The Second Problem of Linear Algebra

Let A denote an n by n matrix. It is clear that for any choice of the scalar, $\lambda,$ the equation $A\vec{x} = \lambda\vec{x},$ has the solution $\vec{x} = \vec{0}.$ However, if λ is such that $A\vec{x} = \lambda\vec{x},$ has nontrivial solutions, $\vec{x},$ then we say that λ is an **eigenvalue for A**, and the corresponding nontrivial solutions are said to be **eigenvectors for A** corresponding to the eigenvalue $\lambda.$ Note that if $\vec{x} \neq \vec{0},$ *and* $A\vec{x} = \lambda\vec{x},$ then for every nonzero scalar $\alpha,$ $A(\alpha\vec{x}) = \lambda(\alpha\vec{x}),$ so any nonzero scalar multiple of an eigenvector for A corresponding to the eigenvalue λ is again an eigenvector for A corresponding to the eigenvalue $\lambda.$ The problem of finding the eigenvalues and eigenvectors for the matrix A is the second problem of linear algebra.

If $\vec{x}$ is an eigenvector for A corresponding to the eigenvalue $\lambda,$ then $\vec{x} \neq \vec{0}$ *and* $A\vec{x} - \lambda\vec{x} = (A - \lambda I)\vec{x} = \vec{0};$ i.e, $\vec{x} \in N_{A-\lambda I} =: N_\lambda$ *and* $\vec{x} \neq \vec{0}.$ Then $\dim N_\lambda > 0$ and the results of the previous section imply that the rank of $A - \lambda I$ is less than n. In this case, it is well known that the determinant of $A - \lambda I$ must be zero. It is also well known that since A is n by n, $\det(A - \lambda I)$ is a polynomial in λ of degree n. Then we have the following equivalent ways of characterizing eigenvalues and eigenvectors,

λ is an eigenvalue *for A if* :

$i)\quad A\vec{x} = \lambda\vec{x},\ \textit{for some}\ \vec{x} \neq \vec{0},$
$ii)\quad \det(A - \lambda I) = P_n(\lambda) = 0$
$iii)\ rank(A - \lambda I\ ) \ < n$
$iv)\ \dim N_\lambda > 0$

$\vec{x}$ is an eigenvector *for A* corresponding to the eigenvalue λ *if* :

$i)\ \vec{x} \neq \vec{0},\ \textit{and}\ A\vec{x} = \lambda\vec{x},$
$ii)\ \vec{x} \in N_{A-\lambda I} = N_\lambda$

Example- Consider the matrix $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, Then

$$\det(A - \lambda I) = \det\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} = 3 - 4\lambda + \lambda^2 = P_2(\lambda)$$

and the eigenvalues of A are the roots of the equation, $P_2(\lambda) = 0;\ i.e.,\ \lambda_1 = 1,\ \lambda_2 = 3.$ Next, we find the eigenvectors of $A.$ The eigenvector corresponding to $\lambda = 1$ lies in the null space of

$$(A - \lambda I)|_{\lambda=1} = (A - I) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Since $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \vec{0}$ if and only if $x_1 + x_2 = 0,$

$\vec{x} \in N\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right)$, if and only if $\vec{x}$ is a multiple of $\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \vec{E}_1$

Then all eigenvectors of A associated with λ_1 are multiples of $\vec{E}_1$. The eigenvector corresponding to $\lambda = 3$ lies in the null space of

$$(A - \lambda I)|_{\lambda=3} = (A - 3I) = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

For the null space of $(A - 3I)$ we have

$\vec{x} \in N\left(\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\right)$, if and only if $\vec{x}$ is a multiple of $\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \vec{E}_2$

Note that both $\lambda_1 = 1,\ \lambda_2 = 3$ are real numbers and $\vec{E}_1 \cdot \vec{E}_2 = 0$. This is not accidental. We will show that for any real symmetric matrix, the eigenvalues are real and the eigenvectors associated with distinct eigenvalues are orthogonal. First, however, we consider another example. For the matrix

$A = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$, the eigenvalues are : $\lambda_1 = 2i,\ \lambda_2 = -2i$

and the eigenvectors are :

$\vec{E}_1 = \begin{bmatrix} -i \\ 1 \end{bmatrix}$ associated with $\lambda = 2i,$

$\vec{E}_2 = \begin{bmatrix} i \\ 1 \end{bmatrix}$ associated with $\lambda = -2i.$

