be bases of a vector space* V*. Then there is a unique* $n \times n$ *matrix* $\underset{C \leftarrow B}{P}$ *such that*

$$[\vec{x}]_C = \underset{C \leftarrow B}{P} \cdot [\vec{x}]_B$$

The columns of $\underset{C \leftarrow B}{P}$ *are the* C*–coordinate vectors of the vectors in the basis* B*. That is,*

$$\underset{C \leftarrow B}{P} = \left[[\vec{b_1}]_C \cdots [\vec{b_n}]_C \right]$$

If we have $\underset{C \leftarrow B}{P}$, how do we go back? The matrix $\underset{C \leftarrow B}{P}$

- is invertible
- the vectors are linearly independent
- is square

$$[\vec{x}]_C = \underset{C \leftarrow B}{P} \cdot [\vec{x}]_B$$

$$\implies \underset{C \leftarrow B}{P^{-1}} \cdot [\vec{x}]_C = [\vec{x}]_B$$

We can also write $\underset{C \leftarrow B}{P^{-1}}$ as $\underset{B \leftarrow C}{P}$.

Different Dimensions

Can we create a change of basis matrix from one dimension to another? We would need a change of basis matrix of $\mathbb{R}^{m \times n}$, implying $m \neq n$.

If such a linear transformation existed, then it must be a **homeomorphism** (continuous mapping with a continuous inverse).

We can come up with mappings to accomplish this, but they won't strictly be a change of basis/coordinates mapping.

Using theorem 15, we can solve the change of basis problem. We only need the old basis relative to the new.

Example

Let $B = \{\vec{b_1}, \vec{b_2}\}$ and $C = \{\vec{c_1}, \vec{c_2}\}$ be bases for $\mathbb{R}^2$.

Find a change of basis matrix from B to C.

$$\vec{b_1} = \begin{bmatrix} -9 \\ 1 \end{bmatrix} \quad \vec{b_2} = \begin{bmatrix} -5 \\ -1 \end{bmatrix} \qquad\qquad \vec{c_1} = \begin{bmatrix} 1 \\ -4 \end{bmatrix} \quad \vec{c_2} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$$

Find $\underset{C \leftarrow B}{P}$.

We can express the b vectors with respect to C like this, with some unknowns

$$[\vec{b_1}]_C = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad [\vec{b_2}]_C = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

And write them with the change of basis matrix of C

$$\begin{bmatrix} \vec{c_1} & \vec{c_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \vec{b_1} \qquad \begin{bmatrix} \vec{c_1} & \vec{c_2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \vec{b_2}$$

We can solve these simultaneously with this augmented matrix

$$\begin{bmatrix} \vec{c_1} & \vec{c_2} & \vdots & \vec{b_1} & \vec{b_2} \end{bmatrix} = \begin{bmatrix} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -4 \end{bmatrix}$$

$$[\vec{b_1}]_C = \begin{bmatrix} 6 \\ -5 \end{bmatrix} \qquad [\vec{b_2}]_C = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

$$\underset{C \leftarrow B}{P} = \begin{bmatrix} [\vec{b_1}]_C & [\vec{b_2}]_C \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}$$

Observe that

$$\begin{bmatrix} \vec{c_1} & \vec{c_2} & \vdots & \vec{b_1} & \vec{b_2} \end{bmatrix} \sim \begin{bmatrix} I & \vdots & \underset{C \leftarrow B}{P} \end{bmatrix}$$

And to find $\underset{B \leftarrow C}{P}$, we just find $\underset{C \leftarrow B}{P^{-1}}$.

This is the cutoff for Exam 2. Only topics before now will be on Exam 2.

2 Orthogonality

Dillhoff talks about orthogonality but doesn't really describe it. Google tells me the definition is

1. of or involving right angles; at right angles.
2. (of variates) statistically independent.

We're interested in the second definition. Dillhoff gave an example of a vector triple which represents some machine learning models confidence that an object in an image is a car, human, or animal. The output of this model could look like (0.8, 3.9, 0.4) or (0.4, 0.6, 0.7). Orthogonality will help us determine outliers like the 3.9 in the first vector. In the second, there is a maximum value but nothing of significance.

Inner Product

Take two vectors in $\mathbb{R}^n$, $\vec{u}$ and $\vec{v}$. They are both $n \times 1$.

We can multiply them as

$$\vec{u} \cdot \vec{v}^T \qquad \in \mathbb{R}^{n \times n}$$

$$\vec{u}^T \cdot \vec{v} \qquad \in \mathbb{R}$$

The second product is just a single scalar. We can think of this as a summary statistic.

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

This is called the **inner product**, or the **dot product**.

$$\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u} = \sum_{i}^{n} u_i v_i$$

Theorem 2 (Chapter 6, Theorem 1) *let $\vec{u}$, $\vec{v}$, $\vec{w}$ be vectors in $\mathbb{R}^n$, let c be a scalar.*

a) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$

b) $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$

c) $(c\vec{u}) \cdot \vec{v} = c(\vec{u} \cdot \vec{v}) = \vec{u} \cdot (c\vec{v})$

d) $\vec{u} \cdot \vec{u} \geq 0$, *and* $\vec{u} \cdot \vec{u} = 0$ *if, and only if,* $\vec{u} = 0$.