20. Diagonalization

Let V and W be vector spaces, with bases $S = \{e_1, \dots, e_n\}$ and $T = \{f_1, \dots, f_m\}$ respectively. Since these are bases, there exist constants v^i and w^j such that any vectors $v \in V$ and $w \in W$ can be written as:

$$\begin{aligned} v &= v^1 e_1 + v^2 e_2 + \dots + v^n e_n \\ w &= w^1 f_1 + w^2 f_2 + \dots + w^m f_m \end{aligned}$$

We call the coefficients $v^1, \dots, v^n$ the *components* of v in the basis $\{e_1, \dots, e_n\}$.

Example Consider the basis $S = \{1-t, 1+t\}$ for the vector space $P_1(t)$. The vector $v = 2t$ has components $v^1 = -1, v^2 = 1$, because

$$v = -1(1-t) + 1(1+t).$$

We may consider these components as vectors in $\mathbb{R}^n$ and $\mathbb{R}^m$:

$$\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \in \mathbb{R}^n, \qquad \begin{pmatrix} w^1 \\ \vdots \\ w^m \end{pmatrix} \in \mathbb{R}^m.$$

Now suppose we have a linear transformation $L : V \to W$. Then we can expect to write L as an $m \times n$ matrix, turning an n-dimensional vector of coefficients corresponding to v into an m-dimensional vector of coefficients for w.

Using linearity, we write:

$$\begin{aligned} L(v) &= L(v^1 e_1 + v^2 e_2 + \dots + v^n e_n) \\ &= v^1 L(e_1) + v^2 L(e_2) + \dots + v^n L(e_n). \end{aligned}$$

This is a vector in W. Let's compute its components in W.

We know that for each e_j, $L(e_j)$ is a vector in W, and can thus be written uniquely as a linear combination of vectors in the basis T. Then we can find coefficients M^i_j such that:

$$L(e_j) = f_1 M^1_j + \dots + f_m M^m_j = \sum_{i=1}^{m} f_i M^i_j.$$

We've written the M^i_j on the right side of the f's to agree with our previous notation for matrix multiplication. We have an "up-hill rule" where the matching indices for the multiplied objects run up and to the left, like so: $f_i M^i_j$.

Now M^i_j is the ith component of $L(e_j)$. Regarding the coefficients M^i_j as a matrix, we can see that the jth column of M is the coefficients of $L(e_j)$ in the basis T.

Then we can write:

$$\begin{aligned} L(v) &= L(v^1 e_1 + v^2 e_2 + \ldots + v^n e_n) \\ &= v^1 L(e_1) + v^2 L(e_2) + \ldots + v^n L(e_n) \\ &= \sum_{i=1}^{m} v^j L(e_j) \\ &= \sum_{i=1}^{m} v^j (M^1_j f_1 + \ldots + M^m_j f_m) \\ &= \sum_{i=1}^{m} f_i [\sum_{j=1}^{n} M^i_j v^j]. \end{aligned}$$

The last equality is the definition of matrix multiplication. Thus:

$$\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \stackrel{L}{\mapsto} \begin{pmatrix} M^1_1 & \ldots & M^1_n \\ \vdots & & \vdots \\ M^m_1 & \ldots & M^m_n \end{pmatrix} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix},$$

and $M = (M^i_j)$ is called the matrix of L. Notice that this matrix depends on a *choice* of bases for V and W.

Example Let $L : P_1(t) \mapsto P_1(t)$, such that $L(a + bt) = (a + b)t$. Since $V = P_1(t) = W$, let's choose the same basis for V and W. We'll choose the basis $\{1 - t, 1 + t\}$ for this example.

Thus:

$$\begin{aligned} L(1-t) = (1-1)t = \quad & 0 \quad = (1-t)\cdot 0 + (1+t)\cdot 0 = \begin{pmatrix} (1-t) & (1+t) \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ L(1+t) = (1+1)t = \quad & 2t \quad (1-t)\cdot 1 + (1+t)\cdot 1 = \begin{pmatrix} (1-t) & (1+t) \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \Rightarrow M \quad & = \quad \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Now suppose we are lucky, and we have $L : V \mapsto V$, and the basis $\{v_1, \ldots, v_n\}$ is a set of linearly independent eigenvectors for L, with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then:

$$\begin{aligned} L(v_1) &= \lambda_1 v_1 \\ L(v_2) &= \lambda_2 v_2 \\ &\vdots \\ L(v_n) &= \lambda_n v_n \end{aligned}$$

As a result, the matrix of L is:

$$M = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

where all entries off of the diagonal are zero.

