Ransom Everglades

Linear Transformations

The Search for a Good Basis

Linear Algebra Project

Author:
Period 6

Teacher:
Mr. Lefebvre

March 20, 2020

Contents

1 What is a Linear Transformation?

1.1 Defining a Linear Transformation

A transformation T follows the same idea as a function, as it assigns an output T(v) to each input vector v from the vector space V. These input and output vectors can be in any space (R^n, the function space, or the matrix space).

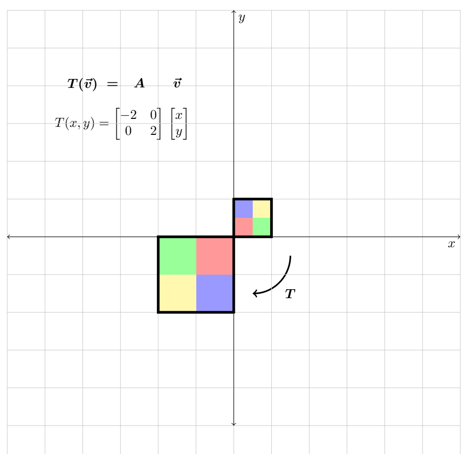
The transformation is linear if it meets this requirement for all $\vec{v}$ and $\vec{w}$:

$$T(c\vec{v} + d\vec{w}) = cT(\vec{v}) + dT(\vec{w})$$

The only exception is when one of the input vectors is 0. This is called the identity transformation: $T(\vec{v}) = \vec{v}$.

1.2 Linear Transformation Using a Matrix

A transformation takes every vector v and multiplies it by A, thus producing the output $T(\vec{v}) = A\vec{v}$. This operation transforms the whole space V.



1.3 Linear Transformation Properties

Rules of linear transformation mimic rules of a linear combination. Here are some facts about linear transformations.

1. Kernel of $T(\vec{v})$ is the set of all inputs in which $T(\vec{v}) = 0$, corresponds to null-space.
2. Range of $T(\vec{v})$ is the set of all outputs $T(\vec{v})$, corresponds to column space.
3. If A is a non-singular matrix, then $T(\vec{v})$ is invertible, thus $T^{-1}(T(\vec{v})) = 1$.

1.4 Input Basis and Output Basis

Let W represent a new basis with vectors $\{\vec{w_1}, \vec{w_2}, ...\vec{w_n}\}$. $T(\vec{v})$ can be written as some combination $b_1\vec{w_1} + b_2\vec{w_2} + ... + b_n\vec{w_n}$. V acts as the input basis and W acts as the output basis.

$$T(\vec{v}) = \begin{bmatrix} | & | & \cdots & | \\ \vec{w_1} & \vec{w_2} & \cdots & \vec{v_n} \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = b_1\vec{w_1} + b_2\vec{w_2} + ... + b_n\vec{w_n} \tag{1}$$

Example:

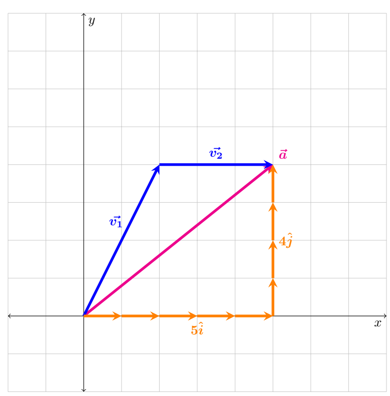
$$T(\begin{bmatrix} 8 \\ 7 \end{bmatrix}) = \underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}}_{W} \underbrace{\begin{bmatrix} 3 \\ 2 \end{bmatrix}}_{b} = 3 \underbrace{\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{\vec{w_1}} +2 \underbrace{\begin{bmatrix} 1 \\ 2 \end{bmatrix}}_{\vec{w_2}}$$

2 Coordinates with Respect to a Basis

2.1 What are Coordinates?

Coordinates are ways to describe vectors as a combination of basis vectors.

