Recall:
\[ \| f(x) - f(x_0) - Df(x_0) \cdot (x - x_0) \| \to 0 \]

as \( x \to x_0 \).

As \( x \to x_0 \),

\[ f(x) \approx Df(x_0) \cdot (x - x_0) + f(x_0) = Df(x_0) \cdot x + (f(x_0) - Df(x_0) \cdot x_0) \]

affine transformation in \( x \).

"Best affine approximation to \( f(x) \) near \( x_0 \)."

Example. \( f : \mathbb{R} \to \mathbb{R} \)

[Diagram showing a tangent line at \( x_0 \).]
Example: \( f: \mathbb{R}^2 \to \mathbb{R} \)

Def’n (Differentiability on \( A \subset \mathbb{R}^n \)) If \( f \) is differentiable at each point \( x \in A \), then we say \( f \) is differentiable on \( A \).

Theorem (Uniqueness of derivative) Let \( A \subset \mathbb{R}^n \) be an open set. Suppose that \( f: A \to \mathbb{R}^m \) is differentiable at \( x_0 \in A \), then \( Df(x_0) \) is uniquely determined by \( f \).

Proof. Let \( L_1 \) and \( L_2 \) be two linear transformations satisfying the definition of derivative, i.e. \( \forall \delta > 0, \exists \delta_1, \delta_2 > 0 \) s.t. \( x \in A \) with \( \|x - x_0\| < \delta_1 \) implies...
$$||f(x) - f(x_0) - L_1(x - x_0)|| < \varepsilon \ \text{and} \ x \in A \ \text{with} \ ||x - x_0|| < \delta_1$$

implies

$$||f(x) - f(x_0) - L_2(x - x_0)|| < \varepsilon \ ||x - x_0||$$

Choose any unit vector in $\mathbb{R}^n$
denoted by $e \in \mathbb{R}^n$, $||e|| = 1$

then let

$$x = x_0 + \lambda e$$

with $\lambda \in \mathbb{R}$

we choose $\lambda$ small enough s.t. $x \in A$

and $||x - x_0|| < \min \{\delta_1, \delta_2\}$

$$||x - x_0|| = ||x_0 + \lambda e - x_0|| = ||\lambda e|| = ||\lambda|| \cdot ||e|| = ||\lambda||$$

$$||L_1e - L_2e|| = \frac{||L_1(\lambda e) - L_2(\lambda e)||}{||\lambda||}$$

$$= \frac{||L_1(x - x_0) - L_2(x - x_0)||}{||\lambda||}$$

$$= \frac{||f(x) - f(x_0) - L_2(x - x_0) - (f(x) - f(x_0) - L_1(x - x_0))||}{||x - x_0||}$$
\[ \begin{align*}
&\leq \frac{\| f(x) - f(x_0) - L_2(x-x_0) \|}{\| x - x_0 \|} \\
&\quad + \frac{\| f(x) - f(x_0) - L_1(x-x_0) \|}{\| x - x_0 \|} \\
&< 2 \varepsilon
\end{align*} \]

Since \( \varepsilon \) is arbitrary,
\[ \| L_1 e - L_2 e \| = 0 \Rightarrow L_1 e = L_2 e \]

By arbitrariness of unit vector \( e \),
we know that \( L_1 = L_2 \).

Remark: A must be open. → HW

\[ f : \mathbb{R} \to \mathbb{R}, \quad f : \{1\} \to \mathbb{R}, \quad f(1) = 1 \]

\[ A \]

\[ A \delta \circ, \quad |x - x_0| < \delta \]
and \( x \in A \)

\[ \Rightarrow x = x_0 = 1 \]
\[ \forall x \in \mathbb{R}: \quad |f(x) - f(x_0) - c(x - x_0)| \]

\[ = \left| \frac{f(x_0) - f(x_0) - c(x - x_0)}{x - x_0} \right| = 0 < \varepsilon \quad |x - x_0| \]

\( \Rightarrow \) any real number \( c \in \mathbb{R} \) can be defined as the derivative of \( f \) at \( x_0 = 1 \).

**HW Problem 27:** If \( f: \mathbb{R}^n \to \mathbb{R}^m \) is linear, then \( \exists M > 0 \) s.t.

\[ \|f(x)\| \leq M \|x\| \quad \text{for all } x \in \mathbb{R}^n. \]

**Theorem:** Suppose \( A \subset \mathbb{R}^n \) is open and \( f: A \to \mathbb{R}^m \) is differentiable at \( x_0 \in A \), then \( f \) is continuous at \( x_0 \).

