7.11, 8.1, 8.2

Diagonalization

P. Danziger

1 Change of Basis

Given a basis of $\mathbb{R}^n$, $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, we have seen that the matrix whose columns consist of these vectors can be thought of as a change of basis matrix.

$$A_B = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{pmatrix}$$

Given a vector $\mathbf{u} \in \mathbb{R}^n$, we write $\mathbf{u}_B$, to represent the components of $\mathbf{u}$ with respect to the basis B. We write S for the standard basis, so $\mathbf{u}_S$ are the components of $\mathbf{u}$ with respect to the standard basis.
Given a vector $\mathbf{u} \in \mathbb{R}^n$, the components of $\mathbf{u}$ with respect to the basis B by finding the solutions to $A_B\mathbf{u}_B = \mathbf{u}_S$. Since B is a basis, A_B is invertible, and we can find $\mathbf{u}_B$ by

$$\mathbf{u}_B = A_B^{-1}\mathbf{u}_S.$$

[Note in the books rather confusing notation A_B would be called $A_{B\to S}$ and A_B^{-1} is $A_{S\to B}$.]
A_B can be thought of as a passive transformation, instead of thinking of the transformation T_{A_B} associated with A_B as actively moving the space, we think of it as passively moving the coordinates.

Now, note that if an $n \times n$ matrix P is invertible then the columns of P form a basis for $\mathbb{R}^n$. So every invertible matrix is a change of basis matrix. P^{-1} can be thought of as the matrix which transforms to the coordinate system consisting of the columns of P. P then transforms such a coordinate system back to the original coordinate system.
Let B be the column vectors of P. Given a vector $\mathbf{u} \in \mathbb{R}^n$, we can find $\mathbf{u}_B$ by solving the system of equations $P\mathbf{u}_B = \mathbf{u}_S$. We can solve this equation using the inverse $\mathbf{u}_B = P^{-1}\mathbf{u}_S$. Thus the components of $\mathbf{u}$ with respect to the basis given by the column vectors of P are

$$\mathbf{u}_B = P^{-1}\mathbf{u}_S.$$

Example 1

1. Find the components of $\mathbf{u} = (2, 1)$ with respect to the basis $B = \{(1, 1), (1, 2)\}$.

$$\text{Consider } P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad P^{-1}\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

So $\mathbf{u}_B = (3, -1)_B$ are the components of the vector $\mathbf{u}$ with respect to the basis $\{(1, 1), (1, 2)\}$.

2. Given that a vector has components $(3, -1)_B$ with respect to the basis $B = \{(1, 1), (1, 2)\}$, find the components of $\mathbf{u}$ with respect to the standard basis.

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

So $(3, -1)_P$ is the vector $(2, 1)$.

2 Similarity

Definition 2 *Two $n \times n$ matrices, A and B, are* similar *if there exists an invertible matrix P such that*

$$B = P^{-1}AP.$$

Notes

- If A is similar to B, so $B = P^{-1}AP$, then $PBP^{-1} = A$, and so B is similar to A.
- If A is similar to B, so $B = P^{-1}AP$, then $PB = AP$.

If A and B are similar matrices, so $B = P^{-1}AP$, then B can be thought of as preforming the same transformation as A, but on vectors given with respect to the basis given by the columns of $P-1$. Let C be the basis formed by the columns of P^{-1} and for any given $\mathbf{u} \in \mathbb{R}^n$ let $\mathbf{v} = T_A(\mathbf{u})$. Now

$$T_B(\mathbf{u}) = B\mathbf{u} = P^{-1}AP\mathbf{u} = P^{-1}A\mathbf{u}_C = P^{-1}\left(T_A(\mathbf{u})\right)_C = P^{-1}\mathbf{v}_C = \mathbf{v}$$

Example 3

Find a matrix similar to $A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$.

$$P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

$$\begin{aligned} \text{So} \quad B = P^{-1}AP &= \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 2 & 3 \end{pmatrix} \\ &= \begin{pmatrix} -2 & -5 \\ 2 & 4 \end{pmatrix} \end{aligned}$$

