1 Examples

Example 1.1

Let

$$A = \begin{bmatrix} 2 & -2 & 2 \\ -2 & -1 & 4 \\ 2 & 4 & -1 \end{bmatrix}$$

Find the eigenvalues and eigenvectors of A.

To begin with, consider

$$\begin{aligned} \det \begin{bmatrix} 2-\lambda & 2 & -2 \\ 2 & -1-\lambda & -4 \\ -2 & -4 & -1-\lambda \end{bmatrix} &= \begin{vmatrix} 2-\lambda & 2 & -2 \\ 2 & -1-\lambda & -4 \\ 0 & -5-\lambda & -5-\lambda \end{vmatrix} \\ &= (2-\lambda)\begin{vmatrix} -1-\lambda & -4 \\ -5-\lambda & -5-\lambda \end{vmatrix} - 2\begin{vmatrix} 2 & -2 \\ -5-\lambda & -1-\lambda \end{vmatrix} \\ &= (2-\lambda)(-(3-\lambda)(5+\lambda)) - 2(-2-2\lambda-10-2\lambda) \\ &= -(\lambda+6)(\lambda-3)^2 \end{aligned}$$

Now let's find eigenvectors for $\lambda = 3$. In this case we are solving

$$(A-3I)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

Note

$$A-3I = \begin{bmatrix} -1 & -2 & 2 \\ -2 & -4 & 4 \\ 2 & 4 & -4 \end{bmatrix}$$

A basic row reduction gives

$$\begin{bmatrix} -1 & -2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and thus the kernel has dimension 2, and a basis is given by

$$v_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

Thus, $\text{span}(v_1, v_2)\backslash 0$ gives the collection of eigenvectors for $\lambda = 3$. We will leave as an exercise to find eigenvectors for $\lambda = -6$.

Example 1.2

Let $V = K[x]_n$ be the space of polynomials in x of max degree n. Find the characteristics polynomial for derivative.

A basis of V is $1, x, ..., x^n$, and in this case $\frac{d}{dx}$ has matrix representation

$$\begin{bmatrix} 0 & 1 & 0 & .... & 0 \\ 0 & 0 & 2 & ... & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & ... & n-1 \\ 0 & 0 & 0 & ... & 0 \end{bmatrix}$$

Thus

$$\det(D - \lambda I) = \det \begin{bmatrix} -\lambda & 1 & 0 & .... & 0 \\ 0 & -\lambda & 2 & ... & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & ... & n-1 \\ 0 & 0 & 0 & ... & -\lambda \end{bmatrix} = (-\lambda)^n$$

Example 1.3

Let $f(x) = x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0$ be irreducible $\mathbb{Q}$ polynomial, $n > 0$. Let ω_0 be a root of f in the complex number. Now view $\mathbb{C}$ as $\mathbb{Q}$-vector space. Define

$$\begin{aligned} F : \mathbb{Q}[x] &\to \mathbb{C} \\ g(x) &\mapsto g(\omega_0) \end{aligned}$$

1. Is F a $\mathbb{Q}$-linear map? If yes, find a basis of $\operatorname{Im} F$
2. Let $A(z) = \omega_0 z$ for all $z \in \operatorname{Im} F$. Is $A : \operatorname{Im} F \to \operatorname{Im} F$ a linear map? If yes, what is a matrix representation of A in your basis in (1)?
3. Is A diagonalizable over $\mathbb{Q}$?
4. Now view A as matrix over $\mathbb{C}$, is it diagonalizable?

(1): Take $g_1, g_2 \in \mathbb{Q}[x]$, then its not hard to see

$$F(g_1 + cg_2) = (g_1 + cg_2)(\omega_0) = g_1(\omega_0) + cg_2(\omega_0)$$

Thus F is $\mathbb{Q}$-linear.

