Generalized Eigenspaces

Key facts. *Sometimes there are not enough eigenvectors to form a basis. There is always a basis of generalized eigenvectors and that allows one to construct fundamental matrices. This also implies the diagonalisability of symmetric matrices.*

1 Generalized eigenvalues.

Suppose that A is a linear transformation from a finite dimensional complex vector space $\mathbb{V}$ to itself and that λ is an eigenvalue of A. Then $\ker(A-\lambda I)$ is the linear space consisting of all eigenvectors with eigenvalue λ together with the zero vector.

The space $\ker(A-\lambda I)^2$ contains that space since if $(A-\lambda I)v=0$ then surely $(A-\lambda I)(A-\lambda I)v = 0$. In this way for $\ell = 1, 2, \ldots$ define a nondecreasing sequence of linear subspaces,

$$K_\ell \ := \ \ker\big(A-\lambda I\big)^\ell, \qquad \text{so} \qquad 0 \neq K_1 \subseteq K_2 \subseteq K_3 \subseteq \cdots .$$

Example 1.1 *The matrix*

$$\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$$

has characteristic polynomial

$$\det(zI - A) \ = \ (z-3)^2,$$

so has only the eigenvalue $\lambda = 3$. For this eigenvalue,

$$K_1 \ = \ \ker(A-3I) \ = \ \mathbb{C}\,(1,0), \qquad K_2 \ = \ \ker(A-3I)^2 \ = \ \ker\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \ = \ \mathbb{C}^2.$$

Proposition 1.1 *There is an $r \leq \dim\mathbb{V}$ so that*

$$0 \ \neq \ K_1 \ \subsetneq \ K_2 \ \subsetneq \ \cdots \ \subsetneq \ K_r \ = \ K_{r+1} \ = \ K_{r+2} \ = \ \cdots .$$

One has,

$$\dim K_r \ \geq \ r\,. \tag{1.1}$$

Proof. Whenever there is strict inclusion the dimension must increase by at least one. Thus, there can be at most $\dim\mathbb{V}$ strict inclusions $K_\ell \neq K_{\ell+1}$. Therefore, there is a first $r \geq 1$ so that $K_{r+1} = K_r$.

It then suffices to prove the following assertion. *If $K_\ell = K_{\ell+1}$, then $K_{\ell+1} = K_{\ell+2}$.* The assertion is proved as follows. If $v \in K_{\ell+2}$ then $(A-\lambda I)^{\ell+1}(A-\lambda I)v = 0$ so $(A-\lambda I)v \in K_{\ell+1} = K_\ell$ so $(A-\lambda I)^\ell(A-\lambda I)v = 0$ proving that $v \in K_{\ell+1}$.

Since the inclusion of the K_ℓ is strict before r, the dimensions satisfy,

$$\dim K_1 \ \geq \ 1, \qquad \dim K_\ell \ \geq \ \dim K_{\ell-1} + 1, \quad \text{for} \quad 2 \leq \ell \leq r,$$

proving (1.1). ■

Definition 1.1 *The space K_r is called the* **generalized eigenspace** *associated to λ and r is its* **index***.*

Proposition 1.2 *If K_r is the generalized eigenspace associated to λ and its index is r, define*

$$Y \;=\; \text{range}\,(A - \lambda I)^r\,.$$

K_r and Y are invariant under A, that is $A(K_r) \subset K_r$ and $A(Y) \subset Y$. In addition,

$$\mathbb{V} = K_r \oplus Y\,. \tag{1.2}$$

Proof. K and Y are the kernel and range of the linear transformation $(A - \lambda I)^r$. It follows that

$$\dim K_r \;+\; \dim Y \;=\; \dim \mathbb{V}\,.$$

To prove (1.2) it therefore suffices to show that

$$K_r \cap Y = 0\,.$$

If $v \in K_r \cap Y$ then,

$$(A - \lambda I)^r v = 0, \qquad \text{and there is a } u \text{ s.t. } \; v = (A - \lambda I)^r u\,.$$