**Proof:** \( \forall \varepsilon > 0 \), by differentiability, \( \exists \delta > 0 \) s.t. \( \|x - x_0\| < \delta \) and \( x \in A \), we have

\[ \|f(x) - f(x_0) - Df(x_0) \cdot (x - x_0)\| < \varepsilon \|x - x_0\| \]
Since $Df(x_0)$ is a linear transformation,
\[ \exists M > 0 \text{ s.t. } \| Df(x_0) \cdot x \| \leq M \| x \|, \forall x \in \mathbb{R}^n. \]

By triangle inequality:
\[ \| f(x) - f(x_0) \| \leq 3 \| x - x_0 \| + \| Df(x_0) \cdot (x-x_0) \| \]
\[ \leq 3 \| x - x_0 \| + M \| x - x_0 \| \]
then if take $\delta < \min \{ \delta, \frac{\epsilon}{3M} \}$,
\[ \| f(x) - f(x_0) \| \leq 2 \epsilon \] whenever $x \in A$
\[ \| x - x_0 \| < \delta \]

Def'n (partial derivative) The partial derivatives of $f$, denoted as $\frac{\partial f}{\partial x_i}$ for $i=1,\ldots,m$ and\[ 1\leq i\leq n, \] is given by
\[ \frac{\partial f_i}{\partial x_i}(x_1,\ldots,x_n) = \lim_{h \to 0} \frac{f_i(x_1,\ldots,x_i+h,\ldots,x_n) - f_i(x_1,\ldots,x_n)}{|h|} \] if the limit exists.
Def'n (Jacobian matrix) Suppose that $A \subset \mathbb{R}^n$ is open, and $f: A \rightarrow \mathbb{R}^m$, with all partial derivatives exist. Then
\[
\begin{bmatrix}
\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\
\frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n}
\end{bmatrix}
\]
is called the Jacobian matrix of $f$.

Theorem (Matrix representation of $Df(x)$) Suppose that $A \subset \mathbb{R}^n$ is open and that $f: A \rightarrow \mathbb{R}^m$ is differentiable, then $Df(x)$ can be represented by the Jacobian matrix of $f$. Moreover, the representation is unique.

Proof. Uniqueness follows from previous theorem.
Assume at some point \( x \in A \subseteq \mathbb{R}^n \), the matrix representation is

\[
Df(x) = \begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\]

It suffices to show \( \forall j = 1, \ldots, m, \ i = 1, \ldots, n \),

\[
A_{ji} = \frac{\partial f_j}{\partial x_i}
\]

Let \( e_i = (0, \ldots, 0, 1, 0, \ldots, 0) \) be the standard \(^i^{th}\) component basis of \( \mathbb{R}^n \), and let \( y = x + he_i \), then

\[
\frac{\|f(x) - f(x) - Df(x) \cdot (y - x)\|}{\|y - x\|}
\]

\[
= \frac{\|f(x_1, \ldots, x_i + h, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_n)\|}{h} - h \frac{Df(x) \cdot e_i}{h}
\]

\[
\rightarrow 0 \quad \text{as} \quad h \rightarrow 0 \quad \text{(by differentiability)}
\]
Notice that
\[ \text{D}(\text{fix}) \cdot e_i = \begin{bmatrix} a_{ii} & a_{i2} & \cdots & a_{in} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \]

\[ = \begin{bmatrix} a_{ii} \\ a_{i1} \\ \vdots \\ a_{in} \end{bmatrix} \]

look at the \( j \)-th component of this vector \( a_{ji} \)

\[ \lim_{h \to 0} \frac{|f_i(x, \ldots, x_i - h, x_i + h, \ldots, x_n) - f_i(x, \ldots, x_n)|}{|h|} \cdot h a_{ji} \]

\[ \to 0 \quad \text{as} \quad h \to 0 \]

\[ \implies a_{ji} = \frac{\partial f_i}{\partial x_i} \]

\( \square \)
Examples: \( f : \mathbb{R}^3 \rightarrow \mathbb{R}^2 \) by
\[
f(x_1, x_2, x_3) = (x_1, x_2, x_2 x_3^3)
\]
then
\[
Df(x_1, x_2, x_3) = \begin{bmatrix}
x_2 & x_1 & 0 \\
0 & x_3 & 3 x_2 x_3^2
\end{bmatrix}
\]
\( f : \mathbb{R}^2 \rightarrow \mathbb{R}^3 \)
\[
f(x_1, x_2) = (x_1^2, x_1 x_2, x_1^4 x_2)
\]
then
\[
Df(x_1, x_2) = \begin{bmatrix}
2 x_1 & 0 \\
3 x_1 x_2 & x_1^3 \\
4 x_1^3 x_2 & x_1^4
\end{bmatrix}
\]

Defn (Gradient) If \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is differentiable, \( Df \) is also called the \underline{gradient} of \( f \).

\[
\nabla f(x) = Df(x) = (\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), \ldots, \frac{\partial f}{\partial x_n}(x))
\]

Example: \( f : \mathbb{R}^2 \rightarrow \mathbb{R} \)
\[
f(x_1, x_2) = \frac{s \sin x_1}{x_2}
\]
\[
\nabla f(x) = \left( \frac{\cos x_1}{x_2}, - \frac{s \sin x_1}{x_2^2} \right)
\]

HW Problem 29-31