Eigenvalues and eigenvectors

Any $n \times n$ matrix $\vec{A}$ defines a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n, \vec{x} \mapsto \vec{A}\vec{x}$. For a nonzero vector $\vec{x} \in \mathbb{R}^n$ we say $\vec{x}$ is an **eigenvector** of $\vec{A}$ if $\vec{A}\vec{x}$ is a scalar multiple of $\vec{x}$, i.e. there exists a scalar $\lambda \in \mathbb{R}$ such that $\vec{A}\vec{x} = \lambda\vec{x}$. Likewise, we say that $\lambda \in \mathbb{R}$ is an **eigenvalue** of $\vec{A}$ if $\vec{A}\vec{x} = \lambda\vec{x}$ for some $\vec{x} \in \mathbb{R}^n$. Iff λ is an eigenvalue of $\vec{A}$, there exists a nonzero solution to $(\vec{A} - \lambda\vec{I})\vec{x}$, and therefore $(\vec{A} - \lambda\vec{I})$ is **not** invertible. Thus, all nonzero vectors $\vec{v} \in \ker(\vec{A} - \lambda\vec{I}) \subset \mathbb{R}^n$ are eigenvectors of $\vec{A}$, with eigenvalue λ.

For example, given

$$\vec{A} = \begin{bmatrix} -1 & 3 \\ 3 & -1 \end{bmatrix}$$

we can see that an eigenvalue of $\vec{A}$ is $\lambda = 2$, and so

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

is a singular matrix, and that the corresponding eigenvector is $\vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Another eigenvalue of $\vec{A}$ is $\lambda = -4$, with eigenvector $\vec{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Note that the two eigenvectors are distinct, as they are linearly independent.

For $\vec{x} \in \mathbb{R}^2$, $\vec{x} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ for unique $c_1, c_2 \in \mathbb{R}$. Therefore,

$$\vec{a}\vec{x} = c_1\vec{A} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\vec{A} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = c_1\lambda_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\lambda_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

This gives us a very easy way to exponentiate $\vec{A}$:

$$\vec{A}^n\vec{x} = c_1\lambda_1^n \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\lambda_2^n \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

As an aside, given $\vec{P} = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}$,

$$\vec{A} = \vec{P} \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix} \vec{P}^{-1}$$

This means that $\vec{A}$ is **diagonalisable**, and it is **similar** to $\begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}$.

Theorem: If $\vec{A}$ is upper triangular, then its eigenvalues are its diagonal entries.

Proof. Given a matrix $\vec{A}_{n\times n}$ and an eigenvalue λ

$$\vec{A} - \lambda\vec{I} = \begin{bmatrix} a_{11}-\lambda & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & a_{nn}-\lambda \end{bmatrix}$$

and so $\det(\vec{A} - \lambda\vec{I}) = \prod_{i=1}^{n}(a_{ii} - \lambda)$. Therefore, $(\vec{A} - \lambda\vec{I})$ is not invertible iff $\lambda = a_{kk}$ for some $1 \leq k \leq n$. $\square$

Theorem: Suppose $\vec{v}_1, \vec{v}_2 \cdots \vec{v}_r$ are eigenvectors with corresponding eigenvalues $\lambda_1, \lambda_2 \cdots \lambda_r$. Assume these eigenvalues are distinct. Then, $\{\vec{v}_1, \vec{v}_2 \cdots \vec{v}_r\}$ are linearly independent.

Proof. Let $1 \leq j \leq r$ be the largest index for which $\{\vec{v}_1, \vec{v}_2 \cdots \vec{v}_r\}$ are linearly independent. If $j = r$, we are done. Otherwise, if $j < r$, then

$$\vec{v}_{j+1} = \sum_{i=1}^{j} c_i \vec{v}_i, c_i \in \mathbb{R}\ \forall\ 1 \leq i \leq j \tag{1}$$

Multiplying (1) by $\vec{A}$ we get

$$\vec{A}\vec{v}_{j+1} = \sum_{i=1}^{j} c_i \vec{A}\vec{v}_i$$

$$\lambda_{j+1}\vec{v}_{j+1} = \sum_{i=1}^{j} c_i \cdot \lambda_i \vec{v}_i$$

but, if we multiply (1) by λ_{j+1} instead of $\vec{A}$, we get

$$\lambda_{j+1}\vec{v}_{j+1} = \sum_{i=1}^{j} c_i \cdot \lambda_{j+1} \vec{v}_i$$