Next we find a basis of $\operatorname{Im} F$. Pick $u \in \operatorname{Im} F$, then there exists $g \in \mathbb{Q}[x]$ so $g(\omega_0) = u$. Now perform long division with $g(x), f(x)$ we get

$$g(x) = h(x)f(x) + r(x)$$

where $\deg(r) < \deg(f) = n$. Suppose $r(x) = c_0 + c_1 x + ... + c_{n-1}x^{n-1}$, then we see we get

$$g(\omega_0) = 0 = h(\omega_0)f(\omega_0) + r(\omega_0)$$

But since ω_0 is a root of f in complex, we get $f(\omega_0) = 0$ and thus

$$u = c_0 + c_1\omega_0 + ... + c_{n-1}\omega_0^{n-1}$$

In other word, $\{1, \omega_0, ..., \omega_0^{n-1}\}$ spans $\operatorname{Im} F$. It remains to show linearly independent. To that end, suppose for a contradiction we can find $k_0, ..., k_{n-1}$ so

$$k_0 + k_1\omega_0 + ... + k_{n-1}\omega_0^{n-1} = 0$$

Then we get a polynomial $q(x) = \sum_{i=0}^{n-1} k_i x^i$ with $q(\omega_0) = 0 = f(\omega_0)$. Thus over the complex numbers q and f shares a common factor $(x - \omega_0)$. In other word, f and q are not coprime over $\mathbb{C}$, and hence they cannot be coprime over $\mathbb{Q}$. Hence we either have $f \mid q$ or $q \mid f$, but f is irreducible, so we must have $f \mid q$. This is a contradiction by degree consideration. This shows $\{1, ..., \omega_0^{n-1}\}$ is linearly independent.

(2): Observe $A(1) = \omega_0$, $A(\omega_0) = \omega_0^2$,...,$A(\omega_0^{n-2}) = \omega_0^{n-1}$ and

$$A(\omega^{n-1}) = \omega_0^n = -a_0 - a_1\omega_0 - ... - a_{n-1}\omega_0^{n-1}$$

Thus $A(z) \in \operatorname{Im}(F)$ if $z \in \operatorname{Im}(F)$. This verifies $A : \operatorname{Im} F \to \operatorname{Im} F$. We leave as an exercise to show it is linear.

The matrix representation of A is clearly given by

$$\begin{bmatrix} 0 & 0 & 0 & ... & 0 & -a_0 \\ 1 & 0 & 0 & ... & 0 & -a_1 \\ 0 & 1 & 0 & ... & 0 & -a_2 \\ 0 & 0 & 1 & ... & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & ... & 0 & -a_{n-1} \end{bmatrix}$$

(3): To see if this is diagonalizable, we need to find its characteristics polynomial. We claim $\det(A - \lambda I) = (-1)^n f(\lambda)$. To prove this, we proceed by induction.

For $n = 2$ we get

$$\begin{vmatrix} -\lambda & -a_0 \\ 1 & -a_1 - \lambda \end{vmatrix} = \lambda^2 + a_1\lambda + a_0 = f(\lambda)$$

Suppose induction holds, then we consider n by n determinant

$$\begin{bmatrix} -\lambda & 0 & 0 & ... & 0 & -a_0 \\ 1 & -\lambda & 0 & ... & 0 & -a_1 \\ 0 & 1 & -\lambda & ... & 0 & -a_2 \\ 0 & 0 & 1 & ... & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & ... & 0 & -\lambda - a_{n-1} \end{bmatrix}$$

Now expand this by the first row, we get

$$-\lambda \cdot \begin{vmatrix} -\lambda & 0 & ... & 0 & -a_1 \\ 1 & -\lambda & ... & 0 & -a_2 \\ 0 & 1 & ... & 0 & -a_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & ... & 0 & -\lambda - a_{n-1} \end{vmatrix} + (-1)^{n+1}(-a_0) \begin{vmatrix} 1 & -\lambda & ... & 0 \\ 0 & 1 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & 1 \end{vmatrix}$$

But then the left term by induction hypothesis is

$$-\lambda \cdot (-1)^{n-1}(\lambda^{n-1} + a_{n-1}\lambda^{n-2} + \ldots + a_2\lambda + a_1) + (-1)^{n+2}a_0$$

This is just

$$(-1)^n f(\lambda)$$

as desired.