Example: In this image, the same vector $\vec{a}$ can be written in multiple ways with different coordinates.



2.1.1 The Standard Basis

A vector $\vec{a}$ written in the standard coordinate system is the same as saying it is a combination of basis vectors $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. The standard coordinate system is the one we have been using up until now and the notation for writing a vector a written in the standard coordinate system is $\vec{a}$.

Example: The vector $\vec{a}$ in the standard basis

$$\vec{a} = \begin{bmatrix} 5 \\ 4 \end{bmatrix} = 5\hat{\boldsymbol{i}} + 4\hat{\boldsymbol{i}} \tag{2}$$

2.1.2 Alternate Basis V

A vector $\vec{a}$ written in an alternate coordinate system is the same as saying it is a combination of basis vectors $V = \begin{bmatrix} | & | & \cdots & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_i} \\ | & | & \cdots & | \end{bmatrix}$. The notation for writing a vector in an alternate coordinate system V is $[\vec{a}]_v$.

Example: The vector $\vec{a}$ in an alternate basis V

$$[\vec{a}]_v = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1\boldsymbol{\vec{v_1}} + 1\boldsymbol{\vec{v_2}} \tag{3}$$

It is important to note that $\vec{v_1}$ and $\vec{v_2}$ can also be written in both coordinate systems.

$$\begin{bmatrix} \vec{v_1} & \vec{v_2} \end{bmatrix}_v = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} \vec{v_1} & \vec{v_2} \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 4 & 0 \end{bmatrix}$$

3 Change of Basis Matrix

3.1 General Equation

Let $I = \left\{\hat{i}, \hat{j}, \hat{k}, \cdots\right\}$. I is the standard basis with n dimensions.
Let $V = \{\vec{v_1}, \vec{v_2}, \cdots, \vec{v_k}\}$. V is a basis with k dimensions and k basis vectors $\vec{v_k}$.
The change of basis matrix B takes you from basis V to the standard basis.

$$\underbrace{[B]}_{nxk} \underbrace{\begin{bmatrix} | & | & \cdots & | \\ [\vec{v_1}]_v & [\vec{v_2}]_v & \cdots & [\vec{v_i}]_v \\ | & | & \cdots & | \end{bmatrix}}_{kxk} = \underbrace{\begin{bmatrix} | & | & \cdots & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_i} \\ | & | & \cdots & | \end{bmatrix}}_{nxk} \tag{4}$$

$$[B] \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_i} \\ | & | & \cdots & | \end{bmatrix} \tag{5}$$

$$B = \begin{bmatrix} | & | & \cdots & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_i} \\ | & | & \cdots & | \end{bmatrix} \tag{6}$$

The change of basis matrix B that takes you from coordinate system V to the standard coordinate system I is the matrix with basis vectors $\vec{v_i}$ as columns.

$$B[\vec{a}]_v = \vec{a} \tag{7}$$

3.2 Special Case: B is invertible

Basis $V = \{\vec{v_1}, \vec{v_2}, \cdots, \vec{v_k}\}$ has k basis vectors.
Standard basis $I = \left\{\hat{i}, \hat{j}, \hat{k}, \cdots\right\}$ has n basis vectors.

This means that coordinates $[\vec{a}]_v$ will have k dimensions and coordinates $\vec{a}$ will have n dimensions.

$$[\vec{a}]_v = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix} \text{ and } \vec{a} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \tag{8}$$

When $k = n$ (basis V and the standard basis I have the same dimensions), B is invertible and you can go from the standard basis I to basis V by multiplying by B^{-1}.