In this case note that $\vec{E}_1 \cdot \vec{E}_1 = \vec{E}_2 \cdot \vec{E}_2 = 0 \ but \ \vec{E}_1 \cdot \vec{E}_2 \neq 0$. This is bad. Neither $\vec{E}_1$ nor $\vec{E}_2$ is the zero vector but each appears to have zero length. The problem here lies in the definition of the dot or inner product. We have to modify the previous definition to account for vectors having complex entries. The new definition for the inner product is,

$$\vec{x} \cdot \vec{z} = x_1\bar{z}_1 + \cdots + x_n\bar{z}_n \qquad where \qquad \bar{z}_k = \text{the complex conjugate of } z_k$$

i.e., if $z_k = \alpha_k + i\beta_k \ then \ \bar{z}_k = \alpha_k - i\beta_k$. With this definition, $\vec{E}_1 \cdot \vec{E}_1 \neq 0,\ \ \vec{E}_2 \cdot \vec{E}_2 \neq 0 \ and$ $\vec{E}_1 \cdot \vec{E}_2 = 0$. In addition, when the vectors involved both have only real components, then the new definition reduces to the old one. For the new definition, we have the following properties,

$$\vec{x} \cdot \vec{z} = \overline{\vec{z} \cdot \vec{x}} \qquad and \qquad (\alpha\vec{x}) \cdot \vec{z} = \alpha(\vec{x} \cdot \vec{z}) \quad hence \quad \vec{x} \cdot (\alpha\vec{z}) = \bar{\alpha}\,(\vec{x} \cdot \vec{z}).$$

Now we can show that for any real symmetric matrix, the eigenvalues are real, i.e.,for every eigenvalue, $\lambda = \bar{\lambda}$.

i) for any matrix A (symmetric or not) and all vectors $\vec{x}, \vec{z}$ $\quad (A\vec{x}) \cdot \vec{z} = \vec{x} \cdot (A^{\top}\vec{z})$
then if A is symmetric (i.e., $A = A^{\top}$) $\quad (A\vec{x}) \cdot \vec{z} = \vec{x} \cdot (A\,\vec{z})$

ii) if $A\vec{x} = \lambda\vec{x}$ then $(A\vec{x}) \cdot \vec{x} = \vec{x} \cdot (A\,\vec{x})$
becomes $(A\vec{x}) \cdot \vec{x} = \vec{x} \cdot (A\,\vec{x})$ or $(\lambda\vec{x}) \cdot \vec{x} = \vec{x} \cdot (\lambda\,\vec{x})$
or, $\lambda(\vec{x} \cdot \vec{x}) = \bar{\lambda}(\vec{x} \cdot \vec{x})$, by the property shown above.

iii) finally, since $\vec{x}$ is an eigenvector, $\vec{x} \cdot \vec{x} \neq 0$ so $\lambda = \bar{\lambda}$.

To show that the eigenvectors associated with distinct eigenvalues of a real symmetric matrix are orthogonal, suppose $\quad A\vec{x} = \lambda\vec{x} \quad and \quad A\vec{z} = \mu\vec{z},\ \lambda \neq \mu.$

Then $\quad (A\vec{x}) \cdot \vec{z} = \vec{x} \cdot (A\,\vec{z}) \quad or \quad (\lambda\vec{x}) \cdot \vec{z} = \vec{x} \cdot (\mu\,\vec{z})$

Since λ, μ are eigenvalues of A, they are both real so that

$$\lambda(\vec{x} \cdot \vec{z}) = \mu(\vec{x} \cdot \vec{z}) \quad or \quad (\lambda - \mu)\,(\vec{x} \cdot \vec{z}) = 0$$

Finally, since $\lambda \neq \mu$, it follows that $(\vec{x} \cdot \vec{z}) = 0.$

These results are part of the following theorem

Spectral Theorem for Symmetric Matrices Suppose A is an n by n symmetric real matrix. Then

i) the eigenvalues of A are all real.

ii) eigenvectors associated with distinct eigenvalues of A are orthogonal

iii) A has n mutually orthogonal eigenvectors $\vec{E}_1, \ldots \vec{E}_n$
which form an orthogonal basis for R^n.

The meaning of iii) is that is $\vec{x}$ is an arbitrary vector in R^n then we can find unique constants $\alpha_1, \ldots, \alpha_n$ such that

$$\vec{x} = \alpha_1\vec{E}_1 + \ldots + \alpha_n\vec{E}_n.$$

In fact, the $\alpha's$ can easily be found by using the orthogonality of the eigenvectors, for each $k,\ 1 \le k \le n,$ $\quad \vec{x} \bullet \vec{E}_k = \left(\alpha_1 \vec{E}_1 + \ldots + \alpha_n \vec{E}_n\right) \bullet \vec{E}_k$

$$= \alpha_1 \vec{E}_1 \bullet \vec{E}_k + \ldots + \alpha_n \vec{E}_n \bullet \vec{E}_k = \alpha_k \vec{E}_k \bullet \vec{E}_k$$

and since the $\vec{E}_k's$ are eigenvectors, $\vec{E}_k \bullet \vec{E}_k \neq 0$ so that

$$\alpha_k = \frac{\vec{x} \bullet \vec{E}_k}{\vec{E}_k \bullet \vec{E}_k} \quad \text{for each } k,\ 1 \le k \le n.$$

The fact that the eigenvectors of A form a basis for R^n is very useful in solving systems of linear constant coefficient differential equations.