Fréchet Derivative on the Finite Dimensional Banach Space

By

Kim, Beyngjun

Department of Mathematics
Graduate School of Education
Cheju National University

Supervised By

Assistant Prof. Han, Chulsoon

June, 1981

Fréchet Derivative on the Finite Dimensional Banach Space

이를 教育学碩士学位 論文으로 提出함

済州大学 教育大学院 数学教育専攻

提 出 者　金　秉　俊

指導教授　韓　哲　淳

1981 年　6月　日

金秉俊의 碩士學位 論文을 認准함

濟 州 大 学　教 育 大 学 院

主 審 ________________ ㊞

副 審 ________________ ㊞

副 審 ________________ ㊞

1 9 8 1 年　6月　　日

감 사 의 글

이 논문이 완성되기까지 바쁘신 가운데도 지도하여 주신 한철순 교수님께 감사드리며, 그동안 물심양면으로 도움을 주셨던 류근식, 현진오 교수님께 진심으로 감사드립니다.

또한 그동안 저에게 사랑과 격려를 하여 주신 주위의 많은 분들께 감사드립니다.

1981년 6월 일

김 병 준

CONTENTS.

국 문 요 약

제 목 : 유한차원 Banach 공간위에서 Fréchet 미분

김 병 준
수학교육전공
제주대학교육대학원

유한차원 normed vector space V 에서 유한차원 normed vector space W 로 보내는 Fréchet 미분에 대한 약간의 성질을 찾는 것이다.

정리 (2 . 1) 은 함수 $f:S \longrightarrow W$ 에 대해, $f_j = w_j^* \circ f$ 라 정의할때 f가 점 a에서 미분가능 하기위한 필요충분 조건은 정수 j에 대해 f_j가 미분 가능하다.

정리 (2 . 4) $V=R^n$ 과 $W=R^m$ 라 가정하고 $f_j = w_j^* \circ f$ 를 $f(x) = \sum_{j=1}^{m} f_j(x) \circ u_j$ 에 의해서 정의된 실 함수라 하자.

또한 f가 한점 a에서 미분 가능하면 $f_1, f_2, \ldots, f_n$ 은 한 점에서 1차 편미분이 되고 $(Df)_a$ 를 나타내는 행렬은

$$\begin{pmatrix} (D_1 f_1)_a & \cdots & \vdots & \cdots & (D_n f_1)_a \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ (D_1 f_m)_a & \cdots & \vdots & \cdots & (D_n f_m)_a \end{pmatrix}$$

이다.

0. INTRODUCTION.

We shall be concerned with a generalization, due to M. Fréchet(1925), of the classical differential calculus of real-valued functions of a real variable. We recall that a real-valued function f on R has derivative m at a point a of R if and only if for each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left| \frac{f(x) - f(a)}{x - a} - m \right| \leq \varepsilon \qquad \text{------------} \quad (*)$$

whenever $0 < |x-a| < \delta$. The inequality (*) can be replaced by the equivalent inequality

$$|f(x) - f(a) - m(x-a)| \leq \varepsilon |x-a| \qquad \text{------------} \quad (**)$$

whenever $|x - a| < \delta$.

Fréchet's generalization of the differential calculus applies to the mapping of a real normed vector space V into a real normed vector space W. Let f be such a mapping. The derivative of f at a point a of V will be defined to be a linear transformation T of V into W which satisfies the inequality

$$\|f(x) - f(a) - T(x-a)\| \leq \varepsilon \|x-a\|$$

whenever $\|x-a\| < \delta$. It is obvious that (***) is a generalization of (**): and the analogy between(***) and (**) becomes clearer when we remark that the mapping $y \longrightarrow my$ is a linear mapping of R into itself.

The purpose of the present paper is to find the Fréchet derivative for a function f of a finite dimensional normed vector space V into a finite dimensional vector space W and investigate its some definitions and properties.

Our paper will be diviede into 2 sections. In the first section, we introduce some definitions and notation which are needed in our further consideration.

In final section, we investigate the Frechet derivative for a function of V into W and its some properties.

