Eigenvalues, Eigenvectors and Eigenspace

Finding Nemo:

For a matrix A, if there exists a scalar λ and a vector $\mathbf{v} \neq 0$, such that $\boxed{A\mathbf{v} = \lambda\mathbf{v}}$, then we call λ an eigenvalue and $\mathbf{v}$ an eigenvector of A. It follows that $A\mathbf{v} - \lambda\mathbf{v} = (A - \lambda I)\mathbf{v} = 0$. $\therefore \boxed{\det(A - \lambda I) = 0}$. It means that applying A to vector $\mathbf{v}$ does not change its direction but scales it by a factor of λ (the eigenvalue). Not all matrices have eigenvalues (and therefore eigenvectors). When one does, the collection of all its eigenvectors is called the eigenspace.

e.g. The eigenspace of a flip matrix is the line or plane it flips over. If it does not shear in other directions, the axis it flips along is an eigenspace as well, with a negative eigenvalue. There are no eigenvalues for a rotation matrix as no vectors point to the same direction afterwards. (Product of two matrices with eigenvalues do not necessarily have eigenvalues.)

For diagonal matrices, the eigenvalues are on the diagonal. If all eigenvalues are distinct, the eigenspace is the axes (less 0).

e.g. A 3×3 diagonal matrix $D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$ has eigenvalues λ_1, λ_2 and λ_3, so $D\mathbf{i} = \lambda_1\mathbf{i}$, $D\mathbf{j} = \lambda_2\mathbf{j}$, and $D\mathbf{k} = \lambda_3\mathbf{k}$.

That means the eigenspace is the union of $a_x\mathbf{i}, a_y\mathbf{j}$, and $a_z\mathbf{k}$, where $a_x, a_y, a_z \in \mathbb{R} - \{0\}$. If there are identical eigenvalues, e.g. $\lambda_1 = \lambda_2$, any vectors on the xy-plane are eigenvectors as they are sheared equally on both axises. Generally, if there are r identical eigenvalues, their corresponding eigenspace will be of r-dimension. If all eigenvalues are identical, the matrix becomes λI and its eigenspace will be $\mathbb{R}^n$.

Similarly, for a triangular matrix A, $\det(A - \lambda I) = \prod_{k=1}^{n}(a_{kk} - \lambda) = 0$, so all eigenvalues are on the diagonal as well. The eigenspace are not the axises, but it is still true that r identical eigenvalues correspond to an eigenspace of r-dimension. (The above analysis implies that an $n \times n$ matrix has at most n distinct eigenvalues.)

Row and column operations do not change the eigenvalues (even though the eigenvectors would change), one can reduce a matrix into a triangular one and find the eigenvalues, then substitute each eigenvalues into $A - \lambda I$ to solve for the eigenvectors.

Characteristic polynomial:

If matrix A has distinct eigenvalues $\lambda_r (r = 1, 2, \ldots n)$, it will have r independent eigenvectors $\mathbf{v_r}$. $\because\ A\mathbf{v_r} = \lambda_r \mathbf{v_r}$,

$$A\big[\mathbf{v_1 v_2} \ldots \mathbf{v_n}\big] = \big[A\mathbf{v_1}\ A\mathbf{v_2} \ldots A\mathbf{v_n}\big] = \big[(\lambda_1 \mathbf{v_1})(\lambda_2 \mathbf{v_2}) \ldots (\lambda_n \mathbf{v_n})\big] = \big[\mathbf{v_1 v_2} \ldots \mathbf{v_n}\big] D, \quad \text{where } D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

$\therefore\ AP = PD$, where $P = \big[\mathbf{v_1 v_2} \ldots \mathbf{v_n}\big]$. As $\mathbf{v_r}$ are distinct, P is invertable, so $A = PDP^{-1}$.

Since $\det(AP) = \det(PD)$, $\det(A)\det(P) = \det(P)\det(D)$, $\boxed{\det(A) = \det(D) = \prod_{k=1}^{n} \lambda_k.}$ If A is invertible, so is D.

We will also have $\boxed{A^n = PD^nP^{-1}}$ as $A^n = PDP^{-1} \cdot PDP^{-1} \ldots PDP^{-1} = PD(P^{-1}P)D(P^{-1}P) \ldots DP^{-1} = PD^nP^{-1}$.

For any polynomial $g(A) = \sum_{k=0}^{n} c_k A^k = \sum_{k=0}^{n} c_k (PD^kP^{-1}) = P \left(\sum_{k=0}^{n} c_k D^k \right) P^{-1} = P\ g(D)\ P^{-1}$, $\boxed{g(A) = P\ g(D)\ P^{-1}.}$

$$D^k = \begin{bmatrix} \lambda_1^k & 0 & \ldots & 0 \\ 0 & \lambda_2^k & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n^k \end{bmatrix}, \quad g(D) = \sum_{k=0}^{n} c_k D^k = \begin{bmatrix} \sum c_k \lambda_1^k & 0 & \ldots & 0 \\ 0 & \sum c_k \lambda_2^k & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sum c_k \lambda_n^k \end{bmatrix} = \begin{bmatrix} g(\lambda_1) & 0 & \ldots & 0 \\ 0 & g(\lambda_2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & g(\lambda_n) \end{bmatrix}$$

Let $f(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix} = (a_{11} - \lambda)(a_{12} - \lambda) \ldots (a_{nn} - \lambda) + \ldots \text{(various terms)} \ldots$

As $f(A) = \det(A - AI) = 0$, $\boxed{f(A) = P\ f(D)\ P^{-1} = 0}$ where $f(\lambda) = \det(A - \lambda I)$. (Cayley-Hamilton Theorem)

The function $f(x) = \det(A - xI)$ is called the characteristic polynomial of matrix A. The eigenvalues of A are the roots of its characteristic polynomial $f(x)$. In other words, given a transformation matrix, one may find its eigenvalues by solving the characteristic polynomial. Like determinants provide a "connection" between matrices and scalars, eigenvalues connect matrices and polynomials.

Note that this idea of "connections" is not a mathematical concept but helps you think. For example, the fact that $\det(AB) = \det(A) \cdot \det(B)$ and $\det(A^{-1}) = \dfrac{1}{\det(A)}$ would lead to A having no inverse if $\det(A) = 0$.