So with respect to the basis $\{(2, -1), (-1, 1)\}$, A looks like $\begin{pmatrix} -2 & -5 \\ 2 & 4 \end{pmatrix}$

Theorem 4 *If A and B are similar matrices, then they have the same characteristic polynomial, and hence the same eigenvalues.*

Proof: Let A and B be similar matrices, so $B = P^{-1}AP$ for some invertible matrix P.

$$\begin{array}{rcll} |\lambda I - B| & = & |\lambda I - P^{-1}AP| & \\ & = & |P^{-1}P(\lambda I) - P^{-1}AP| & (P^{-1}P = I) \\ & = & |P^{-1}(\lambda I)P - P^{-1}AP| & (PI = IP) \\ & = & |P^{-1}(\lambda I - A)P| & \text{(Distribution)} \\ & = & |P^{-1}|\,|\lambda I - A|\,|P| & (|XY| = |X|\,|Y|) \\ & = & |\lambda I - A| & (|P^{-1}|\,|P| = 1)\square \end{array}$$

3 Diagonalization

Definition 5 *A matrix is said to be* diagonalizable *if it is similar to a diagonal matrix.*

Consider an $n \times n$ matrix A with n linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, which thus form a basis B. Recall that the matrix associated with a transformation has columns given by the images of the basis vectors. With respect to this basis, $(T_A(\mathbf{v}_i))_B = \lambda_i (\mathbf{v}_i)_B = \lambda_i \mathbf{e}_i$. If D is the matrix of A with respect to this basis D is diagonal.
Let P be the matrix whose columns are the linearly independent eigenvectors of A, then A in this coordinate system will be D, where $A = PDP^{-1}$, or more usefully $D = P^{-1}AP$.

Theorem 6 *An $n \times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ corresponding to eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectivley (with possible repitition). In this case A is similar to the matrix $D = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$.*
$D = P^{-1}AP$, where P is the matrix whose columns are the eigenvectors of A. i.e. $P = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{pmatrix}$

Note not all matrices have n linearly independent eigenvectors.

Theorem 7 *Given an $n \times n$ matrix A, the following statements are equivalent:*

- *A is diagonalizable.*
- *A has n linearly independent Eigenvectors.*
- *The algebraic multiplicity equals the geometric multiplicity for each eigenvalue of A.*

Example 8

1. Diagonalize $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

   Eigenvalues $\lambda = 0, 2$, with corresponding eigenvectors $(1, -1)$ and $(1, 1)$.

$$P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

$$D = \text{diag}(0,2) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

$$\begin{aligned} P^{-1}AP &= \tfrac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \\ &= \tfrac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 2 \\ 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \end{aligned}$$

2. Diagonalize $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix}$

   Eigenvalues $\lambda = 1$ and $\lambda = 2$.

   But Algebraic multiplicity of $\lambda = 1$ is 2, whereas the geometric multiplicity is 1.

   So A is **not** diagonalizable.

4 Applications

4.1 Powers of Matrices

Theorem 9 *If $D = P^{-1}AP$, then $A^k = PD^kP^{-1}$*

Proof: $D = P^{-1}AP$, so

$$\begin{aligned} D^k &= \underbrace{(P^{-1}AP)(P^{-1}AP)\ldots(P^{-1}AP)}_{k} \\ &= \underbrace{P^{-1}A(PP^{-1})A(PP^{-1})A\ldots A(PP^{-1})AP}_{k} \\ &= P^{-1}\underbrace{AAA\ldots AA}_{k}P \\ &= P^{-1}A^kP \end{aligned}$$

Example 10

1. Find A^{10}, where $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

$$P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

$$D = \text{diag}(0,2) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

So

$$D^{10} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}^{10} = \begin{pmatrix} 0 & 0 \\ 0 & 2^{10} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1024 \end{pmatrix}$$

$$\begin{aligned} PD^{10}P^{-1} &= \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1024 \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 512 & 512 \end{pmatrix} \\ &= \begin{pmatrix} 512 & 512 \\ 512 & 512 \end{pmatrix} \end{aligned}$$

