characteristic values and vectors (of a matrix)

Canonical name	CharacteristicValuesAndVectorsofAMatrix
Date of creation	2013-03-22 17:43:58
Last modified on	2013-03-22 17:43:58
Owner	perucho (2192)
Last modified by	perucho (2192)
Numerical id	6
Author	perucho (2192)
Entry type	Topic
Classification	msc 15A18
Synonym	eigenvalues
Synonym	eigenvectors

Over the spectrum $\sigma(A)$ of a matrix A, its eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_s$ possess multiplicities $n_1, n_2, \ldots, n_s$, respectively, with $\sum_{k=1}^{s} n_k = n$. Its associated characteristic polynomial is then factored as

$$\Delta(\lambda) \equiv |\lambda I - A| = \Pi_{k=1}^{s} (\lambda - \lambda_k)^{n_k}. \tag{1}$$

Let us set $\mathrm{mult}(\lambda_k) = n_k$ for multiplicity of $\lambda_k (k = 1, \ldots, s)$. We will now prove the following theorem.

Theorem 1. *If $\sigma(A) = \{\lambda_k\}_{k=1}^{s}$, $\mathrm{mult}(\lambda_k) = n_k$, and $g(\mu)$ is a scalar polynomial, then $\sigma(g(A)) = \{g(\lambda_k)\}_{k=1}^{s}$, $\mathrm{mult}(g(\lambda_k)) = n_k$.*

Proof. Let $g(\mu)$ be an arbitrary scalar polynomial. We want to find the characteristic values of $g(A)$. For this purpose we split $g(\mu)$ into linear factors

$$g(\mu) = a_0 \Pi_{i=1}^{t} (\mu - \mu_i)^{l_i}, \qquad a_0 \neq 0, \qquad \sum_{i=1}^{t} l_i = l. \tag{2}$$

On substitution $\mu \mapsto A$, we have

$$g(A) = a_0 \Pi_{i=1}^{t} (A - \mu_i I)^{l_i}, \tag{3}$$

being I the identity matrix. Let us compute the determinant of $g(A)$. (Coefficient a_0 will be powered to n, the order of the square matrix A).

$$\begin{aligned} |g(A)| &= a_0^n \Pi_{i=1}^{t} |(-1)(\mu_i I - A)|^{l_i} = a_0^n \Pi_{i=1}^{t} (-1)^{n l_i} |\mu_i I - A|^{l_i} \\ &= a_0^n (-1)^{n \sum_{i=1}^{t} l_i} \Pi_{i=1}^{t} |\mu_i I - A|^{l_i} = a_0^n (-1)^{nl} \Pi_{i=1}^{t} \Delta(\mu_i)^{l_i} \\ &= a_0^n (-1)^{nl} \Pi_{i=1}^{t} [\Pi_{k=1}^{s} (\mu_i - \lambda_k)^{n_k}]^{l_i}, \end{aligned}$$

because on substitution $\lambda \mapsto \mu_i$ in (1). Next we commute the binomial by introducing $(-1)^{nl}$ into the product signs and also we note that $a_0^n = a_0^{\sum_{k=1}^{s} n_k} = \Pi_{k=1}^{s} a_0^{n_k}$, so that

$$|g(A)| = \Pi_{k=1}^{s} [a_0 \Pi_{i=1}^{t} (\lambda_k - \mu_i)^{l_i}]^{n_k},$$

and we may use (2) for $\mu = \lambda_k$ to obtain

$$|g(A)| = \Pi_{k=1}^{s} g(\lambda_k)^{n_k}. \tag{4}$$

Finally we substitute the polynomial $g(\mu)$ by $\lambda - g(\mu)$, where λ is an arbitrary parameter, getting for (4)

$$\Delta(g(A)) \equiv |\lambda I - g(A)| = \Pi_{k=1}^{s} [\lambda - g(\lambda_k)]^{n_k}. \tag{5}$$

This proves the theorem. $\square$

As an important particular case we have: $\sigma(A^m) = \{\lambda_k^m\}_{k=1}^s$, $(m = 0, 1, \cdots)$, $\mathrm{mult}(\lambda_k) = n_k$.

