21. Orthonormal Bases

The canonical/standard basis

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

has many useful properties.

- Each of the standard basis vectors has unit length:

$$||e_i|| = \sqrt{e_i \cdot e_i} = \sqrt{e_i^T e_i} = 1.$$

- The standard basis vectors are *orthogonal* (in other words, at right angles or perpendicular).

$$e_i \cdot e_j = e_i^T e_j = 0 \text{ when } i \neq j$$

This is summarized by

$$e_i^T e_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases},$$

where δ_{ij} is the *Kronecker delta*. Notice that the Kronecker delta gives the entries of the identity matrix.

Given column vectors v and w, we have seen that the dot product $v \cdot w$ is the same as the matrix multiplication $v^T w$. This is the *inner product* on $\mathbb{R}^n$. We can also form the *outer product* vw^T, which gives a square matrix.

The outer product on the standard basis vectors is interesting. Set

$$\begin{aligned}
\Pi_1 &= e_1 e_1^T \\
&= \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \\
&\vdots \\
\Pi_n &= e_n e_n^T \\
&= \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & \dots & 1 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}
\end{aligned}$$

In short, Π_i is the diagonal square matrix with a 1 in the ith diagonal position and zeros everywhere else. [1]

Notice that $\Pi_i \Pi_j = e_i e_i^T e_j e_j^T = e_i \delta_{ij} e_j^T$. Then:

$$\Pi_i \Pi_j = \begin{cases} \Pi_i & i = j \\ 0 & i \neq j \end{cases} .$$

Moreover, for a diagonal matrix D with diagonal entries $\lambda_1, \dots, \lambda_n$, we can write

$$D = \lambda_1 \Pi_1 + \dots + \lambda_n \Pi_n.$$

Other bases that share these properties should behave in many of the same ways as the standard basis. As such, we will study:

[1] This is reminiscent of an older notation, where vectors are written in juxtaposition. This is called a 'dyadic tensor,' and is still used in some applications.

- *Orthogonal bases* $\{v_1, \dots, v_n\}$:

$$v_i \cdot v_j = 0 \text{ if } i \neq j$$

  In other words, all vectors in the basis are perpendicular.

- *Orthonormal bases* $\{u_1, \dots, u_n\}$:

$$u_i \cdot u_j = \delta_{ij}.$$

  In addition to being orthogonal, each vector has unit length.

Suppose $T = \{u_1, \dots, u_n\}$ is an orthonormal basis for $\mathbb{R}^n$. Since T is a basis, we can write any vector v uniquely as a linear combination of the vectors in T:

$$v = c^1 u_1 + \dots c^n u_n.$$

Since T is orthonormal, there is a very easy way to find the coefficients of this linear combination. By taking the dot product of v with any of the vectors in T, we get:

$$\begin{aligned} v \cdot u_i &= c^1 u_1 \cdot u_i + \dots + c^i u_i \cdot u_i + \dots + c^n u_n \cdot u_i \\ &= c^1 \cdot 0 + \dots + c^i \cdot 1 + \dots + c^n \cdot 0 \\ &= c^i, \\ \Rightarrow c^i &= v \cdot u_i \\ \Rightarrow v &= (v \cdot u_1) u_1 + \dots + (v \cdot u_n) u_n \\ &= \sum_i (v \cdot u_i) u_i. \end{aligned}$$

This proves the theorem:

Theorem. *For an orthonormal basis* $\{u_1, \dots, u_n\}$*, any vector* v *can be expressed*

$$v = \sum_i (v \cdot u_i) u_i.$$

Relating Orthonormal Bases

Suppose $T = \{u_1, \dots, u_n\}$ and $R = \{w_1, \dots, w_n\}$ are two orthonormal bases for $\mathbb{R}^n$. Then:

$$\begin{aligned}
w_1 &= (w_1 \cdot u_1)u_1 + \ldots + (w_1 \cdot u_n)u_n \\
&\vdots \\
w_n &= (w_n \cdot u_1)u_1 + \ldots + (w_n \cdot u_n)u_n \\
\Rightarrow w_i &= \sum_j u_j(u_j \cdot w_i)
\end{aligned}$$

As such, the matrix for the change of basis from T to R is given by

$$P = (P_i^j) = (u_j \cdot w_i).$$

Consider the product PP^T in this case.

$$\begin{aligned}
(PP^T)_k^j &= \sum_i (u_j \cdot w_i)(w_i \cdot u_k) \\
&= \sum_i (u_j^T w_i)(w_i^T u_k) \\
&= u_j^T \left[\sum_i (w_i w_i^T) \right] u_k \\
&= u_j^T I_n u_k \qquad (*) \\
&= u_j^T u_k = \delta_{jk}.
\end{aligned}$$