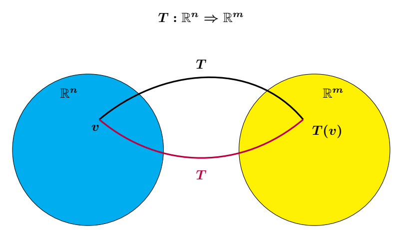
$$B[\vec{a}]_v = \vec{a} \text{ and } B^{-1}\vec{a} = [\vec{a}]_v \tag{9}$$

$$B = \begin{bmatrix} | & | & \cdots & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_i} \\ | & | & \cdots & | \end{bmatrix} \tag{10}$$

4 The Transformation Matrix

4.1 What is the Transformation Matrix?

The transformation matrix is the matrix $A_{v \to w}$ assigned to the linear transformation $T(v)$. This matrix takes the vector $\vec{v}$ from input basis V to output basis W. The input $\vec{v}$ is in $V = \mathbb{R}^n$ and the output $T(\vec{v})$ is in $W = \mathbb{R}^m$. The notation $A_{V \to w}$ denotes just that.



4.2 Transformation Matrix for the Standard Basis

The transformation matrix for the standard basis is the matrix $A_{I\to I}$ that takes an input vector $\vec{v}$ from the standard basis to output $T(\vec{v})$, also in the standard basis $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Often times this matrix is not diagonal and finding A^k is computationally expensive.

$$T(\vec{v}) = A_{I\to I}\vec{v}$$

.

Example: The transformation matrix $A_{I\to I}$ corresponding to the linear transformation $T(\vec{v}) = \langle 2v_1 + 3v_2, v_1 - 4v_2 \rangle$.

$$\begin{bmatrix} 2 & 3 \\ 1 & -4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 2v_1 + 3v_2 \\ 1v_1 - 4v_2 \end{bmatrix} \tag{11}$$

$$A_{I\to I} = \begin{bmatrix} 2 & 3 \\ 1 & -4 \end{bmatrix} \tag{12}$$

4.3 Transformation Matrix from Input Basis $V = \mathbb{R}^n$ to Output Basis $W = \mathbb{R}^m$

4.3.1 Why is this Important?

The reason finding a new transformation matrix besides $A_{I\to I}$ is useful because you can now find a transformation matrix $A_{v\to w}$ that is diagonal. This makes computations with A much more efficient.

$$T([\vec{v}]_v) = A_{v\to w}[\vec{v}]_w$$

$$\underbrace{A}_{mxn}\underbrace{[\vec{v}]_v}_{nx1} = \underbrace{[T(\vec{v})]_w}_{mx1} \tag{13}$$

Finding a general equation for a transformation matrix that takes you from input basis $V = \mathbb{R}^n$ to output basis $W = \mathbb{R}^m$ allows you to choose the best input and output bases.

Example: Say you need to multiply by your transformation matrix A thousands of times. Having a diagonal matrix A makes finding A^k much easier.

4.4 Finding Transformation Matrix $A_{v\to w}$

Goal: Find some matrix $A_{v\to w}$ such that $T([\vec{v}]_v) = A_{v\to w}[\vec{v}]_w$.

What we know: $T(\vec{v}) = A_{I\to I}\vec{v}$

Let $\vec{c}$ and $\vec{d}$ be vectors in the standard basis.

$$A_{I\to I}\vec{c} = \vec{d} = T(\vec{c})$$

Step 1: Multiply by change of basis matrix $B_{in} = V$ to get $[\vec{v}]_v$ from the input basis to the standard basis.

$$B_{in}[\vec{v}]_v = \vec{c}$$

Step 2: Multiply by change of basis matrix $B_{out} = W$ to get $[T(\vec{v})]_w$ from the output basis to the standard basis.

$$B_{out}[T(\vec{v})]_w = \vec{d}$$

Step 3: Plug $\vec{c}$ and $\vec{d}$ back into the original equation.