We call the $n \times n$ matrix of a linear transformation $L : V \mapsto V$ *diagonalizable* if there exists a collection of n linearly independent eigenvectors for L. In other words, L is diagonalizable if there exists a basis for V of eigenvectors for L.

In a basis of eigenvectors, the matrix of a linear transformation is diagonal.

On the other hand, if an $n \times n$ matrix M is diagonal, then the standard basis vectors e_i are already a set of n linearly independent eigenvectors for M.

Change of Basis

Suppose we have two bases $S = \{v_1, \ldots, v_n\}$ and $T = \{u_1, \ldots, u_n\}$ for a vector space V. Then we may write each v_i uniquely as a linear combination of the u_j:

$$v_i = \sum_j u_j P^i_j.$$

Here, the P^i_j are constants, which we can regard as a matrix $P = (P^i_j)$. P must have an inverse, since we can also write each u_j uniquely as a linear combination of the v_i:

$$u_j = \sum_k v_k Q^k_j.$$

Then we can write:

$$v_i = \sum_k \sum_j v_k Q^k_j P^j_i.$$

But $\sum_j v_k Q^k_j P^j_i$ is the i, j entry of the product matrix QP. Since the only expression for v_i in the basis S is v_i itself, then QP fixes each v_i. As a result, each v_i is an eigenvector for QP with eigenvalues 1, so QP is the identity.

The matrix P is then called a *change of basis* matrix.

Changing basis changes the matrix of a linear transformation. To wit, suppose $L : V \mapsto V$ has matrix $M = (M^i_j)$ in the basis $T = \{u_1, \ldots, u_n\}$, so

$$L(u_i) = \sum_k M^k_i u_k.$$

Now, suppose that $S = \{v_1, \ldots, v_n\}$ is a basis of eigenvectors for L, with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then

$$L(v_i) = \lambda_i v_i = \sum_k v_k D^k_i$$

where D is the diagonal matrix whose diagonal entries D^k_k are the eigenvalues λ_k. Let P be the change of basis matrix from the basis T to the basis S. Then:

$$L(v_i) = L(\sum_j u_j P^j_i) = \sum_j L(u_j) P^j_i = \sum_j \sum_k u_k M^k_j P^j_i.$$

Meanwhile, we have:

$$L(v_i) = \sum_k v_k D^k_i = \sum_k \sum_j u_j P^j_k D^k_i.$$

In other words, we see that

$$MP = PD \text{ or } D = P^{-1} M P.$$

We can summarize as follows:

- Change of basis multiplies vectors by the change of basis matrix P, to give vectors in the new basis.

- To get the matrix of a linear transformation in the new basis, we *conjugate* the matrix of L by the change of basis matrix: $M \to P^{-1} M P$.

If for two matrices N and M there exists an invertible matrix P such that $M = P^{-1}NP$, then we say that M and N are *similar*. Then the above discussion shows that diagonalizable matrices are similar to diagonal matrices.

References

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

Wikipedia:

- Change of Basis
- Diagonalizable Matrix
- Similar Matrix

Review Questions

1. Show that similarity of matrices is an *equivalence relation*. (The definition of an equivalence relation is given in Section 2, in the fourth review problem.)

2. When is the 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ diagonalizable? Include examples in your answer.

3. Let $P_n(t)$ be the vector space of degree n polynomials, and $\frac{d}{dt} : P_n(t) \mapsto P_{n-1}(t)$ be the derivative operator. Find the matrix of $\frac{d}{dt}$ in the bases $\{1, t, \ldots, t^n\}$ for $P_n(t)$ and $\{1, t, \ldots, t^{n-1}\}$ for $P_{n-1}(t)$.

4. When writing a matrix for a linear transformation, we have seen that the choice of basis matters. In fact, even the order of the basis matters!

    - Write all possible reorderings of the standard basis $\{e_1, e_2, e_3\}$ for $\mathbb{R}^3$.
    - Write each change of basis matrix between the standard basis $\{e_1, e_2, e_3\}$ and each of its reorderings. What can you observe about these change of basis matrices? (Note: These matrices are known as *permutation matrices*.)

- Given the linear transformation $L(x, y, z) = (2y - z, 3x, 2z + x + y)$, write the matrix M for L in the standard basis, and two other reorderings of the standard basis. Can you make any observations about the resulting matrices?