P AND D IN $P^{-1}XP = \mathrm{dg}(\lambda_1, \ldots, \lambda_\eta) = D$ AS MATRIX FUNCTIONS OF X

R. F. RINEHART

1. Introduction. Let $\mathfrak{M}_C{}^n$ be the algebra of $n \times n$ matrices over the complex field C, $X = \|x_{rs}\|$ a matrix of $\mathfrak{M}_C{}^n$, and $f(X) = \|f_{rs}(x_{11}, x_{12}, \ldots, x_{nn})\|$ a function with domain and range in $\mathfrak{M}_C{}^n$. If the f_{rs} are differentiable with respect to each of the x_{ij}, on some open set R of $\mathfrak{M}_C{}^n$, then the differential $df(X) = \|df_{rs}(x_{ij})\|$ exists, and, moreover, $f(X)$ is Hausdorff-differentiable (HD) **(1, 3, 7)** i.e. $df(X)$ is expressible in the form

$$df(X) = \sum_{i=1}^{n^2} A_i \, dX \, B_i,$$

where $dX = \|dx_{rs}\|$, and the matrices A_i, B_i are independent of dX. The Hausdorff derivative $f^I(X)$ is defined to be

$$f^I(X) = \sum_{i=1}^{n^2} A_i B_i,$$

i.e. the value of $df(X)$ for $dX = I$, the identity matrix **(2)**.

If $\mathfrak{M}_C{}^n$ is topologized by the topology induced by any suitable matrix norm, then the subset $\mathfrak{K}$ of $\mathfrak{M}_C{}^n$ consisting of matrices with n distinct eigenvalues is an open set in $\mathfrak{M}_C{}^n$ since the eigenvalues of a matrix are continuous, and indeed differentiable, functions of the elements of the matrix. Hence, if $X \in \mathfrak{K}$, and $P^{-1}XP = \mathrm{dg}(\lambda_1, \ldots, \lambda_n) = D$, then D is a Hausdorff-differentiable function of X on $\mathfrak{K}$. Further, as will be shown in §3, the matrix P can be so chosen that P is also a differentiable (therefore HD), function of X on $\mathfrak{K}$.

The purpose of this note is to examine the matrices D and P as HD functions, and in particular, to calculate their Hausdorff derivatives, which turn out to have the interesting properties: $D^I = I$ and $P^I = 0$. The first of these results yields a simple and interesting partial differential equation satisfied by the eigenvalues of matrices $X \in \mathfrak{K}$.

2. The function $D(X) = \mathrm{dg}(\lambda_1, \ldots, \lambda_n)$**.** Let X belong to the space $\mathfrak{K}$ of matrices of $\mathfrak{M}_C{}^n$ with distinct eigenvalues. Then X is similar to a diagonal matrix, $P^{-1}XP = \mathrm{dg}(\lambda_1, \ldots, \lambda_n) = D(X)$, and the λ_i are differentiable functions of the elements of X **(8)**. (The ordering of the λ_i may be fixed by selecting an ordering at some fixed $X_0 \in \mathfrak{K}$. The continuity of the λ_i as functions of

Received August 4, 1965. Presented to the American Mathematical Society, April 12, 1965. This work has been supported by Army Research Office (Durham).

$X \in \mathfrak{R}$ thereby determines the ordering for all $X \in \mathfrak{R}$.) The differential of $D(X)$ is $d(D(X);\, dX) = \mathrm{dg}(d\lambda_1, \ldots, d\lambda_n)$, and the Hausdorff derivative $D^I(X)$ is given by $D^I(X) = d(D(X);\, I)$ **(2)**.

THEOREM 2.1. *$D^I(X) = I$ throughout $\mathfrak{R}$.*

Proof. It has been shown in **(9)** that Hausdorff differentiability on an open set of $\mathfrak{M}_C{}^n$ implies differentiability in the following sense (R-differentiability): for norm H sufficiently small,

(a) $D(X + H) - D(X)$ is expressible in the form

$$\sum_{i=1}^{n^2} M_i H N_i,\ M_i, N_i \in \mathfrak{M}_C{}^n,$$

and

(b) $L = \lim_{H \to 0} \sum M_i N_i$ exists. Further, L is equal to the Hausdorff derivative.