$$A_{I\to I}B_{in}[\vec{v}]_v = B_{out}[T(\vec{v})]_w$$

Step 4: Multiply both sides by $B_{out} = W$ and solve for $[T(\vec{v})]_w]$

$$B_{out}^{-1}A_{I\to I}B_{in}[\vec{v}]_v = [T(\vec{v})]_w$$

$$[T(\vec{v})]_w = \underbrace{B_{out}^{-1}A_{I\to I}B_{in}}_{A_{v\to w}(nxm)}[\vec{v}]_v = T([\vec{v}]_v)$$

Transformation matrix $A_{v\to w}$ that takes a vector $\vec{v}$ from input basis $V = B_{in} = \mathbb{R}^n$ to output basis $V = B_{in} = \mathbb{R}^n$ is:

$$A_{v\to w} = \underbrace{\begin{bmatrix} | & | & \cdots & | \\ \vec{w_1} & \vec{w_2} & \cdots & \vec{w_i} \\ | & | & \cdots & | \end{bmatrix}^{-1}}_{B_{out}(mxm)} \underbrace{A_{I\to I}}_{(mxn} \underbrace{\begin{bmatrix} | & | & \cdots & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_i} \\ | & | & \cdots & | \end{bmatrix}}_{B_{in}(nxn)} \tag{14}$$

5 Choosing the Best Basis

We now know that the same linear transformation $T(\vec{v})$ can have different transformation matrices associated with it. The goal now is to find the best input and output bases such that the matrix associated with $T(\vec{v})$ is easy to compute with.

5.1 Input Basis V and Output Basis U

We have already seen in a previous section that when the transformation matrix $A_{v\to w}$ that takes a vector $\vec{v}$ from input space V to output space U is the matrix. To make things easier, we will choose input basis vectors $V = \{\vec{v_1}, \vec{v_2}, \cdots, \vec{v_n}\}$ and output basis vector $U = \{\vec{u_1}, \vec{u_2}, \cdots, \vec{u_m}\}$ as orthonormal.

$$A_{v\to w} = \begin{bmatrix} | & | & \cdots & | \\ \vec{u_1} & \vec{u_2} & \cdots & \vec{u_m} \\ | & | & \cdots & | \end{bmatrix}^{-1} A_{I\to I} \begin{bmatrix} | & | & \cdots & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_n} \\ | & | & \cdots & | \end{bmatrix} = U^{-1}A_{I\to I}V \tag{15}$$

If you recall from the singular value decomposition section, $U^{-1}AV = \Sigma$ where Σ is the singular value matrix with singular values of $A_{I\to I}$: $(\sigma_1, \sigma_2, \cdots, \sigma_r)$ along the diagonal.

$$\Sigma = U^{-1}A_{I\to I}V = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & \sigma_n \end{bmatrix} \tag{16}$$

This matrix $\Sigma = U^{-1}A_{I\to I}V = U^TA_{I\to I}V$ is easy to compute and makes using the linear transformation $T(\vec{v})$ very easy because Σ is diagonal. We have already seen in other chapters that finding Σ^k is easier than finding A^k.

5.2 Eigenvectors as a Basis

In the case that $A_{I\to I}$ is a square matrix and has n independent eigenvectors, the eigenspace $X = \{\vec{x_1}, \vec{x_2}, \cdots, \vec{x_n}\}$ makes for a great basis. In this case both transformation matrices $A_{x\to x}$ and $A_{I\to I}$ take a vector $\vec{v}$ from $\mathbb{R}^n \to \mathbb{R}^n$.

In this scenario, the input basis matrix is X and the output basis matrix is also X.

$$A_{v\to w} = \begin{bmatrix} | & | & \cdots & | \\ \vec{x_1} & \vec{x_2} & \cdots & \vec{x_n} \\ | & | & \cdots & | \end{bmatrix}^{-1} A_{I\to I} \begin{bmatrix} | & | & \cdots & | \\ \vec{x_1} & \vec{x_2} & \cdots & \vec{x_n} \\ | & | & \cdots & | \end{bmatrix} = X^{-1}AX \tag{17}$$

If you recall from the eigenvector section, $X^{-1}A_{I\to I}X = \Lambda$ where Λ is the eigenvector matrix with eigenvalues of $A_{I\to i}$: $(\lambda_1, \lambda_2, \cdots, \lambda_n)$ along the diagonal.

$$\Lambda = X^{-1}A_{I\to I}X = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{bmatrix} \tag{18}$$

This matrix $\Lambda = X^{-1}AX$ makes computing the linear transformation $T(\vec{v})$ easier because Σ is diagonal. We have already seen in other chapters that finding Σ^k is easier than finding A^k.