1. PRELIMINARES.

In this section we establish basic terminology and recall certain known results relevant to our discussion. We omit the proofs of most of them, which have already been known. The following notation will be used throughout the present paper:

R is a set of all real numbers.

$\langle V, \| \ \| \rangle$ is a n-dimensional normed vector space over R.

$\langle W, \| \ \| \rangle$ is a m-dimensional normed vector space over R.

$\{ v_i \mid i=1,2,,,n \}$ is a basis of V.

$\{ w_i \mid i=1,2,,,m \}$ is a basis of W.

V* and W* are dual space of V and W respectively

$\{ v_i^* \mid i=1,2,,,n \}$ is a basis of V*.

$\{ w_i^* \mid i=1,2,,,m \}$ is a basis of W*.

S is a non-empty open subset of V.

Definition(1,1) Let f be a mapping of S into W. Then the mapping f is said to be differentiable at a point $a \in S$ if and only if there is a linear transformation T of V into W which satisfies the following condition: for each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\| f(x) - f(a) - T(x-a) \| \leqq \varepsilon \| x-a \| \quad \text{--------}(*)$$

for all $x \in S$ which $\| x-a \| < \delta$.

By the above definition, we obtained following properties.

Proposition(1,2) If f is differentiable at a point $a \in S$, then there is a unique linear transformation of V and W which satsfies condition (*).

This linear transformation is called the derivative fo f at a and is denoted by $(Df)_a$ is bounded.

Proposition(1,3) Let $y \in W$ and suppose that $f(x)=y$ for all $x \in V$. Then $(Df)_a = 0$ for all $a \in V$.

Proposition(1,4) Let T be a linear transformation of V and W. Then $(Df)_a = T$ for all $a \in V$.

Proposition(1,5) (Linearity)

Let f and G be mapping of S into W that are differentiable at a in S and let α, β in R. Then $h = \alpha f + \beta g$ is differentiable at a in S and $(Dh)_a = \alpha (Df)_a + \beta (Dg)_a$.

Proposition(1,6) (Chain rule)

Let f be a mapping of S into an open subset T of W and g be a mapping of T into a normed linear space U. Suppose that f is differentiable a in S and that g is differentiable at $b=f(a)$ in T. Then $g\circ f$ is differentiable at a and $(D(g\circ f)_a) = (Dg)_b \circ (Df)_a$.

Proposition(1,7) Let L(V,W) be the set of all linear transformation of the vector space V and W. Then L(V,W) is a vector space over R. For $T\in L(V,W)$, define the norm $\|T\|$ of T to be the sup of all numbers $|T(x)|$, where x ranges over all x in with $|x|\leq 1$.
Then $\|T(x)\| \leq \|T\||x|$.

Proposition(1,8) Two vector spaces V and R^n are isomorphic.

Proposition(1,9) Let φ be an isomorphism of a finite vector space, then matrix $M(\varphi)$ is invertible where $M(\varphi)$ is a matrix represented by isomorphism .

Proposition(1,10) The matrix of a composite linear transformation $T\circ S$ is the product of the matrices of the factors:

$$M(T\circ S) = M(T)M(S)$$

2. MAIN THEOREM.

The purpose of this section is to find the Fréchet derivate for a function S into W and we investigate its some properties.

Proposition(2,1) For $f:S \longrightarrow W$ a function, we define $f_j = w_j^* \circ f$. Then f is differentiable at $a \in S$ if and only if f_j are differentiable at a for all j. Moreover $(Df)_a = \sum_{j=1}^{m} w_j \circ (Df_j)_a$.

(Proof) Since $w_j^*: W \longrightarrow R$ is a linear transformation, by proposition(1,4) w_j^* is differentiable and $(Dw_j^*)_a = w_j^*$. Suppose f is differentiable at a, by proposition(1,6) $w_j^* \circ f$ is differentiable at a for all j. Hence f_j is differentiable at a for all j and $(Df_j)_a = w_j^* \circ (Df)_a$.

