1 Haunted by Analysis 2

1.1 The Frechet Derivative

We give a reminder of what a Frechet derivative is.

Definition 1. Let \( f : V \rightarrow W \) be a mapping between two Banach spaces \( V \) and \( W \). The Frechet derivative of \( f \) at some \( x \in V \), denoted by \( (Df)(x) \), is a bounded linear operator \( B(V, W) \) which is an approximation of \( f \) near \( x \) in the following sense:

\[
\lim_{h \to 0} \frac{\|f(x+h) - f(x) - ((Df)(x))(h)\|_W}{\|h\|_V} = 0
\]

\( f \) is called Frechet-differentiable iff \( (Df)(x) \) exists for all \( x \in V \).

Claim 2. Assume that \( V \) and \( W \) are finite dimensional. Then every linear operator is bounded. Furthermore, if all partial derivatives of \( f \) exist and are continuous, then \( f \) is Frechet differentiable and \( (Df)(x) \) is identified with the matrix given with entries \( (\partial_i f_j)(x) \) (the Jacobian matrix). The converse is false as seen in some pathological examples.

Claim 3. If \( f \) is linear itself then \( (Df)(x) \) is independent of \( x \) and is equal to \( f \).

Proof. \( (Df)(x) \) is unique if it exists (…). Then assuming \( f \) is linear, we have

\[
\frac{\|f(x+h) - f(x) - ((Df)(x))(h)\|_W}{\|h\|_V} = \frac{\|f(h) - ((Df)(x))(h)\|_W}{\|h\|_V}
\]

so that \( (Df)(x) := f \) does the job.

Remark 4. Note that \( (Df)(x) \) can also be seen as a map \( V \ni x \mapsto (Df)(x) \in B(V, W) \). In this sense, this map is not generically linear. Indeed, here’s an
example: \( f : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \) given by \( f \left( \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) := \begin{bmatrix} (x_1)^3 \\ (x_2)^3 \\ (x_3)^3 \end{bmatrix} \). The Fréchet derivative of this map is given by the matrix

\[
[(Df)(x)]_{i,j=1}^3 = \begin{bmatrix}
(\partial_1 f_1)(x) & (\partial_2 f_1)(x) & (\partial_3 f_1)(x) \\
(\partial_1 f_2)(x) & (\partial_2 f_2)(x) & (\partial_3 f_2)(x) \\
(\partial_1 f_3)(x) & (\partial_2 f_3)(x) & (\partial_3 f_3)(x)
\end{bmatrix}
\]

\[
= \begin{bmatrix}
3(x_1)^2 & 0 & 0 \\
0 & 3(x_2)^2 & 0 \\
0 & 0 & 3(x_3)^2
\end{bmatrix}
\]

and this matrix, as a function of \( x \), is clearly not linear.

\section{Linear Algebra}

\subsection{Orientation}

Let \( V \) be a finite dimensional vector space. We know that there is an isomorphism \( V \cong \mathbb{R}^n \) for some \( n \in \mathbb{N}_{>0} \).

**Definition 5.** A choice of such an isomorphism \( f : V \rightarrow \mathbb{R}^n \) is an orientation on \( V \).

**Definition 6.** Two orientations \( f_1 : V \rightarrow \mathbb{R}^n \) and \( f_2 : V \rightarrow \mathbb{R}^n \) are called “equivalent” if the linear map \( f_1 \circ f_2^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n \), which is an \( n \times n \) matrix, has positive determinant.

**Claim 7.** There are exactly two equivalence classes for orientations.

**Definition 8.** A map \( f : V \rightarrow V \) is orientation preserving iff \( \det((Df)(x)) > 0 \) for all \( x \in V \).

**Example 9.** Consider the reflection on \( \mathbb{R}^3 \), given by \( -I_{3\times3} \). Its determinant is \((-1)^3 = -1\) so it is not orientation preserving.

**Remark 10.** Deformations of rigid bodies should preserve orientation.

\subsection{Symmetric positive definite matrices}

**Definition 11.** (Cholesky decomposition) A matrix \( P \in Mat_{n \times n}(\mathbb{R}) \) is called positive iff there is some \( A_P \in Mat_{n \times n}(\mathbb{R}) \) such that \( (A_P)^T A_P = P \).