5.3 Generalized Eigenvectors as a Basis

In the case that $A_{I\to I}$ is a square matrix and has $s < n$ independent eigenvectors, then the eigenspace $X = \{\vec{x_1}, \vec{x_2}, \cdots, \vec{x_n}\}$ no longer makes for a great basis because X is not invertible. A is no longer diagonizable so the goal becomes to make the matrix associated with the transformation as diagonal as possible. This matrix that is the "most diagonal" matrix associated with the linear transformation is known as the *Jordan Form* J. When $s = n$, $J = \Lambda$.

In this scenario, the input basis matrix is the generalized eigenvectors and the output basis matrix is also the generalized eigenvector.

$$B_{in} = B_{out} = \text{generalized eigenvectors}$$
$$B_{in}^{-1}A_{I\to I}B_{out} = \text{Jordan Form } J$$

In order to better understand the Jordan Form and where it comes from, we will look at similar matrices.

5.3.1 What are Similar Matrices?

A matrix B is defined as being similar to another matrix A when it can be written in the form of $B = M^{-1}AM$ for some matrix M. All matrices in the same family share the same ***eigenvalues*** and number of independent ***eigenvalues***.

We have already seen this before. When A is a square matrix and has n independent eigenvalues and n eigenvectors $X = \{\vec{x_1}, \vec{x_2}, \cdots, \vec{x_n}\}$, the diagonal matrix Λ can be written as $\Lambda = X^{-1}AX$ and by definition is similar to A.

Example:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \Lambda = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} B = \begin{bmatrix} -2 & -15 \\ 1 & 6 \end{bmatrix} \tag{19}$$

All three of these matrices are similar and share the same eigenvalues. Notices, the diagonal matrix Λ is the best matrix in the family of similar matrices because it is the easiest to compute with.

5.3.2 The Jordan Form

What happens when A does not have n independent eigenvectors? In this case, you cannot construct a similar matrix Λ that is diagonal. Instead, you can construct a matrix known as the **Jordan Form *J*** that is the most diagonal matrix in its family of similar matrices.

Example: Let A be a $2x2$ matrix with $\lambda_1 = \lambda_2 = 4$, then there are two families of matrices that meet this conditon.

Family 1

$$\text{Let } A = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} = 4 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 4I$$

Let B be the family of matrices similar to A, then

$$B = M^{-1}(4I)M = 4M^{-1}IM = 4M^{-1}M = 4I = A$$

$$B = \begin{bmatrix} 2 & -1 \\ 3 & 6 \end{bmatrix}, \begin{bmatrix} 3 & 1 \\ -1 & 5 \end{bmatrix}, \begin{bmatrix} 7 & 1 \\ 5 & 1 \end{bmatrix}, \text{etc.}$$

The family of similar matrices to this A contains every other matrix with trace = 8, determinant = 16, and one independent eigenvectors.

Family 2

$$\text{Let } A = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$$

Let B be the family of matrices similar to A, then

$$B = M^{-1}AM$$

The family of similar matrices to this A contains every other matrix with trace = 8, determinant = 16 and one independent eigenvectors. Notice no matrix in this family is diagonizable and instead, A is the "most diagonal" and thus the best matrix in this family. This matrix is known as the Jordan Form J.