Conversely, suppose that f_j are differentiable at a for $1 \leq j \leq m$. Then for $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\| f_j(x) - f_j(a) - (Df_j)_a(x-a) \| \leq \frac{\varepsilon}{m} \| x-a \|$$

whenever $\| x-a \| < \delta_j$.

Since $\sum_{j=1}^{m} w_j \circ w_j^* = 1_w$, $\sum_{j=1}^{m} w_j \circ f_j = f$ [3] and $\sum_{j=1}^{m} w_j \circ (Df_j)_a = (Df)_a$.

Hence $\left\| f(x)-f(a)-\sum_{j=1}^{m}(w_j\ (Df_j)_a(x-a)\right\|$

$$=\left\| \sum_{j=1}^{m} w_j\circ f_j(x)-\sum_{j=1}^{m} w_j\circ f_j(a)-\sum_{j=1}^{m} w_j\circ (Df_j)_a(x-a)\right\|$$

$$=\left\| \sum_{j=1}^{m} w_j(f_j(x)-f_j(a)-(Df_j)_a(x-a))\right\|$$

$$\leqq \sum_{j=1}^{m}\|w_j\|\ \|f_j(x)-f_j(a)-(df_j)_a(x-a)\|$$

$$\leqq \varepsilon\|x-a\|$$

Therefore, f is differentiable at a and $(Df)_a$

$$= \sum_{j=1}^{m} w_j\circ (Df_j)_a .$$ (Q. E. D)

Proposition(2,2) Suppose $\rho : V \longrightarrow R^n$ and $\varphi : W \longrightarrow R^m$ are isomorphisms. Then for each function $f:V \longrightarrow W$, there exists only one function $g:R^n \longrightarrow R^m$ such that $g=\varphi\circ f\circ\rho^{-1}$. Conversely for each function $g : R^n \longrightarrow R^m$, there exists only one function $f:V \longrightarrow W$ such that $f=\varphi^{-1}\circ g\circ\rho$ this is

$$\begin{array}{ccc} V & \xrightarrow{\ f\ } & W \\ \rho\,\|\wr & \circlearrowright & \varphi\,\|\wr \\ R^n & \xrightarrow{\ g\ } & R^m \end{array}$$

From the above proposition(2,2), we obtained the following; If f is differentiable at a, then g is differentiable at $\rho(a)$ and $(Dg)_{\rho(a)}= \varphi(Df)_a\rho^{-1}$, and if g is differentiable at b, then f is differentiable at $\rho^{-1}(b)$ and $(Df)_{\rho^{-1}(b)}=(Dg)_b\rho$.

Definition(2,3) Let $\{e_i \mid i=1,,,n\}$ and $\{u_i \mid i=1,,,m\}$ be the standard bases of R^n and R^m respectively and f a map R^n into R^m where $f(x)=\sum_{j=1}^{m} u_i \circ f_j(x)$, then T is called the partial derivative of f_j at a: if for each $\varepsilon>0$, there exists $\delta>0$ such that

$$\|f_j(a+te_j) - f(a) - T(te_i)\| \leqq \varepsilon \|te_i\|$$

whenever $|t|<\delta$ and t is real.

Here, we write $T=(D_j f_j)$

Proposition(2,4) Suppose $V=R^n$ and $W=R^m$ and $f_j=w_j^* \circ f$ $(j=1,,,m)$ be the real valued function on non-empty open set S of R^n defined by $f(x)=\sum_{j=1}^{m} w_j \circ f_j(x)$ for all $x \in S$. Suppose also that f is differentiable at a point $a \in S$. Then $f_1, f_2,,, f_m$ have first partial derivatives at a and the matrix which represents $(Df)_a$ is

$$\begin{pmatrix} (D_1 f_1)_a & (D_2 f_1)_a & \vdots & (D_n f_1)_a \\ (D_1 f_2)_a & (D_2 f_2)_a & \vdots & (D_n f_2)_a \\ \cdots & \cdots & \cdots & \cdots \\ (D_1 f_m)_a & (D_2 f_m)_a & \vdots & (D_n f_m)_a \end{pmatrix}$$

(Proof) Suppose first that m=1. Let $\sum_{j=1}^{n} \zeta_j = \sum_{j=1}^{n} w_j \circ (Df_j)_a$ be the matrix representing $(Df)_a$. Then

$$(Df)_a(x)= \sum_{j=1}^{n}(w_j\circ(Df_j)_a)(\xi_j)= \sum_{j=1}^{n}\tau_j\xi_j \qquad ---(*1)$$

for all $x=(\xi_1,\xi_2,,,\xi_n)\in R^n$

and $\quad w_j\circ(Df)_a(x)=(Df_j)_a(\xi_j)$.