However, by assumption f is irreducible over $\mathbb{Q}$, thus $\det(A - \lambda I) = (-1)^n f(\lambda)$ has no root over $\mathbb{Q}$, thus A has no eigenvalues over $\mathbb{Q}$, thus cannot be diagonalizable over $\mathbb{Q}$.

(4): Since f is irreducible over $\mathbb{Q}$, f has no repeated roots over $\mathbb{C}$. Thus all eigenvalues of f are distinct, and hence it must be diagonalizable.

To see this, consider

Lemma 1.4

Let λ_1, λ_2 be distinct eigenvalues of A. Let v_1 be eigenvector of λ_1 and v_2 eigenvector of λ_2, then v_1, v_2 are linearly independent.

Proof. Indeed, note $Av_i = \lambda_i v_i$. Thus suppose $av_1 + bv_2 = 0$, then

$$A(av_1 + bv_2) = 0 = a\lambda_1 v_1 + b\lambda_2 v_2$$

Since λ_i are distinct, WLOG we can assume $\lambda_1 \neq 0$. Then we get

$$\lambda_1(av_1 + bv_2) = 0 = a\lambda_1 v_1 + b\lambda_1 v_2$$

Now subtract the two equations we get

$$0 = a\lambda_1 v_1 + b\lambda_1 v_2 - (a\lambda_1 v_1 + b\lambda_2 v_2) = b(\lambda_1 - \lambda_2)v_2$$

But v_2 is not the zero vector, $\lambda_1 - \lambda_2 \neq 0$, and thus $b = 0$. Now repeat this argument for a, we conclude $a = b = 0$. Thus they are linearly independent.

Corollary 1.4.1

Let A be n by n matrix with n distinct eigenvalues. Then A is diagonalizable.

Proof. Each λ_i has at least one non-zero eigenvector. Say $v_1, .., v_n$, but then they are linearly independent by the above result. Since dim of V is n, this means $v_1, ..., v_n$ is a basis and thus A is diagonalizable.

2 Enrichment: Linear Code

Suppose you and your friend wants to order takeout.

You decided to go for 15 hot wings, and when you order it, what happens in secret is that your phone translate this into binary code, which is 01111 (amount other things, of course), and send it to the restaurant.

But what happens in real life is that information get lost all the time (a simpler example would be radio signals, which can be affected by other sources of electromagnetic waves), and hence it is possible that somehow your binary code get messed up and became 11111. Now you get 31 hot wings on your plate.

This is very bad, because you are paying double the amount.

Hence, it is only natural to want to design a protocol of transmitting data so that not only we are getting the message, we want to be able to check if this message has been altered (for natural reason) or not.

For the purpose of exposition, let's assume the message we are trying to send is only 4 bits. In other word, we are looking at 4-tuple (a_1, a_2, a_3, a_4) with $a_i \in \{0, 1\}$. In other word, this 4-tuple lies in $\mathbb{Z}_2^4$.

In real life, if your friend texted you the string

amachronistic

It will not be too hard for you to figure out what they meant is

anachronistic

(p.s. it means chronologically misplaced).

Thus, longer the word, easier the task to deduce the correct message, assuming the error rate is small.

Hence, it is natural to devise the protocol so that we contain extra information. Thus, for a message (a_1, a_2, a_3, a_4), let's append three more elements c_1, c_2, c_3, so that it becomes a vector in $\mathbb{Z}_2^7$, where

$$\begin{cases} c_1 = a_1 + a_2 + a_3 \\ c_2 = a_2 + a_2 + a_4 \\ c_3 = a_1 + a_3 + a_4 \end{cases}$$

This induces linear map

$$\sigma : \mathbb{Z}_2^4 \to \mathbb{Z}_2^7$$

which is clearly injective.

This map σ is called a ***encoding***, and $C := \operatorname{Im} \sigma$ is called a ***code***, and an element of $C := \operatorname{Im} \sigma$ is called a ***codeword***.