5.3.3 How to Construct the Jordan Form

If matrix A (n x n) has $s < n$ independent eigenvectors, it is similar to the Jordan Form matrix J that has s Jordan Blocks $\{J_1, J_2, ..., J_s\}$ on its diagonal. Some matrix B puts A into the Jordan Form. Each Jordan Block corresponds to an independent eigenvector.

$$\textbf{Jordan Form} = B^{-1}AB = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{bmatrix} \tag{20}$$

$$\textbf{Jordan Block} = J_1 = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} \tag{21}$$

Matrices are similar if they share the same Jordan Form J.

6 The Fourier and Circulant Matrices

6.1 The Fourier Matrix

We choose eigenvectors to be the basis because naturally, eigenvectors do not change when a transformation is applied. Thus, if B is a basis matrix whose columns are eigenvectors, $B_{in} = B_{out}$.

This matrix then becomes what is called the Fourier Matrix F. Fx is then called a Discrete Fourier Transform of x. Because F is made up of eigenvectors, we can then ask: Which matrices will be diagonalized by F? What matrices have the columns of F as eigenvectors? Here is an important example:

We want to find all matrices that have eigenvectors.

$$x = \begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \\ \lambda^3 \end{bmatrix}$$

Each column of F will have this form.

Let's look at a permutation matrix P, where λ is an eigenvalue of P and x is an eigenvector of P.

$$Px = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \\ \lambda^3 \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \\ \lambda^3 \end{bmatrix} = \lambda x$$

Here, the 4^{th} row gives us the equation for $\lambda^4 = 1$, which then gives us 4 distinct eigenvalues, which then gives us 4 distinct eigenvectors. The eignvectors (columns of F) are orthogonal because they are the eigenvectors of an orthogonal matrix.

$$\Lambda = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -i \end{bmatrix} \quad F = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & i^2 & 1 & (-1)^2 \\ 1 & i^3 & -1 & (-i)^3 \end{bmatrix}$$

This gives us a diagonal eigenvalue matrix Λ. This also gives us F, an orthogonal eigenvector matrix. Notice how every column of F is in the form $\langle 1, \lambda^2, \lambda^3, \lambda^4 \rangle$ because that is what we were looking for. Because P is a permutation matrix, we can come to the conclusion that F will diagonalize any power of P.

6.2 The Circulant Matrix

We know that P^2, P^3, P^4 all have the same eigenvectors of P. Furthermore, we know the eigenvalues of P^k are just λ^k. Let's look at P^2:

$$P^2x = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \\ \lambda^3 \end{bmatrix} = \lambda^2 \begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \\ \lambda^3 \end{bmatrix} = \lambda^2 x$$

Now, the third row gives $\lambda^4 = 1$. P^4 is very special because $P^4 = I$. When we do the "cyclic permutation" 4 times, P^4x is the same vector x we started with.

[**Important**] If $P, P^2, P^3, \text{and} P^4 = I$ all have the same eigenvector matrix F, so does any combination $C = c_1P + c_2P^2 + c_3P^3 + c_4P^4 + c_0I$. C represents the **circulant matrix**. This represents every matrix whose eigenvectors are the columns (Fourier vectors) in F.

$$C = \begin{bmatrix} c_0 & c_1 & c_2 & c_3 \\ c_3 & c_0 & c_1 & c_2 \\ c_2 & c_3 & c_0 & c_1 \\ c_1 & c_2 & c_3 & c_0 \end{bmatrix}$$

It is important to note:

- c_0 is in every spot where there is a 1 in $P^4 = I$.

- c_1 is in every spot where there is a 1 in P.
- c_2 is in every sport where there is a 1 in P^2,
- c_3 is in every spot where there is a 1 in P^3.

