Eigenvalues and Eigenvectors

Given a linear operator $X \xrightarrow{A} X$, a scalar $\lambda \in \mathbb{F}$ is called an eigen value of A iff the equation $Ax = \lambda x$ has a non-zero solutions; put another way, $\ker(A - \lambda I) := \{x \in \mathbb{X} \mid Ax = \lambda x\}$ is a non-zero subspace of X. Thus for dim $X = n < \infty$, λ is an eigen space of A iff λ is a solution of the equation.

$$x_A(\lambda) := det(A - \lambda I) = 0 \qquad ...(1)$$

(Since an operator $X \xrightarrow{B} X$ is invertible iff B is an isomorphism which happens, as observed in ***......Reference from handout 1****, iff for some matrix representation b for B, we have b invertible, i.e. $detb \neq 0$)*

The equation (1) is

$$x_A(\lambda) := a_0 + a_1\lambda + \ldots + a_n\lambda^n = 0$$

which is obtained from the polynomial

$$x_A(\theta) := a_0 + a_1\theta + \ldots + a_n\theta^n \in \mathbb{F}\theta$$

for some coefficients $a_0, \ldots, a_n \in \mathbb{F}$. This polynomial $x_A(\theta)$ is known as the characteristic polynomial of the operator A. When a is some matrix representation of A, there is the same polynomial $x_a(\theta)$ though we call it the characteristic polynomial of the matrix a. To see this, suppose b is some other matrix representation of A. Then $b = p^{-1}ap$ for some invertible matrix *($\because$ a change-of-basis matrix is the same thing as an invertible matrix, see the change-of-basis module.)* Then

$$\begin{aligned} x_b(\lambda) &= det(b - \lambda I_n) \\ &= det(p^{-1}ap - \lambda I_n) \\ &= det(p^{-1}(a - \lambda I_n)p) \\ &= (det(p^{-1}))det(a - \lambda I_n)(det(p)) \\ &= (det\ p)^{-1}x_a(\lambda)(det\ p) \\ &= x_a(\lambda) \end{aligned}$$

Thus λ is an eigenvalue of A

iff λ is an eigenvalue of a

iff $x_a(\lambda) \ = 0$

iff $x_b(\lambda) \ = 0$

iff λ is an eigenvalue of b

iff λ is an eigenvalue of B.

1. We note that the polynomial $x_A(\theta) := a_0 + a_1\theta + \ldots + a_n\theta^n$ is in fact a monic polynomial i.e. $a_n = 1$. Supposing $\mathbb{F} = \mathbb{C}$, this can be written $x_A(\theta) := (\theta - \lambda_1)^{n_1} \ldots (\theta - \lambda_k)^{n_k}$ where $\lambda_1, \ldots, \lambda_k$ exhausts all the eigen values of A; further, $n_1 + \ldots + n_k \ = n$ =dim X.

2. Comment The definition supplied in (1) is obviously concealing something: What ia after all the meaning of the determinant of an operator $T \in L(X,X)$*(in this case* $T = A - \lambda I$*)*?

One may then say, in view of the fact *(established above)* that since $\det(a - \lambda I_n) = \det(b - \lambda I_n)$ for any two matrices $a,\ b$ representing the operator $A \in L(X,X)$, the characteristic polynomial $x_A(\theta)$ of A is the same as $x_a(\theta) = det(a - \lambda I_n)$ for any matrix a representing A. This again is problematic since it is far from clear what the meaning of the determinant of a square matrix is. However, since you have been working with determinants anyway, this may be taken as a working tool, made available off the shelf.

A clean treatment of these things is out of question here(*and you are strongly requested to cease and desist from looking, e.g., at chapter 9, Vol II, of the book by Ronald Shaw mentioned in the first handout where this is a variable)* but once you have suppressed your mathematical conscience, there is nothing easier than this topic in linear algebra; as you know it from your school days

3. The collection $\{\lambda_1, \ldots, \lambda_k\} \ \subseteq \mathbb{F}$ mentioned in 1 for the linear transformation $X \ \xrightarrow{A} \ X$, dim $X = n \ < \infty$, is called the eigen spectrum of A; we shall denote it by A_e; its members are, as mentioned before, are called eigen values of A. Since there are no 'spectral values' other than these eigenvalues for the case dim $X = n \ < \infty$, a course on finite dimensional vector spaces may, and usually does, call it simply the 'spectrum' of A. But using the correct name never hurts.

