Additional notes on Fréchet derivatives

(To accompany Section 3.10 of the AMATH 731 Course Notes)

Let $X, Y$ be normed linear spaces. The Fréchet derivative of an operator $F : X \to Y$ is the bounded linear operator $DF(a) : X \to Y$ which satisfies the following relation,

$$
\lim_{h \to 0} \frac{\|F(a + h) - F(a) - DF(a)h\|}{\|h\|} = 0.
$$

It is a generalization of the derivative of a function $f : \mathbb{R} \to \mathbb{R}$ encountered in first-year calculus and the Jacobian of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ studied in advanced calculus.

Indeed, for functions $f : \mathbb{R} \to \mathbb{R}$, the connection is clear if we go back to the definition of $f'(a)$:

$$
f'(a) = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h}.
$$

We may rewrite this relation as

$$
\lim_{h \to 0} \frac{|f(a + h) - f(a) - f'(a)h|}{|h|} = 0.
$$

The Fréchet derivative of $f$ is the scalar $f'(a)$, which multiplies the scalar $a \in \mathbb{R}$ – as such, $f'(a)$ is a linear operator in $\mathbb{R}$.

For functions $F : \mathbb{R}^n \to \mathbb{R}^m$, the Fréchet derivative $DF(a)$ is the Jacobian of $F$, a linear operator which is represented by an $m \times n$ matrix, as written in Example 3.12 of the Course Notes and reproduced below,

$$
DF(a) = \begin{bmatrix}
\frac{\partial F_1}{\partial x_1}(a) & \cdots & \frac{\partial F_1}{\partial x_n}(a) \\
\vdots & \ddots & \vdots \\
\frac{\partial F_m}{\partial x_1}(a) & \cdots & \frac{\partial F_m}{\partial x_n}(a)
\end{bmatrix}.
$$

Here, the rate of change of $F : \mathbb{R}^m \to \mathbb{R}^n$ in the direction $h \in \mathbb{R}^m$ is measured at the point $a \in \mathbb{R}^m$. In fact, the term,

$$
\frac{DF(a)h}{\|h\|} = DF(a)\hat{h},
$$

in Eq. (1) is, by definition, the directional derivative of $F$ at $a$.

The Fréchet derivative, as defined in Eq. (1) extends the above concepts of the derivative to operators in general normed spaces, for example, infinite-dimensional function spaces. This is of great importance to computational methods for solving nonlinear operator equations.

We consider a few examples below. In all cases, it is best to employ the formal definition in Eq. (1). In the analysis of an operator $F : X \to Y$, the usual procedure is to examine the difference $F(a + h) - F(a)$. All terms that are linear in $h$ (and possibly its derivatives) will comprise the Fréchet derivative. Higher-order terms in $h$ (and derivatives) will comprise a remainder term, i.e.

$$
F(a + h) - F(a) = Lh + R(a, h),
$$

where $L$ is a linear operator. (It may be, for example, an integral operator or a differential operator, or an expression involving both.) From Eq. (1), it then remains to show that

$$
\lim_{h \to 0} \frac{\|R(a, h)\|}{\|h\|} = 0.
$$
If this can be done, then the linear operator \( L \) is the Fréchet derivative of \( F \), i.e.,

\[
L = DF(a)
\]  

(8)

**Example 1:** Let \( X = Y = C[a, b] \) with \( \| \cdot \|_\infty \) norm and let \( T : X \to X \) be the linear integral operator defined by

\[
(Tu)(x) = \int_a^b K(x, s)u(s) \, ds,
\]

(9)

where \( K(x, s) \) is continuous on \([a, b] \times [a, b]\).