Then $(A - \lambda I)^{2r} u = 0$. That is, $u \in K_{2r} = K_r$ so, $v = (A - \lambda I)^r u = 0$.
For the invariance of K, if $v \in K_r$ then $(A - \lambda I)^r v = 0$ so,

$$(A - \lambda I)^r A v \;=\; A(A - \lambda I)^r v \;=\; A\,0 \;=\; 0,$$

proving that $Av \in K_r$.
For the invariance of Y, if $w \in Y$ then there is a u so that $w = (A - \lambda I)^r u$ then

$$Aw \;=\; A(A - \lambda I)^r u \;=\; (A - \lambda I)^r (Au) \;\in\; \text{Range}\,(A - \lambda I)^r \;:=\; Y\,.$$

■

This result splits A into two pieces, the part in K and the part in Y. The next result shows that that split separates the part with eigenvalue λ from the rest.

Proposition 1.3 *λ is the only eigenvalue of $A|_{K_r}$, and, λ is not an eigenvalue of $A|_Y$.*

Proof. To prove the first assertion suppose that $\widetilde{\lambda} \neq \lambda$ and $v \in K_r$ satisfies $Av = \widetilde{\lambda} v$. Then

$$(A - \lambda I)v = (\widetilde{\lambda} - \lambda)v\,, \qquad \text{and therefore,} \qquad (A - \lambda I)^r v \;=\; (\widetilde{\lambda} - \lambda)^r v\,.$$

Since $v \in K_r$ one has,

$$0 \;=\; (A - \lambda I)^r v \;=\; (\widetilde{\lambda} - \lambda)^r v,$$

so $v = 0$.
To prove the second it is sufficient to show that $\ker(\lambda I_Y - A|_Y) = 0$. If $w \in Y$ and $(\lambda I_Y - A|_Y)w = 0$, then $w \in K_1 \subset K_r$ so $w \in K_r \cap Y = 0$. ■

2 Completeness of the generalized eigenspaces.

Theorem 2.1 *Suppose that A is a linear transformation from a finite dimensional complex vector space $\mathbb{V}$ to itself with characteristic polynomial*

$$\det(zI - A) \;=\; \Pi_{j=1}^{k}(z-\lambda_j)^{m_j}, \qquad \lambda_j \ \text{distinct}\,, \qquad \sum m_j = \dim \mathbb{V}\,.$$

Denote by X_j the generalized eigenspaces associated to λ_j.
i. *The index of the generalized eigenspace of λ_j is $\le m_j$, so,*

$$X_j \;:=\; \ker\,(A - \lambda_j I)^{m_j}\,.$$

ii. $\dim X_j \;=\; m_j$.
iii. $\mathbb{V} \;=\; X_1 \;\oplus\; X_2 \;\oplus \cdots \oplus X_k\,.$

Proof. **i,ii.** For ease of reading fix one of the λ_j and call it λ. Proposition 1.2 shows that

$$A \;=\; A|_{K_r} \oplus A|_Y\,,$$

the summands being the restrictions of A to the invariant subspaces. It follows that for all z,

$$\det(zI - A) \;=\; \det(zI_{K_r} - A|_{K_r})\; \det(zI_Y - A|_Y)\,. \tag{2.1}$$

Since λ is the only eigenvalue of $A|_K$ one has

$$\det(zI_{K_r} - A|_{K_r}) \;=\; (z-\lambda)^{\dim K_r}\,.$$

Since λ is not an eigenvalue of $A|_Y$ one has

$$\det(\lambda I_Y - A|_Y) \;\neq\; 0\,.$$

Therefore (2.1) shows that the multiplicity of the root λ is equal to $\dim K_r$. Since the mulitplicity of the root λ is equal to m, one has

$$m \;=\; \dim K_r \;\ge\; r\,,$$

the last estimate from (1.1).
Then,

$$X \;:=\; K_m \;=\; K_r, \qquad \text{whence} \qquad \dim X \;=\; \dim K_r \;=\; m\,, \tag{2.2}$$

proving **i.** and **ii.**
iii. [1] Using (2.2),

$$\dim X_1 \times X_2 \times \cdots \times X_k \;=\; \sum \dim X_j \;=\; \sum_1^k m_j \;=\; \dim \mathbb{V}\,.$$