Since L exists, let us choose H in the special form $H = hI$, $h \in C$. Then, since the eigenvalues of $X + hI$ are $\lambda_1 + h, \lambda_2 + h, \ldots, \lambda_n + h$,

$$D(X + hI) - D(X) = hI = HI \text{ and } L = I = D^I(X).$$

Theorem 2.1 yields an interesting universal partial differential equation satisfied by the eigenvalues of matrices with distinct eigenvalues.

COROLLARY 2.1. *Let $X \in \mathfrak{R}$ and let λ_j be any eigenvalue of X. Then*

$$\sum_{i=1}^{n} \frac{\partial \lambda_j}{\partial x_{ii}} = 1.$$

This result follows directly from the fact **(2)** that the Hausdorff derivative of an HD function $f(X)$ on $\mathfrak{M}_C{}^n$ is given by

$$\sum_{i=1}^{n} \frac{\partial f(X)}{\partial x_{ii}}.$$

3. The transforming matrix P as a function of X. Let $X \in \mathfrak{R} \subset \mathfrak{M}_C{}^n$ with $n > 1$ and let Q be a matrix such that

$$Q^{-1} X Q = \mathrm{dg}(\lambda_1, \ldots, \lambda_n) = D.$$

Even for a chosen ordering of the λ_i, the matrix Q is not unique. Column i of Q is an eigenvector of X corresponding to λ_i. However, since X has distinct eigenvalues, the null space of $X - \lambda_i I$ has dimension 1 and all eigenvectors associated with λ_i are scalar multiples of one such eigenvector. Since $X - \lambda_i I$ is of rank $n - 1$, this fixed eigenvector can be taken to be the set of cofactors of the elements of any row of $X - \lambda_i I$, as long as not all of these cofactors are zero. At least one such non-zero row of cofactors exists since $X - \lambda_i I$ is of rank $n - 1$.

To arrive at a "stabilized" choice of the columns of Q, we proceed as follows. For $n = 1$, choose $Q = 1$. For $n > 1$, let X_0 be a fixed matrix of $\mathfrak{K}$. For each $i = 1, \ldots, n$ choose a non-zero column j_i of $(X - \lambda_i I)^A$ as an eigenvector. Let Δ_{j_i} be the set of all matrices X of $\mathfrak{M}_C{}^n$ for which column j_i of $(X - \lambda_i I)^A$ is non-zero. Since the elements of column j_i are polynomials in λ_i and the elements of X, and are hence continuous (indeed differentiable) functions of the elements of X, it follows that Δ_{j_i} is an open set of $\mathfrak{M}_C{}^n$. The set

$$\Delta = \bigcap_{i=1}^{n} \Delta_{j_i}$$

is therefore also an open set that contains X_0 as an interior point. The matrix P whose ith column is column j_i of $(X - \lambda_i I)^A$ is therefore a Q which is a single-valued HD function of X defined in Δ.

THEOREM 3.1. *For any given fixed matrix $X_0 \in \mathfrak{K}$, there exists an open set Δ containing X_0 on which the matrix P such that $P^{-1}XP$ is diagonal can be chosen as an* HD *function of X. Its Hausdorff derivative*

$$P^I = \sum_{i=1}^{n} \frac{\partial P}{\partial x_{ii}}$$

is zero throughout Δ.

The first part of this theorem has already been proved by Portmann **(3)** in less elementary fashion. For purposes of the last part of the theorem, a more explicit formulation of P than that of Portmann is required.

The last assertion is all that remains to be proved. Before taking up its proof, some background observations about Hausdorff derivatives are in order.