We first calculate \( T(u + h) - T(u) \) for an arbitrary \( h \in X \):

\[
[T(u + h) - T(u)](x) = \int_a^b K(x, s)[u(s) + h(s)] \, ds - \int_a^b K(x, s)u(s) \, ds
\]

\[
= \int_a^b K(x, s)[u(s) + h(s) - u(s)] \, ds
\]

\[
= \int_a^b K(x, s)h(s) \, ds.
\]

(10)

Note that the final term is a linear operator on \( h \), which may not have been unexpected – after all, \( T \) is a linear operator. But let us go through the formalities. We may rearrange the above result to read

\[
\frac{1}{\|h\|} \left[ T(u + h) - T(u) - \int_a^b K(x, s)h(s) \, ds \right] = 0.
\]

(11)

Since this equation is true for all \( h \neq 0 \), it follows that the relation in Eq. (1) is satisfied. Therefore, the Fréchet derivative is given by

\[
DT(u) = \int_a^b K(x, s)h(s) \, ds,
\]

(12)

independent of \( u \), i.e., the bounded linear operator \( T \) itself. This illustrates Proposition 3.8, p. 46 of the Course Notes: The Fréchet derivative of a bounded linear operator \( L \) is \( L \) itself.

**Example 2:** As before, let \( X = Y = C[a, b] \) with \( \| \cdot \|_\infty \) norm. Now let \( T : X \to X \) be the nonlinear integral operator defined by

\[
(Tu)(x) = u(x) \int_a^b K(x, s)u(s) \, ds,
\]

(13)

where \( K(x, s) \) is continuous on \([a, b] \times [a, b]\).

Once again, we calculate \( T(u + h) - T(u) \) for an arbitrary \( h \in X \):

\[
[T(u + h) - T(u)](x) = [u(x) + h(x)] \int_a^b K(x, s)[u(s) + h(s)] \, ds - u(x) \int_a^b K(x, s)u(s) \, ds
\]

\[
= u(x) \int_a^b K(x, s)h(s) \, ds + h(x) \int_a^b K(x, s)u(s) \, ds + R(u, h)(x),
\]

(14)

where

\[
R(u, h)(x) = h(x) \int_a^b K(x, s)h(s) \, ds.
\]

(15)
Note that the remainder term \( R(u, h) \) is nonlinear in \( h \). If \( \| R(u, h) \| / \| h \| \to 0 \) as \( h \to 0 \), then the first two terms in the last line of (14) will define the Fréchet derivative of \( T \). We have

\[
\| R(u, h) \| = \max_{x \in [a,b]} \left| h(x) \int_a^b K(x, s)h(s) \, ds \right| \leq M \| h \|^2. \tag{16}
\]

where

\[
M = (b - a) \max_{[a,b] \times [a,b]} |K(x, s)|. \tag{17}
\]

Thus, the Fréchet derivative of \( T \) is given by

\[
[DT(u)]h(x) = u(x) \int_a^b K(x, s)h(s) \, ds + h(x) \int_a^b K(x, s)u(s) \, ds. \tag{18}
\]

Note that it is a linear operator on \( h \). It is also bounded (why?).

**Example 3:** Let \( X = C^1_0[0,1] \) be the space of all \( C^1 \) functions on \([0,1]\) which vanish at the endpoints, with norm

\[
\| u \| = \left[ \int_0^1 \left( u^2 + (u')^2 \right) \, dx \right]^{1/2}. \tag{19}
\]

This norm is often called the energy norm – we shall study it later in this course.

Now consider the operator \( K : X \to \mathbb{R} \) defined by

\[
K(u) = \int_0^1 \left( u^3 + (u')^2 \right) \, dx. \tag{20}
\]

The operator \( K \) is called a functional – it is a real-valued mapping of the space \( X \). The goal is to compute the Fréchet derivative of \( K \).

After a little calculation, we find that

\[
K(u + h) - K(u) = \int_0^1 \left[ 3u^2h + 2u'h \right] \, dx + R(u, h), \tag{21}
\]

where

\[
R(u, h) = \int_0^1 \left[ 3uh^2 + h^3 + (h')^2 \right] \, dx. \tag{22}
\]

Note that, once again, the RHS of Eq. (21) has been arranged so that the first term includes all terms that are linear in \( h \), whereas the remainder, \( R(u, h) \), includes all terms that are nonlinear in \( h \). We suspect that the first term represents the Fréchet derivative, but in order to prove this we must show that \( \| R(u, h) \| / \| h \| \to 0 \) as \( \| h \| \to 0 \). This is, however, somewhat complicated with the energy norm selected for this problem.