This also gives us a good formula for the eigenvalues of C. The 4 eigenvalues of C are given by the Fourier transform Fc.

$$Fc = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & i^2 & 1 & (-1)^2 \\ 1 & i^3 & -1 & (-i)^3 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} c_0 + c_1 + c_2 + c_3 \\ c_0 + ic_1 - c_2 - c_3 \\ c_0 - c_1 + c_2 - c_3 \\ c_0 - ic_1 - c_2 + ic_3 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix}$$

Note that circulant matrices have the same number c0 going down the main diagonal. The rest of the values c1, c2, and c3 create their own constant diagonals by circling through the rest of the positions. This explains the name "circulant" and it indicates that these matrices are periodic or cyclic. Even the powers of λ cycle around!

7 Basis for a Function Space

7.1 What is a Function Space?

Similar to a vector space being the space occupied by the set of vectors, the function space is the space occupied by a set of functions. A function space is denoted as $f(x)$. While a vector space R^n is finite-dimensional because it has n dimensions, **a function space is infinite-dimensional**. It takes infinitely many basis functions to perfectly recreate a function $f(x)$. Think of Taylor expansions to better understand this concept.

Example: A function space $f(x) = e^x$
is infinite-dimensional The Taylor expansion for this function space $f(x)$ is an infinite sum of polynomials.

$$f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

7.2 Parallels Between Function Space and Vector Space

7.2.1 Basis for Vector Space vs Basis for Function Space

A basis for a function space $f(x)$ should also follow the same conditions as a basis for a vector space: any combination of the basis functions should make up the entire function space. Additionally, just like with Vector spaces, Function spaces can be defined by different bases.

Example 1: Basis for Vector Space
A basis for the space $\mathbb{R}^2$ could the standard basis $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
A basis for the space $\mathbb{R}^2$ could also be $V = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

Example 2: Basis for Function Space
A basis for $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots, a_n x^n$ could be $\{1, x, x^2, \cdots, x^n\}$.
A basis for $f(x) = c_0 + c_1 \sin(x) + c_2 \cos(x) + \cdots,$ could be $\{1, \sin(x), \cos(x)\}$.

7.2.2 Inner Product for Vector Space vs Basis for Function Space

Similar to with Vectors, you can also take the inner product of two functions.

Example 1: Inner Product for Vectors
The inner product of the vectors x and y is $x^T y = x_1 y_1 + x_2 y_2 + ... + x_n y_n$

Example 2: Inner Product Functions
Inner product = $(\boldsymbol{f}, \boldsymbol{g}) = \int f(x) g(x) dx$
Complex inner product = $(\boldsymbol{f}, \boldsymbol{g}) = \int \bar{f(x)} g(x) dx, \bar{f}$ = complex conjugate
Weighted inner product = $(\boldsymbol{f}, \boldsymbol{g})_w = \int w(x) \bar{f(x)} g(x) dx, w(x)$ = weight function

7.2.3 Testing for a Good Basis with Vectors vs Testing for a Good Basis with Functions

Similar to with sectors, you can also test if a basis is a good basis for a function space. In both cases, the best basis is when the inner product of a basis and its transpose is equal to the identity.

Example 1: Testing for a Good Basis with Vectors
We have already seen in previous sections how to find a good basis. Generally the best basis B is when it is orthonormal or as close to orthonormal as possible. The way to test if B is orthonormal is to check $B^T B$. If $B^T B = I$, then basis B is orthonormal and thus a great basis. Otherwise, you want to minimize the $\frac{\text{Largest } \lambda \text{ of } BTB}{\text{Smallest } \lambda \text{ of } B^T B}$

Example 2: Testing for a Good Basis with Functions
Similarly, the test for a good basis with functions is to check the inner product. The best basis is orthonormal which occurs when $f(x)^T g(x) = \int f(x) g(x) dx = 0$

7.3 The Search for a Good Basis

If we look at the basis $\{1, x, x^2, \cdots, x^n\}$ for function space $f(x)$, we can now check if it is a good basis or not. Unfortunately this is a terrible basis because each function x^n is barely independent. You can almost make x^n with some combination of $1, x, x^2, \cdots, x^{n-1}$. The basis also produces inner product $B^T B*$ with a very large ratio between its largest and smallest eigenvalue. This inner product matrix $B^T B$ is known as the $textbf{HilbertMatrix}$.