For primary functions **(4, 5)**, i.e. those functions arising from the classical extension of scalar functions $f(z)$ of a complex variable to complex matrices, it is true that $f^I(X) = 0$ implies $f(X)$ is a constant (in this case a constant scalar matrix), since $f^I(X)$ in this case is equal to the primary function $f'(X)$, the extension of $f'(z)$ to $\mathfrak{M}_C{}^n$. However, for more general functions, this assertion is false. For example,

$$f\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} x_{12} & 0 \\ 0 & 0 \end{pmatrix}$$

is a non-constant HD function of

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix},$$

whose Hausdorff derivative is zero. Even for the highly restricted class of intrinsic functions **(5, 6)**, i.e. functions that "admit" all automorphisms and anti-automorphisms of $\mathfrak{M}_C{}^n$, which includes the class of primary functions, a non-constant function may have a zero Hausdorff derivative. An example is

$$f\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = 2\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} - (x_{11} + x_{22})I.$$

https://doi.org/10.4153/CJM-1966-083-0 Published online by Cambridge University Press

Thus the class of functions with a zero Hausdorff derivative is not a trivial class.

Proof of Theorem 3.1. For $n = 1$ the last assertion of the theorem is trivially true. For $n > 1$ and for $X \in \Delta$, each element of P is a cofactor X_i^{rs} of some element of some one of the matrices $X - \lambda_i I$. $P^I = 0$ will therefore follow if the relation

$$\sum_{j=1}^{n} \frac{\partial X_i^{rs}}{\partial x_{jj}} = 0$$

is satisfied by every such cofactor. Applying the operator

$$\sum_{j=1}^{n} \frac{\partial}{\partial x_{jj}}$$

to X_i^{rs} viewed as an $(n-1)$-rowed determinant yields a sum of $n - 1$ determinants in each of which one row consists of the operator applied to the elements of that row of X_i^{rs}. The elements of the differentiated row will be either of the form

$$\sum \frac{\partial x_{pq}}{\partial x_{jj}}, \qquad p \neq q,$$

or of the form

$$\sum \frac{\partial}{\partial x_{jj}} (x_{pp} - \lambda_i).$$

The first of these is clearly zero; the second is also zero by virtue of Corollary 2.1. Hence,

$$\sum_{j=1}^{n} \frac{\partial X_i^{rs}}{\partial x_{jj}} = 0 \qquad \text{for all } i, r, s$$

and hence $P^I = 0$.

Remarks on the stabilization of P. The process described for the construction of a single-valued P does not define a unique function P or a unique Δ containing X_0. A different selection of columns j_i could yield a different single-valued function P with a different Δ. Further, any selection of P could be modified by multiplying its ith column by a differentiable function of the elements of $X - \lambda_i I$ (with a possible associated restriction on Δ_{j_i} to ensure non-vanishing of the multiplying function). For any of these variants, Theorem 3.1 remains valid.

4. The case of real argument matrices. Let $\mathfrak{M}_R^n$ be the set of real $n \times n$ matrices, and let $f(X)$ be a function with domain in $\mathfrak{M}_R^n$ and with range in $\mathfrak{M}_C^n$. The elements f_{rs} of $f(X)$ are complex-valued functions of the n^2 real variables x_{ij}. If the f_{rs} are differentiable functions of the x_{ij} in the domain of f, then the concepts of Hausdorff differentiability and Hausdorff derivative of **(1, 2, 7, 9)** are directly extensible to such functions. The same is true of the generalized pointwise difference quotient definition of derivative in **(9)** and of the equality of the two types of derivative when the Hausdorff derivative exists in a neighbourhood of the point (9).

Hence if $\mathfrak{R}_R$ denotes the open set of matrices $X \in \mathfrak{M}_R{}^n$ with distinct eigenvalues, the results of §2 remain valid, since the λ_i, being differentiable with respect to the x_{rs} for X belonging to $\mathfrak{R}$, are differentiable with respect to the x_{rs} for $X \in \mathfrak{R}_R$.

