Notes on Change of Basis

David Meyer

dmm@{1-4-5.net,uoregon.edu}

Last update: January 24, 2019

1 Introduction

This document contains a few of my notes made when I was reviewing change of basis in linear algebra.

2 Definitions

Consider a set $B \subseteq V$ where V is a n-dimensional vector space. B is a *basis* of V if every element of V may be written in a unique way as a (finite) linear combination of elements of B. The coefficients of this linear combination are referred to as components or *coordinates* on B of the vector. The coordinates of a given vector $x \in V$ expressed in basis B is denoted as $[x]_B$.

3 The Basis of a Vector Space

A basis B of a vector space V over a field $\mathbb{F}$ (e.g. $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$) is a linearly independent subset of V that spans V. This means that a subset B of V is a basis if the following two conditions hold:

- Linear Independence
  - for every finite subset of $\{b_1, \cdots, b_n\}$ of B and $\{a_1, \cdots, a_n\}$ of $\mathbb{F}$ if $a_1 b_1 + \ldots + a_n b_n = 0$ then $a_1 = \ldots = a_n = 0$
- *B spans V*
  - for every vector $\mathbf{v} \in V$ it is possible to choose $\{a_1, \cdots, a_n\}$ in $\mathbb{F}$ and $\{b_1, \cdots, b_n\}$ in B such that $\mathbf{v} = a_1 b_1 + \cdots + a_n b_n$.

The elements of B are called basis vectors and importantly both B and V are ordered sets.

The summary is that the elements of a basis B are linearly independent and every element of V can be represented as a linear combination of the elements of B with coefficients from the field $\mathbb{F}$.

4 The Standard Basis

The standard basis (also called natural basis) for a Euclidean space is the set of unit vectors pointing in the direction of the axes of a Cartesian coordinate system. For example, the standard basis for three-dimensional space is formed by the vectors

$$\mathbf{e}_x = (1,0,0), \quad \mathbf{e}_y = (0,1,0), \quad \mathbf{e}_z = (0,0,1)$$

These vectors are a basis in the sense that any other vector can be expressed uniquely as a linear combination of these. For example, every vector $\mathbf{v}$ in three-dimensional space can be written uniquely as

$$\mathbf{v} = v_x\mathbf{e}_x + v_y\mathbf{e}_y + v_z\mathbf{e}_z$$

where $v_x, v_y, v_z \in \mathbb{F}$ are the scalar "coordinates" of $\mathbf{v}$ in the standard basis. The *coordinate vector* for $\mathbf{v}$ is denoted $[\mathbf{v}]_B$ for basis B. Here $[\mathbf{v}]_{STD} = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$.

In general, the standard basis (which is sometimes called the computational basis) for the n-dimensional Euclidean space consists of the ordered set of n distinct vectors

$$\{\mathbf{e}_i : 1 \leq i \leq n\}$$

where $\mathbf{e}_i$ is the i^{th} basis vector which has a 1 in the i^{th} coordinate and 0's elsewhere.[1] That is, the $\mathbf{e}_i$ are the $n \times 1$ column vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \ldots, \mathbf{e}_{n-1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

[1] This is also called a "one-hot" encoding in machine learning parlance, where the i's might be the classes in a classification problem.

Standard bases can be defined for other vector spaces, such as polynomials and matrices. In both cases, the standard basis consists of the elements of the vector space such that all coefficients but one are 0 and the non-zero one is 1. For polynomials, the standard basis thus consists of the monomials and is commonly called monomial basis.

For example, the standard basis for $\mathcal{P}^2$, the set of polynomials of degree ≤ 2 over some field $\mathbb{F}$ is the ordered set

$$B_{\mathcal{P}^2} = \left\{ \begin{bmatrix} x^0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x^1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ x^2 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ x^2 \end{bmatrix} \right\}$$

By the definition of basis any polynomial $a + bx + cx^2 \in \mathcal{P}^2$ can be written as a linear combination of vectors from $B_{\mathcal{P}^2}$. That is

$$a + bx + cx^2 = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ x \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ x^2 \end{bmatrix}$$

for $a, b, c \in \mathbb{F}$.

We can also see this in matrix form by loading the vectors of $B_{\mathcal{P}^2}$ into a matrix. We call this matrix $P_{\mathcal{P}^2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x^2 \end{bmatrix}$. In matrix form:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a \cdot 1 & + & b \cdot 0 & + & c \cdot 0 \\ a \cdot 0 & + & b \cdot x & + & c \cdot 0 \\ a \cdot 0 & + & b \cdot 0 & + & c \cdot x^2 \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ x \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ x^2 \end{bmatrix} = a + bx + cx^2$$

This example shows an interesting property of vector spaces: once a basis is chosen for a n-dimensional vector space V we get an isomorphism between V and $\mathbb{F}^n$. For polynomials $a + bx + cx^2 \in \mathcal{P}^2$ over a field $\mathbb{F}$, we get the isomorphism $\mathcal{P}^2 \cong \mathbb{F}^3$. Here the mapping is

$$a + bx + cx^2 \longmapsto \begin{bmatrix} a \\ b \\ c \end{bmatrix} \text{ and } \begin{bmatrix} a \\ b \\ c \end{bmatrix} \longmapsto a + bx + cx^2$$

where $a + bx + cx^2 \in \mathcal{P}^2$, $\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{F}^3$, and $\mathbb{F}$ could be any field.