So $A^{10} = \begin{pmatrix} 512 & 512 \\ 512 & 512 \end{pmatrix}$

4.2 Application to Differential Equations

Suppose we have n functions of some variable t, $x_1(t), x_2(t), \ldots x_n(t)$. We may represent them with the vector

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$

$\mathbf{x}(t)$ is called a *vector function.*
We wish to solve a system of differential equations of the form

$$\begin{aligned} x_1'(t) &= a_{11}x_1(t) + a_{12}x_2(t) + \ldots + a_{1n}x_n(t) \\ x_2'(t) &= a_{21}x_1(t) + a_{22}x_2(t) + \ldots + a_{2n}x_n(t) \\ &\vdots \\ x_n'(t) &= a_{n1}x_1(t) + a_{n2}x_2(t) + \ldots + a_{nn}x_n(t) \end{aligned}$$

Where x_i' indicates the derivative of the function $x_i(t)$. We can represent this in matrix notation as $\mathbf{x}'(t) = A\mathbf{x}(t)$.
Given such a system and a set of initial conditions we would like to find the functions $x_i(t)$.
Recall for an arbitrary function $x(t)$,

$$x'(t) = \lambda x(t) \Rightarrow x(t) = Ce^{\lambda t},$$

where C is a constant of integration.

Example 11
Solve the system

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 9 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$$

Where $x(0) = 1$, $y(0) = 3$, $z(1) = 2e^9$.

So

$$x'(t) = 3x(t),\ y'(t) = 2y(t),\ \ z'(t) = 9z(t),$$

Thus

$$x(t) = C_1e^{3t},\ y(t) = C_2e^{2t},\ \ z(t) = C_3e^{9t},$$

Substituting in the initial conditions gives

$$C_1 = 1,\ C_2 = 3,\ C_3 = 2.$$

Now suppose that A is diagonalizable, so there exists invertable P and diagonal D such that $D = P^{-1}AP$. Suppose $\mathbf{x}(t)$ satisfies the linear system of differential equations

$$\mathbf{x}'(t) = A\mathbf{x} \ \ (*)$$

Define $\mathbf{y}(t) = P^{-1}\mathbf{x}(t)$.

$$\begin{array}{llcll} \text{Now} & \mathbf{x}'(t) & = & A\mathbf{x} & (*). \\ \text{So, premultiplying by } P^{-1}, & P^{-1}\mathbf{x}'(t) & = & P^{-1}APP^{-1}\mathbf{x}. & (PP^{-1} = I) \\ \Rightarrow & \mathbf{y}'(t) & = & D\mathbf{y}(t). & \end{array}$$

Thus to solve (*), we first solve

$$\mathbf{y}'(t) = D\mathbf{y}(t).$$

to find $\mathbf{y}(t)$. Then

$$\mathbf{x}'(t) = C\mathbf{y}'(t).$$

Example 12
Solve

$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

Where $x_1(0) = 0$, $x_2(0) = 1$.
We know $A = P^{-1}DP$, where

$$P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \text{ and } D = \text{diag}(0, 2) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

Consider

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}$$

So

$$y_1'(t) = 0 \Rightarrow y_1(t) = C_1$$
$$y_2'(t) = 2y(t) \Rightarrow y_2(t) = C_2e^{2t}$$

Now,

$$\begin{aligned} \mathbf{x}(t) &= P\mathbf{y}(t) \\ &= \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2e^{2t} \end{pmatrix} \\ &= \begin{pmatrix} C_1 + C_2e^{2t} \\ -C_1 + C_2e^{2t} \end{pmatrix} \end{aligned}$$

Now apply the initial conditions, $x_1(0) = 0$, so $C_1 + C_2 = 0$, and $x_2(0) = 1$ gives $-C_1 + C_2 = 1$ Solving gives $C_1 = -\frac{1}{2}$, $C_2 = \frac{1}{2}$.

$$\text{So } \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} e^{2t} + 1 \\ e^{2t} - 1 \end{pmatrix}$$