The first four elements in $v \in \operatorname{Im}\sigma$ contains the actual information, and the last three elements are just for checking correctness.

Let's study σ. Clearly the defining equation for c_1, c_2, c_3 is given by

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}$$

Denote this matrix by A, then we see $v \in \operatorname{Im}\sigma$ if and only if

$$v \in \ker H$$

where

$$H = \begin{bmatrix} 1 & 1 & 1 & 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & -1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & -1 \end{bmatrix}$$

This shows the code C is a vector subspace of $\mathbb{Z}_2^7$.

Now, suppose we received a text $v \in \mathbb{Z}_2^7$, and we want to know if it is correct or not, then we just need to compute Hv. If $Hv \neq 0$, then we know for sure that v must being incorrect.

In this case, we know something is wrong, but is it possible for us to correct this code so that we recover the original message?

Since we always assume the channel we are sending message is always reasonable, in the sense that we allow errors, but the chance of getting a bunch of bits altered is very low (e.g. so 1111 becomes 0000 is very low).

Thus the natural way to correct this code is to find a codeword $w \in C$, so that the number of bits differ between v and w is minimal.

Definition 2.1

Let $v, w \in \mathbb{Z}_2^n$, define ***Hamming distance*** $h(v, w)$ as the number of 1's in $v - w$, i.e. it is the number of places where v_i is not equal w_i.

Thus, if we get message v, and $Hv \neq 0$, then what we do is to find $w \in C$ so that $d(v, w)$ is minimal.

At the current stage, the computation is fine, but if we want to transmit larger data, we are not fine. Hence we do a little bit more analysis.

Assume the original message is a, the received message is c and $b \in C$ minimalise $d(b, c)$. Then define

$$e = c - a$$

as the error vector and we see

$$He = H(c - a) = Hc - Ha = Hc$$

For a vector $v \in \mathbb{Z}_2^7$, we call Hv the check vector. Then we see the error vector e and c has the same check vector.

More generally, a, b has the same check vector iff $Ha = Hb$ iff $H(a-b) = 0$ iff $a - b \in C$ iff $a + C = b + C$, i.e. $a = b$ in $\mathbb{Z}_2^7/C$.

In other word, a and b has the same check vector if and only if they are the same element in the quotient vector space $\mathbb{Z}_2^7/C$. In particular, $e + C = c + C$, and thus the most likely correct code must be the smallest (smallest value of $d(x, 0)$) vector in the equivalence class $c + C$.

Thus, we shall translate received code c as $c - e$ where e is the smallest vector in the equivalence class $c + C$.

Let's see an example in action.

Suppose we are encoding $\mathbb{Z}_2^2$, with check matrix

$$H = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

Then $\dim C = 2$, which contains four vectors $(0,0,0,0)$, $(1,0,1,0)$, $(0,1,1,1)$ and $(1,1,0,1)$. Thus the quotient space $\mathbb{Z}_2^4/C$ is also of dimension 2, which says we have 2^2 equivalence classes. They are given by

0000	1010	0111	1101
1000	0010	1111	0101
0100	1110	0011	1001
0001	1011	0110	1100

Each row is an equivalence class, and now we just find the check matrix for the smallest vector in each of the equivalence class. This gives

0000	1010	0111	1101	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$
1000	0010	1111	0101	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$
0100	1110	0011	1001	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$
0001	1011	0110	1100	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Thus, now if we received the message $c = 1110$, then we compute $Hc = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. This tells us 1110 lies in the row with check vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Once we find 1110 in the table, we just take the vector in the toppest row in the same column as 1110 as our output "corrected" vector.