In an effort to express \( \| R(u, h) \| \) in terms of \( \| h \| \), we try the following:

\[
\| R(u, h) \| = |R(u, h)| \leq 3 \max_{[0,1]} |u(x)| \int_0^1 h^2 \, dx + \max_{[0,1]} |h(x)| \int_0^1 h^2 \, dx + \int_0^1 (h')^2 \, dx. \tag{23}
\]

Now note, from the definition of the energy norm in Eq. (19), that

\[
\int_0^1 h^2 \, dx \leq \| h \|^2, \quad \int_0^1 (h')^2 \, dx \leq \| h \|^2. \tag{24}
\]

We use this result in Eq. (23):

\[
\| R(u, h) \| \leq (3 \| u(x) \|_{\infty} + \| h(x) \|_{\infty} + 1) \| h \|^2. \tag{25}
\]
A simple rearrangement yields

\[ \frac{\|R(u, h)\|}{\|h\|} \leq (3\|u(x)\|_\infty + \|h(x)\|_\infty + 1) \|h\|. \]  (26)

It is now tempting to let \( \|h\| \to 0 \) and conclude that the ratio on the left vanishes in this limit, but there is one complication: Can we guarantee that \( h(x) \) is bounded, so that the middle term on the right hand side does not “blow up”?

In fact, \( h(x) \) must be bounded, since it is a \( C^1 \) function on \([a, b]\), i.e., there exists an \( M > 0 \) such that \( |h(x)| \leq M \). But for each \( h \), there is an \( M \) – what is necessary is to connect \( M \) with \( \|h\| \). This is made possible with the following result.

**Lemma:** If \( h \in C^1[0, 1] \) and \( h(0) = 0 \), then

\[ \|h\|_\infty = \max_{[0, 1]} |h(x)| \leq 2 \left[ \int_0^1 (h'(s))^2 \, ds \right]^{1/2}. \]  (27)

**Proof:** If \( h = 0 \) on \([0, 1]\), then the equality is satisfied. We now consider the case that \( h \) does not vanish identically over \([0, 1]\). From the Fundamental Theorem of Calculus,

\[ \int_0^x h(s)h'(s) \, ds = \frac{1}{2}h(x)^2 - \frac{1}{2}h(0)^2 = \frac{1}{2}h(x)^2. \]  (28)

Applying the Cauchy-Schwartz inequality to the integral on the left yields

\[ \frac{1}{2}h(x)^2 \leq \left[ \int_0^x h(s)^2 \, ds \right]^{1/2} \left[ \int_0^x (h'(s))^2 \, ds \right]^{1/2}. \]  (29)

Thus

\begin{align*}
\frac{1}{2}h(x)^2 &\leq 2 \left[ \int_0^1 h(s)^2 \, ds \right]^{1/2} \left[ \int_0^1 (h'(s))^2 \, ds \right]^{1/2} \\
&\leq 2 \max_{[0, 1]} |h(x)| \left[ \int_0^1 (h'(s))^2 \, ds \right]^{1/2} \\
&= 2\|h\|_\infty \left[ \int_0^1 (h'(s))^2 \, ds \right]^{1/2}.
\end{align*}

Since this inequality holds for all \( x \in [0, 1] \), it follows that

\[ \max_{[0, 1]} h(x)^2 = \|h\|_\infty^2 \leq 2\|h\|_\infty \left[ \int_0^1 (h'(s))^2 \, ds \right]^{1/2}. \]  (30)

Division on both sides by \( \|h\|_\infty > 0 \) yields the desired result.

From the Lemma and the second inequality in (24), it follows that

\[ \|h(x)\|_\infty \leq 2\|h\|_\infty. \]  (31)

Using this result in Eq. (26) yields

\[ \frac{\|R(u, h)\|}{\|h\|} \leq (3\|u(x)\|_\infty + 2\|h\| + 1) \|h\|. \]  (32)
It now follows that
\[ \frac{\|R(u, h)\|}{\|h\|} \to 0 \quad \text{as} \quad h \to 0. \] (33)

Therefore, the Fréchet derivative of the nonlinear functional \( K \) in Eq. (20) is given by
\[ [DK(u)]h = \int_0^1 [3u^2 h + 2u'h'] \, dx. \] (34)