The map

$$X_1 \times X_2 \times \cdots \times X_k \;\ni\; (x_1, x_2, \ldots, x_k) \quad\mapsto\quad x_1 + x_2 + \cdots + x_k \;\in\; \mathbb{V}$$

[1] The proof of this step on page 279 in Brauer and Nohel is incorrect.

is a linear map of spaces of the same dimension. To prove the direct sum assertion of the Theorem it is sufficient to show that it is injective.

If $x_1 + \cdots + x_k = 0$, multiplying by $\Pi_{j\neq\mu}(A - \lambda_j I)^{m_j}$ annihilates all the summands except the x_μ term to yield,

$$\Pi_{j\neq\mu}(A - \lambda_j I)^{m_j} x_\mu \;=\; 0\,.$$

Proposition 1.3 shows λ_μ is the only eigenvalue of $A|_{X_\mu}$, so for $j \neq \mu$, $A - \lambda_j I$ is and invertible map of X_μ to itself. Therefore $\Pi_{j\neq\mu}(A - \lambda_j I)^{m_j}$ is invertible from X_μ to itself. It follows that $x_\mu = 0$. Since this is true for each μ one has $x_1 = x_2 = \cdots = x_k = 0$ proving injectivity. ■

3 Diagonalisability Symmetric Matrices.

It is not hard to derive the fact that symmetric matrices have an orthonormal basis of eigenvectors from the result of the preceding section. Recall that the analogue in a general scalar product space of symmetric matrices are the linear transformations which satisfy

$$\langle Av, w\rangle \;=\; \langle v, Aw\rangle$$

for all v, w where $\langle\ ,\ \rangle$ denotes the scalar product.

Example 3.1 *If $\mathbb{V} = \mathbb{C}^n$ the standard scalar product is*

$$\langle v, w\rangle \;=\; \sum x_j\, \overline{y}_j\,,$$

and the corresponding symmetric linear transformations are the matrices satisfying

$$A_{ij} \;=\; A^*_{ji}\,.$$

They are hermitian symmetric. For real matrices that reduces to symmetry.

Theorem 3.2 *Suppose that $\mathbb{V}$ is a complex scalar product space and A is a symmetric linear transformation from $\mathbb{V}$ to itself.*

i. *The eigenvalues of A are real.*

ii. *For each eigenvalue, the index $r = 1$ that is $K_\ell = K_1$ for all ℓ so every generalized eigenvector is an eigenvector.*

iii. *Eigenvectors with distinct eigenvalues are orthogonal.*

Proof. **i.** If λ is an eigenvalue, choose a unit eigenvector v to find,

$$\lambda = \lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle = \overline{\lambda}\,.$$

ii. Must show that $K_2 = K_1$. That is if $(A-\lambda)^2 v = 0$ then $(A - \lambda I)v = 0$. Compute,

$$\|(A - \lambda I)v\|^2 = \langle (A - \lambda I)v, (A - \lambda I)v\rangle = \langle (A - \lambda I)^2 v, v\rangle = \langle 0, v\rangle = 0\,.$$

iii. If $Av = \lambda_1 v$ and $Aw = \lambda_2 w$ with $\lambda_1 \neq \lambda_2$ must show that $\langle v, w\rangle = 0$. Compute using the fact that the eigenvalues are real,

$$\lambda_1\langle v, w\rangle = \langle \lambda_1 v, w\rangle = \langle Av, w\rangle = \langle v, Aw\rangle = \langle v, \lambda_2 w\rangle = \lambda_2\langle v, w\rangle\,.$$

Therefore $(\lambda_1 - \lambda_2)\langle v, w\rangle = 0$ showing that $\langle v, w\rangle = 0$. ■