Subtracting the two equations, we get

$$0 = \sum_{i=1}^{j} c_i(\lambda_i - \lambda_{j+1})\vec{v}_i$$

Since we know that $\{\vec{v}_1, \vec{v}_2 \cdots \vec{v}_j\}$ are linearly independent, $c_i(\lambda_i - \lambda_{j+1}) = 0 \ \forall \ 1 \leq i \leq j$. Per equation (1), we know that at least one coefficient $c_i \neq 0$, so for some i, $\lambda_i - \lambda_{j+1} = 0 \implies \lambda_i = \lambda_{j+1}$. This is a contradiction, as we assumed that all the eigenvalues are distinct. Therefore, $j = r$, and we are done. $\square$

Eigenspaces

Given a matrix $\vec{A}_{n\times n}$, the **eigenspace** of $\vec{A}$ is the vector space containing the eigenvectors of $\vec{A}$ for a certain eigenvalue λ, equivalent to

$$v_\lambda = \ker(\vec{A} - \lambda\vec{I})$$

We can say that λ is an eigenvalue iff $\ker(\vec{A} - \lambda\vec{I}) \neq \{\vec{0}\}$, that is, iff $(\vec{A} - \lambda\vec{I})$ is singular, which in turn is equivalent to

$$\det(\vec{A} - \lambda\vec{I}) \neq 0$$

From this we get the **characteristic polynomial** of $\vec{A}$ with regards to λ:

$$P_A(\lambda) = \det(\vec{A} - \lambda\vec{I})$$

and we can define the eigenvalues of $\vec{A}$ as the roots of $P_A(x) = 0$.

For example, for $\vec{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$:

$$\begin{aligned} P_A(\lambda) = \det(\vec{A} - \lambda\vec{I}) = \det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} - 2\vec{I}\right) = \det\begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} \\ = (a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda + ad - bc \end{aligned}$$

In fact, the final constant of P_A will always be $\det(\vec{A})$. This can be seen by setting $\lambda = 0$ in the definition for the characteristic equation. The penultimate coefficient will be the **trace** of $\vec{A}$, which is the sum of its diagonal elements. In general, for an upper triangular matrix:

$$P_A(\lambda) = \prod_{i=1}^{n}(a_{ii} - \lambda) \tag{2}$$

Note that, if a root of $P_A(\lambda) = 0$ is multiple, the multiplicity of the root denotes the upper limit of the dimension of the corresponding eigenspace, that is, the size of the largest set of linearly independent vectors $\vec{v}$ such that $\det(\vec{A} - \lambda\vec{I}) = 0$.

Similarity

As mentioned above, there is a concept of **similarity** between matrices. $\vec{A}_{n\times n}$ and $\vec{B}_{n\times n}$ are similar, that is, $\vec{A}\ \vec{B}$, iff there exists an invertible matrix $\vec{P}_{n\times n}$ such that $\vec{A} = \vec{P}\vec{B}\vec{P}^{-1}$. This is an **equivalence relation**, that is, for any $\vec{A}, \vec{B}, \vec{C}$:

1. $\vec{A}\ \vec{A}$
2. $\vec{A}\ \vec{B} \iff \vec{B}\ \vec{A}$
3. $\vec{A}\ \vec{B}, \vec{B}\ \vec{C} \implies \vec{A}\ \vec{C}$

$$\boxed{\text{Theorem: } \vec{A}\ \vec{B} \implies P_A(\lambda) = P_B(\lambda)}$$

Proof. $P_A = \det(\vec{A} - \lambda\vec{I}) = \det(\vec{P}\vec{B}\vec{P}^{-1} - \lambda\vec{P}\vec{I}\vec{P}^{-1}) = \det(\vec{P}^{-1}(\vec{B} - \lambda\vec{I})\vec{P}^{-1}) = \det(\vec{B} - \lambda\vec{I}) = P_B$ □