**Claim 12.** The following are equivalent:

1. \( P \) is positive.
2. $P$ is symmetric and has eigenvalues in $[0, \infty)$.

3. $P$ is symmetric and $\langle x, Px \rangle \geq 0$ for all $x \in \mathbb{R}^n$.

Proof. 1. implies 2.: Assume that $P$ is positive. Then $P = A^T A$ for some $A$. Then $P^T = (A^T A)^T = A^T (A^T)^T = A^T A = P$ so that $P$ is symmetric. Let $\lambda \in \sigma(P)$. Then there is some $v \in \mathbb{R}^n \setminus \{0\}$ such that $Pv = \lambda v$. If $\lambda = 0$ we are finished. Otherwise, $A^T Av = \lambda v$ implies

$$1 = \frac{\|v\|^2}{\|v\|^2} = \frac{\langle v, v \rangle}{\|v\|^2} = \frac{1}{\|v\|^2} \langle v, \lambda v \rangle = \frac{1}{\|v\|^2} \langle v, A^T Av \rangle = \frac{1}{\|v\|^2} \frac{\|Av\|^2}{\|v\|^2}$$

which implies that $\lambda = \frac{\|Av\|^2}{\|v\|^2} > 0$. Since $\lambda$ was an arbitrary eigenvalue of $P$, we find $\sigma(P) \subseteq \mathbb{R}_{\geq 0}$.

2. implies 3.: Any symmetric matrix may be orthogonally diagonalized: $P = O^T DO$ where $O \in O(n)$ and $D$ is a diagonal matrix whose entries are the eigenvalues of $P$. Since we assume $\sigma(P) \in \mathbb{R}_{\geq 0}$, the entries of $D$ are in $\mathbb{R}_{\geq 0}$. Then if $x \in \mathbb{R}^n$ is given,

$$\langle x, Px \rangle = \langle x, O^T DOx \rangle = \langle Ox, DOx \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} (Ox)_i D_{ij} (Ox)_j \quad \text{(D is diagonal)}$$

$$= \sum_{i=1}^{n} (Ox)_i D_{ii} (Ox)_i = \sum_{i=1}^{n} [(Ox)_i]^2 D_{ii} \geq 0$$

Each term in the last sum is non-negative numbers due to it being the product of two non-negative numbers.
3. implies 1.: \( P \) is symmetric, so we diagonalize it as \( P = O^T D O \) as above. Then note that by the above calculation, \( D_{ii} \geq 0 \) for all \( i \) (otherwise we reach a contradiction). As a result, \( \sqrt{D} \) is defined and is a diagonal matrix whose entries are \( \sqrt{D_{ij}} \). Define \( A := \sqrt{DO} \). Then

\[
A^T A = (\sqrt{DO})^T \sqrt{DO} \\
= O^T \sqrt{D} \sqrt{DO} \\
= O^T DO \\
= P
\]

\[
\square
\]

2.3 Polar Decomposition

Let \( A \in \text{Mat}_{n \times n}(\mathbb{R}) \) be given. As we’ve seen in the lecture, there are unique left and right polar decompositions given by

\[
A = O |A| \\
= |A^T| O
\]

where \( |A| \equiv \sqrt{A^T A} \) and \( O := A |A|^{-1} = |A^T|^{-1} A \).

**Example 13.** (Thanks to Hansueli) Note that in general \( |A| \neq |A^T| \). Indeed, let \( A = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix} \). We have

\[
A^T A = \left( \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix} \right)^T \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix} \\
= \frac{1}{2} \begin{bmatrix} 2 & 2 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix} \\
= \frac{1}{2} \begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix}
\]

whereas

\[
AA^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix} \left( \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix} \right)^T \\
= \frac{1}{2} \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix} \\
= \frac{1}{2} \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix} \\
= \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}
\]
This example corresponds to $|A^T| \equiv \sqrt{AA^T}$ being stretch along the $e_1$ axis and then $A|A|^{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ being rotation by 45 degrees.