The subspace $\ker(A - \lambda I)$ of X is called the eigen space of X corresponding to the eigenvalue λ.

4. The spectrum A_e may be empty: for instance if $\mathbb{F} = \mathbb{R}$, $X = \mathbb{R}^2$, $\left\{ e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ is chosen as a basis, $X \xrightarrow{A} X$ is defined by $A(e_1) = -e_2, \ A(e_2) = e_1$ then $x_A(\theta) \ = x_a(\theta) \ = \theta^2 + 1$ *(Verify that the matrix a with respect to the basis will be* $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ *since its columns are to be given by* $A(e_1) = -e_2, \ A(e_2) = e_1$ *then* $x_A(\lambda) \ = \begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix} = -(\lambda^2 + 1)$*)* and $x_A(\theta) = 0$ has no roots in $\mathbb{R}$.

But this will not be the case if the field is algebraically closed; for instance if $\mathbb{F} = \mathbb{C}$. Our illustrations be almost always for the case $\mathbb{F} = \mathbb{C}$.

5. The minimal polynomial of a linear operator $X \xrightarrow{A} X$, $\dim X = n \ < \infty$, is defined uniquely by the requirements*(denoting this by $\mu_A(\theta)$)*

1. $\mu_A(A) \ = 0 \ \in L(X, X)$
2. $q(A) \ = 0, \ q \neq 0 \Rightarrow \deg q \ \geq \deg \mu_A(\theta)$ for any $q = q(\theta) \ \in \ \mathbb{F}[\theta]$, and
3. $\mu_A(\theta)$ is a monic polynomial

The existence and uniqueness of $\mu_A(\theta)$ can be seen as follows: We consider the collection

$$\mathfrak{I} := \{\mathbb{F}[\theta] \mid X \xrightarrow{f(A)} X = 0 \ \in L(X, X)\}$$

If $f(\theta), \ g(\theta) \ \in \ \mathfrak{I}$ then clearly $\alpha f(\theta) + \beta g(\theta)$ and $f(\theta)g(\theta) \ \in \ \mathfrak{I}$ for any given $\alpha, \ \beta \ \in \mathbb{F}$. This means that $\mathfrak{I}$ is an ideal in the ring of polynomials; say it is the annihilator of A and denote it by $\text{Annh}(A)$. Since $\dim L(X, X) = n^2 < \infty$, the $n^2 + 1$ operators $Id, A, \ldots, A^{n^2}$ cannot be linearly independent in the vector space $L(X, X)$ and there must exist scalars $a_0, \ldots, a_n$ such that, with $A^0 = Id$,

$$a_0 A^0 + \ldots + a_{n^2} A^{n^2} = 0, \ a_j \neq 0 \text{ for some } j, \ 1 \leq j \leq n^2.$$

Therefore, $\text{Annh}(A)$ contains some non-zero polynomial $a(\theta)$, degree $a(\theta) \leq n^2$ and thus $\text{Annh}(A) \ \neq \{0\}$.

Now every ideal in $\mathbb{F}[\theta]$ is a principal ideal and hence there exists a single generator of $\text{Annh}(A)$, there will be exactly one which is monic. This is our $\mu_A(\theta)$.

Thus (3) fixes $\mu_A(\theta)$ uniquely while (1) and (2) determine it up to a scalar multiple.

6. More information about the minimal polynomial is available from the Cayley-Hamilton Theorem which says that the characteristic polynomial $x_A(\theta) \in \text{Annh}(A)$. We will accept this result off the shelf; any clean treatment needs more more machinery than we can afford to build her. But note that since degree But note that since degree $(x_A(\theta)) = n$, and $\mu_A(\theta)$ must divide $x_A(\theta)$*($\therefore$ $\mu_A(\theta)$ divides each $f(\theta) \in$ Annh(A))*, so that deg$\mu_A(\theta) \leq n$. Then since

$$x_A(\theta) = (\theta - \lambda_1)^{n_1} \ldots (\theta - \lambda_k)^{n_k}$$

we get

$$\mu_A(\theta) = (\theta - \lambda_1)^{p_1} \ldots (\theta - \lambda_k)^{p_k},\ 1 \leq p_i \leq n_i \text{ for } 1 \leq i \leq k.$$