A significant difference occurs, however, in the case of the transforming matrix P, for P may now be defined globally over $\mathfrak{R}_R$ by choosing column i of P as a unit eigenvector of X corresponding to λ_i. This amounts to choosing a non-zero column of $(X - \lambda_i I)^A$ normalized to hermitian length 1. Since the columns of $(X - \lambda_i I)^A$ are proportional, the same vector (except possibly for sign) will be obtained no matter which non-zero vector of $(X - \lambda_i I)^A$ is selected. The sign ambiguity in selecting the normalized column i of P can be removed by making the selection for some fixed X_0 and thereafter selecting the sign at other X as dictated by the continuity of the elements of P.

The elements of the normalized column of P will be of the form

$$p_{rs} = X_i^{rs}\left[\sum_{s=1}^{n} X_i^{rs}\bar{X}_i^{rs}\right]^{-\frac{1}{2}}.$$

Since $\bar{X}_i^{rs}$ is equal to the corresponding cofactor of $X - \bar{\lambda}_i I$, and since $\bar{\lambda}_i$, being an eigenvalue of X, is also a differentiable function of the x_{rs}, the quantity in brackets will be a differentiable function of the x_{rs}. Further, the argument employed in the proof of Theorem 3.1 shows that

$$\sum_{j=1}^{n} \frac{\partial}{\partial x_{jj}} X_i^{rs}\bar{X}_i^{rs} = \left(\sum_j \frac{\partial}{\partial x_{jj}} X_i^{rs}\right)\bar{X}_i^{rs} + X_i^{rs}\left(\sum_j \frac{\partial}{\partial x_{jj}} \bar{X}_i^{rs}\right) = 0,$$

for all r, s, i. Hence it follows that

$$\sum_{j=1}^{n} \frac{\partial p_{rs}}{\partial x_{jj}} = 0$$

throughout $\mathfrak{R}_R$. This yields

THEOREM 4.1. *Let X belong to the open set $\mathfrak{R}_R$ of real $n \times n$ matrices with distinct eigenvalues. Let P be a matrix whose columns are unit complex vectors such that $P^{-1}XP$ is diagonal. With the sign of each column of P appropriately chosen, P is an* HD *function of X defined throughout $\mathfrak{R}_R$, and its Hausdorff derivative*

$$P^I = \sum_{j=1}^{n} \frac{\partial P}{\partial x_{jj}}$$

is zero throughout $\mathfrak{R}_R$.

In the case of the complex argument matrix domain of §3, this normalization would not yield an eigenvector whose elements are differentiable functions of the elements of X (essentially because $z\bar{z}$ is not a differentiable function of the complex variable z). For this reason, this otherwise attractive normalization had to be eschewed there.

REFERENCES

1. F. Hausdorff, *Zur Theorie der Systeme complexer Zahlen*, Ber. Sächs. Akad. Wiss. Leipzig, *52* (1900), 43–61.
2. W. O. Portmann, *A derivative for Hausdorff-analytic functions*, Proc. Amer. Math. Soc., *10* (1959), 101–105.
3. ——— *Hausdorff-analytic functions of matrices*, Proc. Amer. Math. Soc., *11* (1960), 97–101.
4. R. F. Rinehart, *The equivalence of definitions of a matric function*, Amer. Math. Monthly, *62* (1955), 395–414.
5. ——— *Intrinsic functions on matrices*, Duke Math. J., *28* (1961), 291–300.
6. ——— *Differentiable intrinsic functions of complex matrices*, Proc. Amer. Math. Soc., *12* (1961), 565–573.
7. F. Ringleb, *Beiträge zur Funktionentheorie in hyperkomplexen Systemen*, Rend. Circ. Mat. Palermo, *57* (1933), 311–340.
8. E. J. Townsend, *Functions of a complex variable* (New York, 1915).
9. J. C. Wilson and R. F. Rinehart, *Two types of differentiability of functions on algebras*, Rend. Circ. mat. Palermo, *11* (1962), 204–216.

U.S. Naval Postgraduate School,
Monterey, California