Connection between the characteristic polynomial $\Delta(\lambda)$ and the adjugate matrix $B(\lambda)$ of A.

As it is well known, the adjugate matrix B of a matrix A there corresponds to the algebraic complement or cofactor matrix of the transpose of A. From this definition we have

$$B(\lambda)(\lambda I - A) = \Delta(\lambda)I \qquad \text{and} \qquad (\lambda I - A)B(\lambda) = \Delta(\lambda)I. \tag{6}$$

Let us suppose $\Delta(\lambda)$ is given by

$$\Delta(\lambda) = \lambda^n - \sum_{k=1}^{n} c_k \lambda^{n-k}. \tag{7}$$

It is clear that the difference $\Delta(\lambda) - \Delta(\mu)$ is divisible by $\lambda - \mu$ without remainder, hence

$$\delta(\lambda, \mu) \equiv \frac{\Delta(\lambda) - \Delta(\mu)}{\lambda - \mu} = \lambda^{n-1} + (\mu - c_1)\lambda^{n-2} + (\mu^2 - c_1\mu - c_2)\lambda^{n-3} + \cdots \tag{8}$$

is a polynomial in λ, μ. If we replace in (8) (λ, μ) by the permutable matrices $(\lambda I, A)$ and recalling that from Cayley-Hamilton theorem $\Delta(A) = 0$, then

$$\delta(\lambda I, A)(\lambda I - A) = \Delta(\lambda)I, \tag{9}$$

which by comparing it with $(6)_1$ we conclude that

$$B(\lambda) = \delta(\lambda I, A) \tag{10}$$

is the desired formula by virtue of the uniqueness of the quotient. Therefore (10) and (8) let to write the adjugate $B(\lambda)$ as the matrix polynomial

$$B(\lambda) = I\lambda^{n-1} + \sum_{k=1}^{n-1} B_k \lambda^{n-k-1}, \tag{11}$$

where ($\mu \mapsto A$ in (8))

$$B_k = A^k - \sum_{i=1}^{k} c_i A^{k-i}, \qquad (k = 1, \ldots, n-1), \tag{12}$$

which can also be obtained from the recurrence equation

$$B_k = AB_{k-1} - c_k I, \qquad (k = 1, \dots, n-1; \quad B_0 = I). \tag{13}$$

What is more,

$$AB_{n-1} - c_n I = 0 \equiv B_n. \tag{14}$$

(13) as well as (14) follow inmediately from $(6)_2$ if we equate the coefficients of equal powers of λ on both sides. Also, if we substitute B_{n-1} from (12), into (14), we get $\Delta(A) = 0$ (Cayley-Hamilton), an implicit consequence of generalized Bézout theorem. On the other hand, by setting $\lambda = 0$ in (7) we obtain $c_n = \Delta(0)/(-1) = |-A|/(-1) = (-1)^{n-1}|A| \neq 0$, whenever A be non- singular. From this and from (14) follow that

$$A^{-1} = \frac{1}{c_n} B_{n-1}. \tag{15}$$

Let now λ_c be a characteristic value of A, then $\Delta(\lambda_c) = 0$ and $(6)_2$ becomes

$$(\lambda_c I - A)B(\lambda_c) = 0. \tag{16}$$

Let us assume that $B(\lambda_c) \neq 0$ and denote by $\mathbf{b}$ an arbitrary non-zero column of this matrix. From (16) we have $(\lambda_c I - A)\mathbf{b} = \mathbf{0}$. That is,

$$A\mathbf{b} = \lambda_c \mathbf{b}. \tag{17}$$

Therefore every non-zero column of $B(\lambda_c)$ determines a characteristic vector corresponding to the characteristic value λ_c. Moreover, if to the characteristic value λ_c there correspond l linearly independent characteristic vectors, $n - l$ will be the rank of $\lambda_c I - A$ and so the rank of $B(\lambda_c)$ does not exceed l. In particular, if only one characteristic vector there corresponds to λ_c, then in $B(\lambda_c)$ the elements of any two columns will be proportional (In such a case $l = 1$, hence the rank of $\lambda_c I - A$ will be $n - 1$).
In conclusion: *If the coefficients of the characteristic polynomial are known, then the adjugate matrix may be found by (10). In addition, if the given matrix A is non-singular, then the inverse matrix A^{-1} can be found from (15). Also if λ_c is a characteristic value of A, the non-zero columns of $B(\lambda_c)$ are characteristc vectors of A for $\lambda = \lambda_c$.*