Given $\varepsilon>0$, there exists $\delta>0$ such that

$$|f(x)-f(a)-(Df)_a(x-a)|\leqq \varepsilon\|x-a\|$$

for $\|x-a\|<\delta$ and so (*1) gives

$$|f(x)-f(a)- \sum_{j=1}^{n}(w_j\circ(Df_j)_a)(\xi_j-\alpha_j)|$$

$$=|f(x)-f(a)- \sum_{j=1}^{n}\tau_j(\xi_j-\alpha_j)|\leqq \varepsilon\|x-a\| \qquad ----(*2)$$

for $\|x-a\|<\delta$, where $a=(\alpha_1,,,\alpha_n)$. Let $1\leqq k\leqq n$, $t\in R$

and $x=(\xi_{1k},,,\xi_{nk})$ where $\xi_{jk}=\alpha_j$ for $j\neq k$ and

$\xi_{kk}=t$, then $\|x-a\| = |t-\alpha_k|$, and so if $|t-\alpha_k|<\delta$,

we obtain from (*2)

$$|f(x_k)-f(a)-\tau_k(t-\alpha_k)|\leqq \varepsilon|t-\alpha_k|$$

Consequently

$$\left|\frac{f(\alpha_1,,\alpha_{k-1},t,\alpha_{k+1},,\alpha_n)-f(\alpha_1,,\alpha_{k-1},\alpha_k,\alpha_{k+1},,\alpha_n)}{t-\alpha_k}-\tau_k\right|\leqq\varepsilon$$

for $0<|t-\alpha_k|<\delta$. This shows that f is differentiable with respect to kth variable at a, and that $(d_kf)_a=\tau_k=w_k\circ(Df_k)_a$. We have now proved the theorem in the case when $m = 1$.

Consider now to general case. By proposition(2,1) the real-valued function $f_1, f_2,,, f_m$ are (Frechet) differentiable at a and therefore, by what we have proved above, f_j is differentiable with respect to the kth variable at a for j=1,2,,,,m and k=1,2,,,n.

It remains only to identify the matrix $(\tau_{jk}) = w_j \circ (D_k f_j)_a$ which represents $(Df)_a$.

Since

$$(Df)_a(x) = \left(\sum_{k=1}^{n} \tau_{1k}\xi_k,,,,,, \sum_{k=1}^{n} \tau_{mk}\xi_k \right) \quad ---(*3)$$

for all $x = (\xi_1, \xi_2,,, \xi_n) \in R^n$.

Also proposition(2,1) gives

$$(Df)_a(x) = \left(\sum_{j=1}^{m} w_j \circ (Df_j)_a \right)(x)$$
$$= \sum_{j=1}^{m} \sum_{k=1}^{n} w_j \circ (D_k f_j)_a (\xi_k) \quad -----(*4)$$

for all $x \in R^n$. Finally by (*1) and the first part of the proof we have

$$w_j \circ (Df_j)_a(x) = \sum_{k=1}^{n} w_j \circ (D_k f_j)_a(\xi_k). \quad ------(*5)$$

for all $x = (\xi_1, \xi_2,,, \xi_n) \in R^n$.

From (*3), (*4) and (*5) we obtain

$$\sum_{k=1}^{n} \tau_{jk}\xi_k = w_j \circ (Df_j)_a(x) = \sum_{k=1}^{n} w_j \circ (D_k f_j)_a(\xi_k)$$

for j = 1,2,,,m amd all $x = (\xi_1,,, \xi_n)$ R^n.