In the equality $(*)$ is explained below. So assuming $(*)$ holds, we have shown that $PP^T = I_n$, which implies that

$$P^T = P^{-1}.$$

The equality in the line $(*)$ says that $\sum_i w_i w_i^T = I_n$. To see this, we examine $(\sum_i w_i w_i^T)v$ for an arbitrary vector v. We can find constants c^j such that $v = \sum_j c^j w_j$, so that:

$$\begin{aligned}
(\sum_i w_i w_i^T)v &= (\sum_i w_i w_i^T)(\sum_j c^j w_j) \\
&= \sum_j c^j \sum_i w_i w_i^T w_j \\
&= \sum_j c^j \sum_i w_i \delta_{ij} \\
&= \sum_j c^j w_j \text{ since all terms with } i \neq j \text{ vanish} \\
&= v.
\end{aligned}$$

Then as a linear transformation, $\sum_i w_i w_i^T = I_n$ fixes every vector, and thus must be the identity I_n.

Definition A matrix P is *orthogonal* if $P^{-1} = P^T$.

Then to summarize,

Theorem. *A change of basis matrix P relating two orthonormal bases is an orthogonal matrix. i.e.*

$$P^{-1} = P^T.$$

Example Consider $\mathbb{R}^3$ with the orthonormal basis

$$S = \left\{ u_1 = \begin{pmatrix} \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{6}} \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, u_3 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \right\}.$$

Let R be the standard basis $\{e_1, e_2, e_3\}$. Since we are changing from the standard basis to a new basis, then the columns of the change of basis matrix are exactly the images of the standard basis vectors. Then the change of basis matrix from R to S is given by:

$$\begin{aligned} P = (P_i^j) = (e_j u_i) &= \begin{pmatrix} e_1 \cdot u_1 & e_1 \cdot u_2 & e_1 \cdot u_3 \\ e_2 \cdot u_1 & e_2 \cdot u_2 & e_2 \cdot u_3 \\ e_3 \cdot u_1 & e_3 \cdot u_2 & e_3 \cdot u_3 \end{pmatrix} \\ = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} &= \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} \\ \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}. \end{aligned}$$

From our theorem, we observe that:

$$\begin{aligned} P^{-1} = P^T &= \begin{pmatrix} u_1^T \\ u_2^T \\ u_3^T \end{pmatrix} \\ &= \begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}. \end{aligned}$$

We can check that $P^T P = I_n$ by a lengthy computation, or more simply, notice that

$$\begin{aligned}
(P^T P)_{ij} &= \begin{pmatrix} u_1^T \\ u_2^T \\ u_3^T \end{pmatrix} \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\end{aligned}$$

We are using orthonormality of the u_i for the matrix multiplication above.

Orthonormal Change of Basis and Diagonal Matrices. Suppose D is a diagonal matrix, and we use an orthogonal matrix P to change to a new basis. Then the matrix M of D in the new basis is:

$$M = PDP^{-1} = PDP^T.$$

Now we calculate the transpose of M.

$$\begin{aligned}
M^T &= (PDP^T)^T \\
&= (P^T)^T D^T P^T \\
&= PDP^T \\
&= M
\end{aligned}$$

So we see the matrix PDP^T is symmetric!

References

- Hefferon, Chapter Three, Section V: Change of Basis

Wikipedia:

- Orthogonal Matrix
- Diagonalizable Matrix
- Similar Matrix

Review Questions

1. Let $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

   i. Write D in terms of the vectors e_1 and e_2, and their transposes.

   ii. Suppose $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible. Show that D is similar to

   $$M = \frac{1}{ad - bc} \begin{pmatrix} \lambda_1 ad - \lambda_2 bc & (\lambda_1 - \lambda_2) bd \\ (\lambda_1 - \lambda_2) ac & -\lambda_1 bc + \lambda_2 ad \end{pmatrix}.$$

   iii. Suppose the vectors $\begin{pmatrix} a & b \end{pmatrix}$ and $\begin{pmatrix} c & d \end{pmatrix}$ are orthogonal. What can you say about M in this case?

2. Suppose $S = \{v_1, \ldots, v_n\}$ is an *orthogonal* (not orthonormal) basis for $\mathbb{R}^n$. Then we can write any vector v as $v = \sum_i c^i v_i$ for some constants c^i. Find a formula for the constants c^i in terms of v and the vectors in S.

3. Let u, v be independent vectors in $\mathbb{R}^3$, and $P = \operatorname{span}\{u, v\}$ be the plane spanned by u and v.

   i. Is the vector $v^\perp = v - \frac{u \cdot v}{u \cdot u} u$ in the plane P?

   ii. What is the angle between $v^\perp$ and u?

   iii. Given your solution to the above, how can you find a third vector perpendicular to both u and $v^\perp$?

   iv. Construct an orthonormal basis for $\mathbb{R}^3$ from u and v.

   v. Test your abstract formulae starting with

   $$u = \begin{pmatrix} 1 & 2 & 0 \end{pmatrix} \text{ and } v = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix}.$$