Change of basis, orthogonality and Gram-Schmidt algorithm

Francesco Preta

07/22/2020

Basis of $\mathbb{R}^n$

We have defined a basis of a vector space as a minimal set of vectors that span the entire space. The canonical basis is a basis for $\mathbb{R}^n$, but it's not the only one.

Example:

$$\mathscr{B} = \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Finding out if a given set of vectors is a basis

Theorem

Let $\mathscr{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ be a set of vectors in $\mathbb{R}^n$. Then $\mathscr{B}$ is a basis of $\mathbb{R}^n$ if and only if they are linearly independent, which happens if and only if the determinant of the matrix having i-th column $\mathbf{b}_i$ is not zero.

Changing basis

By the definition of a basis, we know that if $\mathscr{B} = \{\mathbf{b}_i\}_{i=1}^n$ is a basis, for each vector $\mathbf{v}$ there exists a unique set of coefficients y_i such that

$$\mathbf{v} = \sum_{i=1}^{n} y_i \mathbf{b}_i$$

Then the vector $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix}$ corresponds to the vector $\mathbf{v}$ in the coordinates given by the basis $\mathscr{B}$.

Change of basis matrix

We want to be able to write every vector $\mathbf{v}$ as the corresponding linear combination $\mathbf{v} = \sum_{i=1}^{n} y_i \mathbf{b}_i$.

In order to do so, we need to find $\mathbf{y}$ by solving the system $B\mathbf{y} = \mathbf{v}$.

Change of basis matrix (continued)

Since $\mathscr{B}$ is a basis, the solution exists for every $\mathbf{v}$ and is given by
$\mathbf{y} = B^{-1}\mathbf{v}$.

The matrix B^{-1} is called the **change of basis matrix.**

Example:

$$\mathscr{B} = \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Example: (continued)

write the vector $\mathbf{v} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$ in the coordinates given by the basis $\mathscr{B}$

Orthonormal basis

Definition

Let $\mathcal{U} = \{\mathbf{u}_i\}_{i=1}^n$ be a basis for $\mathbb{R}^n$. We say that $\mathcal{U}$ is an **orthogonal basis** if

$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ for every $i \neq j$.

Moreover, we say that $\mathcal{U}$ is **orthonormal** if it is orthogonal and

$||\mathbf{u}_i|| = 1$ for every $i = 1,\ldots,n$.

Kronecker delta

In short, one can say that a basis $\{\mathbf{u}_i\}_{i=1}^{n}$ is orthonormal if

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{ij}$$

Where δ_{ij} is called the **Kronecker delta**, taking value 1 when $i = j$ and 0 otherwise.

The canonical basis is orthonormal

The canonical basis is an orthonormal basis, since

- $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0$ for $i \neq j$,
- $\langle \mathbf{e}_i, \mathbf{e}_i \rangle = 1$ for every $i = 1,\ldots,n$.

However, there are other orthogonal basis which are not the canonical basis.

Example:

$$\mathscr{U} = \left\{ \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Orthonormal basis and orthogonal matrices

The advantage of an orthogonal basis $\mathscr{U} = \{\mathbf{u}_i\}_{i=1}^n$ is that the matrix U having i-th column equal to $\mathbf{u}_i$ is such that $U^T = U^{-1}$. We call such matrices **orthogonal**.

Proof:

Change of basis matrix for orthonormal basis

Since for any basis $\mathscr{B}$, the change of basis matrix is the inverse of the matrix B having the element $\mathbf{b}_i$ as i-th column, for orthonormal basis we have the following:

Theorem

Let $\mathscr{U} = \{\mathbf{u}_i\}_{i=1}^n$ be an orthonormal basis. Then the change of basis matrix from the canonical basis to $\mathscr{U}$ is the matrix U^T having $\mathbf{u}_i$ as i-th row.

Example:

Write the vector $\mathbf{v} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$ in the basis

$$\mathscr{U} = \left\{ \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

The Gram-Schmidt algorithm

Having an orthonormal basis decreases computational complexity with respect to basis that are not orthonormal. The Gram-Schmidt algorithm "normalizes" any given basis by returning an orthonormal basis associated to it.

Algorithm steps

Given a basis $\{\mathbf{b}_i\}_{i=1}^n$ of $\mathbb{R}^n$. Initialize the algorithm by normalizing the first vector, that is

$$\mathbf{u}_1 = \frac{\mathbf{b}_1}{||\mathbf{b}_1||}$$

Then for each $i = 2,\ldots,n$, we first create a vector that is orthogonal to all the previous vectors

$$\mathbf{u}'_i = \mathbf{b}_i - \sum_{j<i} \langle \mathbf{b}_i, \mathbf{u}_j \rangle \mathbf{u}_j$$

and then we normalize it

$$\mathbf{u}_i = \frac{\mathbf{u}'_i}{||\mathbf{u}'_i||}.$$

Proof that $\mathbf{u}'_k$ is orthogonal to $\mathbf{u}_j$ for all $j < k$:

Example:

Apply the Gram-Schmidt algorithm to the basis

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \qquad \mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{b}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Gram-Schmidt on independent vectors

Notice that the Gram-Schmidt algorithm can be applied to any set of independent vectors, even if they are not a basis. The output will still be given by orthonormal vectors (which, of course, do not constitute a basis).

Example:

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix} \qquad \mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix} \qquad \mathbf{b}_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$