For matrices $\mathcal{M}_{m\times n}$, the standard basis consists of the $m \times n$-matrices with exactly one non-zero entry, which is 1. For example, the standard basis for 2×2 over $\mathbb{R}$ matrices is formed by the 4 matrices

$$\mathbf{e}_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{e}_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{e}_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{e}_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

5 Examples

Consider the bases $B = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix} \right\}$ and $C = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ of the vector space $\mathbb{R}^2$. Suppose further that we know that $[x]_B = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

First question: How do we find $[x]_C$? The brute force method is to convert

$$[x]_B \to [x]_{STD} \to [x]_C$$

To do this, first we multiply $[x]_B$ by the corresponding basis elements:

$$[x]_B = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \implies x = [x]_{STD} = 3 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \cdot \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

So $[x]_{STD} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$. Next we have to solve

$$[x]_C = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \implies x_1 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} + x_2 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

So

$$x_1 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} + x_2 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix} \tag{1}$$

First, recall that if $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\det(A) \neq 0$ then[2]

[2] Recall also that a $n \times n$ matrix A is invertible (A^{-1} exists) iff $\det(A) \neq 0$.

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{2}$$

and

$$\det(A) = ad - bc \tag{3}$$

Aside: The set of 2×2 invertible matrices over the real numbers with the group operation being matrix multiplication is called $GL(2, \mathbb{R})$, the *General Linear Group* over $\mathbb{R}$. That is,

$$\begin{aligned} GL(2, \mathbb{R}) &= \{2 \times 2 \text{ matrices } A \text{ over } \mathbb{R} \mid \det(A) \neq 0\} \\ &= \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid a, b, c, d \in \mathbb{R} \text{ and } ad - bc \neq 0 \right\} \end{aligned}$$

Next, rewriting Equation 1 in matrix form we get

$$\begin{aligned} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 5 \\ 3 \end{bmatrix} && \text{\# solve for } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ in matrix form} \\ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \frac{1}{\det\left(\begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}\right)} \begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \end{bmatrix} && \text{\# multiply both sides by } \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}^{-1}, \text{ Equation 2} \\ &= \tfrac{1}{-1} \begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \end{bmatrix} && \text{\# } \det\left(\begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}\right) = 0 \cdot 2 - 1 \cdot 1 = -1, \text{ Equation 3} \\ &= \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \end{bmatrix} && \text{\# scalar multiplication} \\ &= \begin{bmatrix} -7 \\ 5 \end{bmatrix} && \text{\# } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -7 \\ 5 \end{bmatrix} \end{aligned}$$

Note that since we need to compute A^{-1} the determinate of A cannot be zero. The set of $n \times n$ invertible matrices S are just those for which $A \in S \implies \det(A) \neq 0$.

Here the matrix $\begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$ is called P_C. Similarly, the matrix $\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ is called P_B. So we have

$$P_B[\mathbf{e}_1]_B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, P_B[\mathbf{e}_2]_B = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, P_C[\mathbf{e}_1]_C = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \text{ and } P_C[\mathbf{e}_2]_C = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

where $\mathbf{e}_i$ is the i^{th} element of the (ordered) basis. What we can observe here is that in general $P_B[x]_B = [x]_{STD}$ and $P_C[x]_C = [x]_{STD}$ and so $P_B[x]_B = P_C[x]_C$.

Note that the method above converts $[x]_B$ to $[x]_{STD}$ then to $[x]_C$. In general you wouldn't want to do this since we can compute the conversion matrix $P_{B\to C}$ directly without having to go through the standard basis.

So how do we compute $P_{B\to C}$? We saw above that $P_B[x]_B = P_C[x]_C$, so if we multiply both sides by P_C^{-1} we get $[x]_C = P_C^{-1}P_B[x]_B$, which tells us that $P_{B\to C} = P_C^{-1}P_B$.

Applying this result to our example above

$$
\begin{aligned}
[x]_C &= \underbrace{\frac{1}{-1}\begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix}}_{P_C^{-1}} \underbrace{\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}}_{P_B} \underbrace{\begin{bmatrix} 3 \\ 1 \end{bmatrix}}_{[x]_B} && \#\ [x]_C = P_C^{-1}P_B[x]_B = P_{B\to C}[x]_B \\
&= \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} && \#\ \text{multiply by -1} \\
&= \begin{bmatrix} -1 & -4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} && \#\ P_{B\to C} = \tfrac{1}{-1}\begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -4 \\ 1 & 2 \end{bmatrix} \\
&= \begin{bmatrix} -7 \\ 5 \end{bmatrix} && \#\ [x]_C = P_{B\to C}[x]_B = \begin{bmatrix} -1 & -4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} -7 \\ 5 \end{bmatrix}
\end{aligned}
$$

6 Acknowledgements