Thus for $X \xrightarrow{A} X$, $\lambda \in \mathbb{F}$

7. We note that an eigenvector of A is a non zero vector *($\because$ $Ax = \lambda x$ must have a non zero solution by definition.)* Assuming that whenever $\lambda_1, \ldots, \lambda_{r-1}$ are distinct eigen values of A, the corresponding eigen vectors $x_1, \ldots, x_{r-1}$ are linearly independent, we find that in

$$\alpha_1 x_1 + \ldots + \alpha_{r-1} x_{r-1} + \alpha_r x_r \ = 0$$

subject to $\alpha_1, \ldots, \alpha_r$ not all zero, we must have $\alpha_r \neq 0$. *($\because$ otherwise, $\alpha_1 x_1 + \ldots + \alpha_{r-1} x_{r-1} = 0$ and this forces $\alpha_1 = \ldots = \alpha_{r-1} = 0$ since $x_1, \ldots, x_{r-1}$ are linearly independent by hypothesis; as $\alpha_r = 0$ is already there, the condition "subject to $\alpha_1, \ldots, \alpha_r$ not all zero" is violated).* If x_r is an eigenvector with λ_r as eigenvalue *(so that $Ax_r = \lambda_r x_r$)* with $\lambda_r \neq \lambda_j$ for $1 \leq j \leq r-1$, we obtain

$$\begin{aligned} 0 &= A0 \\ &= A(\alpha_1 x_1 + \ldots + \alpha_r x_r) \\ &= \alpha_1 A x_1 + \ldots + \alpha_r A x_r \\ &= \alpha_1 \lambda_1 x_1 + \ldots + \alpha_r \lambda_r x_r \end{aligned}$$

Now $\lambda_r = 0$ means no $\lambda_j = 0$ for $1 \leq j \leq r-1$ *($\because$ $\lambda_r \neq \lambda_j$ for $1 \leq j \leq r-1$)* but then we are left with $0 = \beta_1 x_1 + \ldots + \beta_{r-1} x_{r-1}$ where $\beta_j = \alpha_j \lambda_j$ for $1 \leq j \leq r-1$, which forces $\beta_j = 0$ for each j which forces $\alpha_j = 0$ for each j*(since $\lambda_1, \ldots, \lambda_{r-1}$ are non-zero).* But then $\alpha_r = 0$ must be true and "subject to $\alpha_1, \ldots, \alpha_r$ not all zero" is again violated. If $\lambda_r \neq 0$, we have $-\alpha_r x_r = \frac{1}{\lambda_r}[\alpha_1 \lambda_1 x_1 + \ldots + \alpha_{r-1} \lambda_{r-1} x_{r-1}]$ which means $(\lambda_r - \lambda_1)\alpha_1 x_1 + \ldots + (\lambda_r - \lambda_{r-1})\alpha_{r-1} x_{r-1} = 0$. Since $x_1, \ldots, x_{r-1}$ are assumed linearly independent, we have $(\lambda_r - \lambda_j)\alpha_j = 0$ for $1 \leq j \leq r-1$

which forces $\alpha_j = 0$ for $1 \leq j \leq r-1$. But if $\alpha_1, \ldots, \alpha_{r-1} = 0$, we get $\alpha_r = 0$ again, which is a contradiction as we know. We conclude that it is not possible to have $0 = \alpha_1 x_1 + \ldots + \alpha_r x_r$ unless $\alpha_1 = \ldots = \alpha_r = 0$ when $x_1, \ldots, x_r$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ i.e. the vectors $\{x_1, \ldots, x_{r-1}\}$ are.

Since the empty set is always linearly independent *(which corresponds to* $r = 0$ *in the preceding)*; we conclude

The eigenvectors associated with distinct eigenvalues of $X \to AX$ are linearly independent.

and therefore

If $\lambda_1, \ldots, \lambda_r$ are distinct eigenvalues of $X \to AX$, the sum

$$\oplus_{j=1}^{r} \mathbf{ker}(A - \lambda_r Id)$$

is a direct sum