When the limits of integration are from $x = 0$ to $x = 1$, the inner product of x^i and x^j becomes:

$$\int_0^1 x^i x^j dx = \frac{x^{i+j+1}}{i+j+1}\Big|_{x=0}^{x=1} = \frac{1}{i+j+1} = \text{entries of Hilbert Matrix } B^T B$$

Changing to the symmetric interval from $x = -1$ to $x = 1$ produces orthogonality between all even and all odd functions. This change makes half of the basis vector orthogonal to the other half.

$$\int_{-1}^1 x^3 x^4 dx = \frac{x^8}{8}\Big|_{x=-1}^{x=1} = 0 \Rightarrow \int_{-1}^1 \mathbf{even}(x)\mathbf{odd(x)}dx = 0$$

This is a good start in the search for an orthogonal basis.

7.4 Orthogonal Bases for Function Space

1. The Fourier basis

   - $1, \sin(x), \cos(x), \sin(2x), \cos(2x), \cdots$
   - Fourier basis functions are all periodic, and repeat over every 2π interval.
   - This basis is orthogonal as every sine and cosine is orthogonal to every other sine and cosine.
   - This basis is also a great basis because sine and cosine are great for approximations (Taylor series expansion error bounds).

2. The Legendre basis

   - $1, x, x^2 - \frac{1}{3}, x^3 - \frac{3}{5}x, \cdots$
   - The basis is achieved by starting with $\{1, x, x^2, \cdots, x^n\}$ and applying Grahm-Schmidt to get orthogonal an orthogonal basis.
   - Even powers of x: $\frac{(x^2,1)}{(1,1)} = \frac{(x^2)^T(1)}{(1)^T(1)} = \frac{\int x^2 dx}{\int 1 dx} = \frac{2/3}{2} = \frac{1}{3}$ gives $x^2 - \frac{1}{3}$
   - Odd powers of x: $\frac{(x^3,x)}{(x,x)} = \frac{(x^3)^T(x)}{(x)^T(x)} = \frac{\int x^4 dx}{\int x^2 dx} = \frac{2/5}{3/5} = \frac{3}{5}$ gives $x^3 - \frac{3}{5}x$
   - Continuing the process of Graham-Schmidt for every $x^4, x^5, \cdots, x^n$ produces the rest of the Legendre basis.

3. The Chebyshev basis

   - $1, x, 2x^2 - 1, 4x^3 - 3x, \cdots$

- The Chebyshev basis comes from the basis $1, \cos(\theta), \cos(2\theta), \cos(3\theta), \cdots$ which is orthogonal
- It gives a huge computational advantage because the Fast Fourier Transform can now be used.
- This basis found by setting x = $\cos(\theta)$
- The n^{th} degree Chebyshev polynomial $\boldsymbol{T_n(x)}$ converts to Fourier's $\boldsymbol{\cos(n\theta) = T_n cos(\theta)}$
- $T_0(\cos(x)) = \cos(0) = 1$
- $T_1(\cos(x)) = \cos(\theta) = \cos(\theta)$
- $T_2(\cos(x)) = 2(\cos(\theta))^2 - 1 = \cos(2\theta)$
- $T_3(\cos(x)) = 4(\cos(\theta))^3 - 3(\cos(\theta)) = \cos(3\theta)$

References

[1] `http://web.mit.edu/18.06/www/Spring09/legendre.pdf`
S. G. Johnson, MIT course 18.06, Spring 2009

[2] *Khan Academy*. Accessed March 19, 2020.
`https://www.khanacademy.org/math/linear-algebra/alternate-bases`

[3] Strang , G. (2016). *Introduction to Linear Algebra (5th ed.)*. Wellesley - Cambridge Press.