Consequently, $\tau_{jk} = w_j \circ (D_k f_j)_a$ for j=1,2,,,m and k = 1,2,,,n .

(Q. E. D.)

Definition(2,5) Two matrices M_1 and M_2 are equivalant if and only if there are invertible square matrices P and Q with $M_2 = QM_1P^{-1}$.

Proposition(2,6) Suppose $f:V \longrightarrow W$ is a function. Suppose $\rho:V \longrightarrow R^n$ and $\varphi:W \longrightarrow R^m$ are isomorphism with $g=\varphi \circ f \circ \rho^{-1}$. Then if is differentiable at a, $(Dg)_{\rho(a)} = \varphi(Df)_a\rho^{-1}$ and two matrices $M((Dg)_{\rho(a)})$ and $M((Df)_a)$ are equivalent.

(Proof) Since $(Dg)_{\rho(a)}, \varphi, \rho$ are linear transformation, by the proposition(1,10)

$$M((Dg)_{\rho(a)}) = M(\varphi)M((Df)_a)M(\rho^{-1})$$

and, by the proposition(1,8), $M(\varphi)$ and $M(\rho)$ are invertible matrices. Hence $M((Dg)_{\rho(a)})$ and $M((Df)_a)$ are equivalent. (Q, E, D)

Here, if

$$\rho\left(\sum_{i=1}^{n} \alpha_i v_i\right)=\sum_{i=1}^{n} \alpha_i e_i \quad \text{and} \quad \left(\sum_{i=1}^{m} \beta_i w_i\right)= \sum_{i=1}^{m} \beta_i u_i ,$$

then $M(\rho)$ and $M(\varphi)$ are identify matrices.

Hence $M((Df)_a) = M((Dh)_{\rho(a)}) = M((D_jf_i)_a)$

where $1 \leqq i \leqq n$, $1 \leqq j \leqq m$.

Corollary(2,7) In the above proposition(2,2), if

$$\rho(v_i) = e_i,\ \varphi(w_j) = u_j \text{ for } 1 \leqq i \leqq n,\ 1 \leqq j \leqq m,$$

$$M((Df)_a) \doteq M((Dg)_a).$$

3. REFERENCES.

1. Principles of Mathematical Analysis: Walter Rudin's Third Edition, International Student Edition 1976.

2. Mathematical Analysis: Apostol's Sencond Edition, Addision-Wessley, Publishing Company 1974.

3. Algebra: Saunders Maclane and Garrett Birkhoff, The Macmillan Company 1965.

4. Rovert G. Bartle " The Element of Real Analysis Second Edition " John Wiley & Sonc, Inc. New York 1976.

5. Brown & Page " Elements of Functional Analysis " Van Nostrand Reihold, London. 1970.

6. The Frechet derivative on the Banach Algebra: K. P. Hong. YON SEI University 1979.

7. The Frechet derivative for the rational function on the Banach Algbra : K. S. Ryu, K. P. Hong and Y. H. Par. Che Ju university Journal 13, 1981 (Unpublished)

ABSTRACT.

Frechet derivative on the finite dimensional Banach Space

Kim, Beyng Jun
Department of Mathematics
Graduate School of Education
Che Ju National University

We investigate some properties of Frechet derivative on the finite dimensional normed vector space V into the finite dimensional normed vector space W.

Here, proposition(2,1): if $f:S \longrightarrow W$, define $f_j = w_j^* \circ f$, then f is differentiable at a in S if and only if f_j are differentiable at a for $1 \leqq j \leqq m$.

And also, proposition(2,4): if $V = R^n$ and $W = R^m$, let $f_j = w_j^* \circ f$ is a real-valued function defined by $f(x) = \sum_{j=1}^{m} f_j(x) \circ u_j$, then, if f is differentiable at a, $f_1, f_2, ,,, f_n$ have first partial derivative at a and the matrix which represents is

$$\begin{pmatrix} (D_1 f_1)_a & \cdots\cdots & (D_n f_1)_a \\ \cdots & \cdots\cdots\cdots & \cdots \\ (D_1 f_m)_a & \cdots\cdots & (D_n f_m)_a \end{pmatrix}$$