Diagonalization Theory

Previously

- We have been working towards trying to find for which linear mappings there exists a basis $\mathcal{B}$ such that the matrix of the linear mapping with respect to the basis $\mathcal{B}$ is diagonal.

Diagonalization Theory

Definition: An $n \times n$ matrix A is said to be diagonalizable if there exists an invertible matrix P such that $P^{-1}AP = D$ is diagonal. We say that P diagonalizes A.

Diagonalization Theory

Recall that we previously showed that if $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis for $\mathbb{R}^n$ of eigenvectors of a matrix A, and $\lambda_1, \dots, \lambda_n$ are eigenvalues of A corresponding to $\vec{v}_1, \dots, \vec{v}_n$ respectively, then taking $P = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$ gives

$$P^{-1}AP = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$

Diagonalization Theory

Lemma 6.3.1

Suppose that A is $n \times n$ and that $\lambda_1, \dots, \lambda_k$ are distinct eigenvalues with corresponding eigenvectors $\vec{v}_1, \dots, \vec{v}_k$, then $\{\vec{v}_1, \dots, \vec{v}_k\}$ is linearly independent.

Proof

If $k = 1$, then we have that $\{\vec{v}_1\}$ is linearly independent since $\vec{v}_1 \neq \vec{0}$ by definition of an eigenvector.
Assume that the result is true for some $k \geq 1$.
To show $\{\vec{v}_1, \dots, \vec{v}_k, \vec{v}_{k+1}\}$ is linearly independent we consider

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k + c_{k+1}\vec{v}_{k+1} = \vec{0} \qquad (1)$$

Since $A\vec{v}_i = \lambda_i\vec{v}_i$ we get $(A - \lambda_i I)\vec{v}_i = \vec{0}$ and

$$(A - \lambda_i I)\vec{v}_j = A\vec{v}_j - \lambda_i\vec{v}_j = \lambda_j\vec{v}_j - \lambda_i\vec{v}_j = (\lambda_j - \lambda_i)\vec{v}_j$$

Multiplying both sides of (1) by $A - \lambda_{k+1}I$ gives

$$c_1(\lambda_1 - \lambda_{k+1})\vec{v}_1 + \cdots + c_k(\lambda_k - \lambda_{k+1})\vec{v}_k + \vec{0} = \vec{0}$$

By our induction hypothesis, $\{\vec{v}_1, \dots, \vec{v}_k\}$ is linearly independent and so $c_i(\lambda_i - \lambda_{k+1}) = 0$ for $1 \leq i \leq k$.
But, $\lambda_i \neq \lambda_{k+1}$ for $1 \leq i \leq k$, so we must have $c_1 = \cdots = c_k = 0$.
Thus, (1) becomes

$$0 + c_{k+1}\vec{v}_{k+1} = \vec{0}$$

Since $\vec{v}_{k+1} \neq \vec{0}$ we get $c_{k+1} = 0$.
Consequently, $\{\vec{v}_1, \dots, \vec{v}_{k+1}\}$ is linearly independent as required. □

Diagonalization Theory

Lemma 6.3.2

If A is matrix with distinct eigenvalues $\{\lambda_1, \dots, \lambda_k\}$ and $\mathcal{B}_i = \{\vec{v}_{i,1}, \dots, \vec{v}_{i,g_{\lambda_i}}\}$ is a basis for the eigenspace of λ_i for $1 \leq i \leq k$, then

$$\mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$$

is a linearly independent set.

Theorem 6.3.3 - The Diagonalization Theorem

Let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of a matrix A. Then, A is diagonalizable if and only if $g_{\lambda_i} = a_{\lambda_i}$ for $1 \leq i \leq k$.

Corollary 6.3.4

If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

Diagonalization Theory

Notes

1. If A is diagonalizable, then there exists an invertible matrix P and diagonal matrix D such that
$$P^{-1}AP = D$$
Notice that we can multiply by P on the left and by P^{-1} on the right to get a matrix factorization of A:
$$A = PDP^{-1}$$
2. Since we are currently only concerned with real numbers, we should modify the Diagonalization Theorem to say that a real matrix A is diagonalizable over the real numbers if and only if all the eigenvalues of A are real and $g_{\lambda_i} = a_{\lambda_i}$ for $1 \leq i \leq k$. If A has at least one non-real eigenvalue and hence a non-real eigenvector, then we say that A is not diagonalizable over $\mathbb{R}$.

Diagonalization Theory

Algorithm

Let $L : \mathbb{R}^n \to \mathbb{R}^n$ be a linear mapping with standard matrix $A = [L]$.
To diagonalize the $n \times n$ matrix A, or show that A is not diagonalizable:

1. Find and factor the characteristic polynomial $C(\lambda) = \det(A - \lambda I)$.
2. Let $\lambda_1, \dots, \lambda_n$ denote the n-roots of $C(\lambda)$ (repeated according to multiplicity). If any of the eigenvalues λ_i are not real, then A is not diagonalizable over $\mathbb{R}$.
3. Find a basis for the eigenspace of each λ_i by finding a basis for the nullspace of $A - \lambda_i I$.
4. If $g_{\lambda_i} < a_{\lambda_i}$ for any λ_i, then A is not diagonalizable. Otherwise, form a basis $\{\vec{v}_1, \dots, \vec{v}_n\}$ for $\mathbb{R}^n$ of eigenvectors of A by combining the eigenvectors in the bases for each eigenspace of A. Let $P = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$. Then,
$$P^{-1}AP = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$
where λ_i is an eigenvalue corresponding to the eigenvector $\vec{v}_i$ for $1 \leq i \leq n$. That is, if we take $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$, then $[L]_{\mathcal{B}} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.