3 Quiz Questions

Example 3.1

Compute the following n by n determinant:

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ a_1 & a_2 & a_3 & \dots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \dots & a_n^2 \\ a_1^3 & a_2^3 & a_3^3 & \dots & a_n^3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \dots & a_n^{n-1} \end{vmatrix}$$

The options are:

1. $$\prod_{i=1}^{n-1}(a_{i+1}-a_i)$$
2. $$\sum_{1\le i<j\le n}(a_j-a_i)$$
3. $$\prod_{1\le i<j\le n}(a_j-a_i)$$
4. $$\sum_{\sigma\in S_n}\prod_{i=1}^{n}(a_i-a_{\sigma(i)})$$

Solution. Use induction to show

$$\det V_n = \prod_{1\le j<i\le n}(a_i-a_j)$$

Clearly it holds for $n=2$. Thus we assume it holds for $n-1$. Then

$$\det V_n = \begin{vmatrix} 1 & 1 & \dots & 1 \\ 0 & a_2-a_1 & \dots & a_n-a_1 \\ 0 & a_2^2-a_1a_2 & \dots & a_n^2-a_1a_n \\ 0 & \vdots & \ddots & \vdots \\ 0 & a_2^{n-1}-a_2^{n-2}a_1 & \dots & a_n^{n-1}-a_{n-1}^{n-2}a_1 \end{vmatrix}$$

where we add $-a_1$ times $(n-1)$th row to the nth row, then add -1 times $(n-2)$th row to the $n-1$th row and so on. Now expand the first column we get

$$\begin{aligned} \det V_n &= \begin{vmatrix} a_2-a_1 & \dots & a_n-a_1 \\ \vdots & \ddots & \vdots \\ a_2^{n-1}-a_1a_2^{n-2} & \dots & a_n^{n-1}-a_{n-1}^{n-2}a_1 \end{vmatrix} \\ &= \begin{vmatrix} (a_2-a_1)1 & \dots & (a_n-a_1)1 \\ (a_2-a_1)a_2 & \dots & (a_n-a_1)a_n \\ \vdots & \ddots & \vdots \\ (a_2-a_1)a_2^{n-2} & \dots & (a_n-a_1)a_n^{n-2} \end{vmatrix} \end{aligned}$$

Now pull out the $a_2 - a_1$ on the first column, $a_3 - a_1$ on the second and so on, we get our induction hold.

Example 3.2

Compute the following n by n determinant where $a \neq b$ and $a, b \neq 0$:

$$\begin{vmatrix} a+b & ab & 0 & 0 & \dots & 0 & 0 \\ 1 & a+b & ab & 0 & \dots & 0 & 0 \\ 0 & 1 & a+b & ab & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & a+b \end{vmatrix}$$

The options are

1. $$\frac{a^{n+1} - b^{n+1}}{a - b}$$
2. $$a^n + (n-1)ab + b^n$$
3. $$(a+b)^n - (n-1)ab$$
4. $$\frac{a^n + b^n - a^{n-1} - b^{n-1}}{a - b}$$

Proof. We use induction to show

$$D_n = \frac{a^{n+1} - b^{n+1}}{a - b}$$

Indeed, expand D_n by the first column we get

$$D_n = (a+b)D_{n-1} + (-1)^{1+2} ab \cdot 1 \cdot D_{n-2}$$

Thus

$$D_n - aD_{n-1} = b(D_{n-1} - aD_{n-2})$$

Thus, we get $\{D_n - aD_{n-1}\}_{n \geq 2}$ is a geometric series with ratio b. Thus we get closed formula

$$D_n - aD_{n-1} = (D_2 - aD_1)b^{n-2}$$

where an easy computation shows

$$D2 - aD_1 = b^2$$

and thus

$$D_n - aD_{n-1} = b^n$$

Now use

$$D_n = (a+b)D_{n-1} - abD_{n-2}$$

again to get

$$D_n - bD_{n-1} = a(D_{n-1} - bD_{n-2})$$

we conclude

$$D_n - bD_{n-1} = a^n$$

Solve the system

$$\begin{cases} D_n - aD_{n-1} = b^n \\ D_n - bD_{n-1} = a^n \end{cases}$$

we conclude

$$D_n = \frac{a^{n+1} - b^{n+1}}{a-b}$$

$$\begin{vmatrix} 2 & 1 \\ 0 & 3 \end{vmatrix}$$