Example. *We find out the characteristic values and vectors from the matrix*

$$A = \begin{bmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{bmatrix}.$$

From (1),

$$\Delta(\lambda) = |\lambda I - A| = \begin{vmatrix} \lambda - 3 & 3 & -2 \\ 1 & \lambda - 5 & 2 \\ 1 & -3 & \lambda \end{vmatrix} = \lambda^3 - 8\lambda^2 + 20\lambda - 16.$$

Comparing with (7), we have

$$c_1 = 8, \qquad c_2 = -20, \qquad c_3 = 16.$$

Next we use (8),

$$\delta(\lambda, \mu) = \frac{\Delta(\lambda) - \Delta(\mu)}{\lambda - \mu} = \lambda^2 + (\mu - 8)\lambda + \mu^2 - 8\mu + 20,$$

so that from (11)

$$B(\lambda) = \delta(\lambda I, A) = \lambda^2 I + (\underbrace{A - 8I}_{B_1})\lambda + \underbrace{A^2 - 8A + 20I}_{B_2}.$$

We will now evaluate B_1 and B_2 by using (12) and (13), respectively.

$$B_1 = A - 8I = \begin{bmatrix} -5 & -3 & 2 \\ -1 & -3 & -2 \\ -1 & 3 & -8 \end{bmatrix}, \qquad B_2 = AB_1 + 20I = \begin{bmatrix} 6 & 6 & -4 \\ 2 & 2 & 4 \\ 2 & -6 & 12 \end{bmatrix},$$

thus $B(\lambda)$ is

$$B(\lambda) = \begin{bmatrix} \lambda^2 - 5\lambda + 6 & -3\lambda + 6 & 2\lambda - 4 \\ -\lambda + 2 & \lambda^2 - 3\lambda + 2 & -2\lambda + 4 \\ -\lambda + 2 & 3\lambda - 6 & \lambda^2 - 8\lambda + 12 \end{bmatrix}.$$

Also $|A| = 16$ and A^{-1} is obtained from (15), i.e.

$$A^{-1} = \frac{1}{16} B_2 = \frac{1}{8} \begin{bmatrix} 3 & 3 & -2 \\ 1 & 1 & 2 \\ 1 & -3 & 6 \end{bmatrix}.$$

Furthermore,

$$\Delta(\lambda) = (\lambda - 2)^2(\lambda - 4).$$

We notice the eigenvalue $\lambda = 2$ *possesses multiplicity* 2 *and also that all the entries of the adjugate* $B(\lambda)$ *are divisible by the binomial* $\lambda - 2$ *(*$|B(2)| = 0$*, i.e.* $\lambda = 2$ *annihilates it), therefore it can be reduced which makes instructive this problem. Thus,*

$$C(\lambda) = \begin{bmatrix} \lambda - 3 & -3 & 2 \\ -1 & \lambda - 1 & -2 \\ -1 & 3 & \lambda - 6 \end{bmatrix},$$

which for $\lambda = 2$ *it becomes*

$$C(2) = \begin{bmatrix} -1 & -3 & 2 \\ -1 & 1 & -2 \\ -1 & 3 & -4 \end{bmatrix}.$$

From this we get the charactreristic vectors $(1, 1, 1)$ *by multiplying the first colum by* -1*, and also* $(-3, 1, 3)$*, both correponding to* $\lambda = 2$*. Third column is a linear combination of the first two (subtract it). Likewise we find for the another characteristic value* $\lambda = 4$

$$C(4) = \begin{bmatrix} 1 & -3 & 2 \\ -1 & 3 & -2 \\ -1 & 3 & -2 \end{bmatrix},$$

whence we get the eigenvector $(1, -1, -1)$*, being the remaining two columns clearly proportional to the first one.*