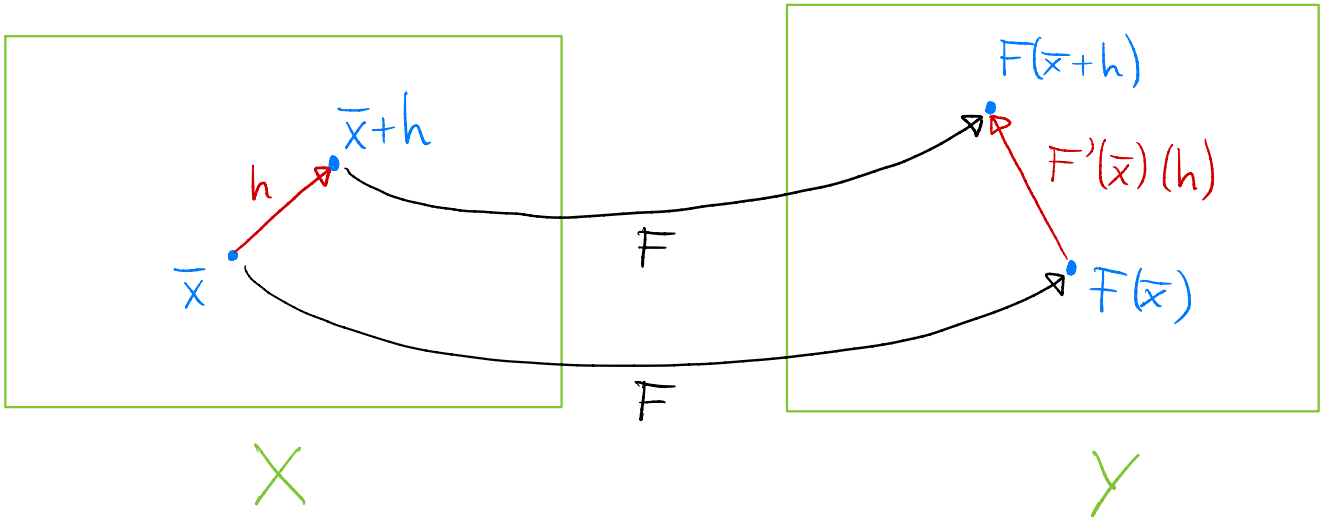
THE GATEAUX DERIVATIVE

The Fréchet derivative $F'(\bar{x})$ of F at $\bar{x}$ is a bounded linear operator such that

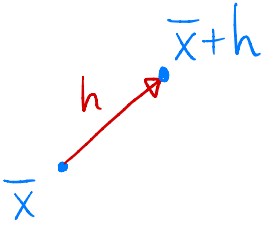
$$F(\bar{x}+h) \approx F(\bar{x}) + F'(\bar{x})(h) \quad \text{for small } h.$$



The Fréchet derivative $F'(\bar{x})$ of F at $\bar{x}$ is a bounded linear operator such that

$$F(\bar{x}+h) \approx F(\bar{x}) + F'(\bar{x})(h) \quad \text{for small } h.$$

- $F'(\bar{x})$ gives us (almost) the relation between F at $\bar{x}$ and F at nearby values $\bar{x}+h$, in any direction h.
- But $F'(\bar{x})$ might be difficult to compute directly!
- What if we only compute the "derivative" in one particular direction?



Definition

Let $F: O \to Y$ be a function and $\bar{x} \in O$.

The Gateaux derivative (or directional derivative) of F at $\bar{x}$ in direction $r \in X$ is

$$F'(\bar{x}; r) = \lim_{\substack{t \to 0 \\ t \in \mathbb{R}}} \frac{F(\bar{x} + tr) - F(\bar{x})}{t}$$

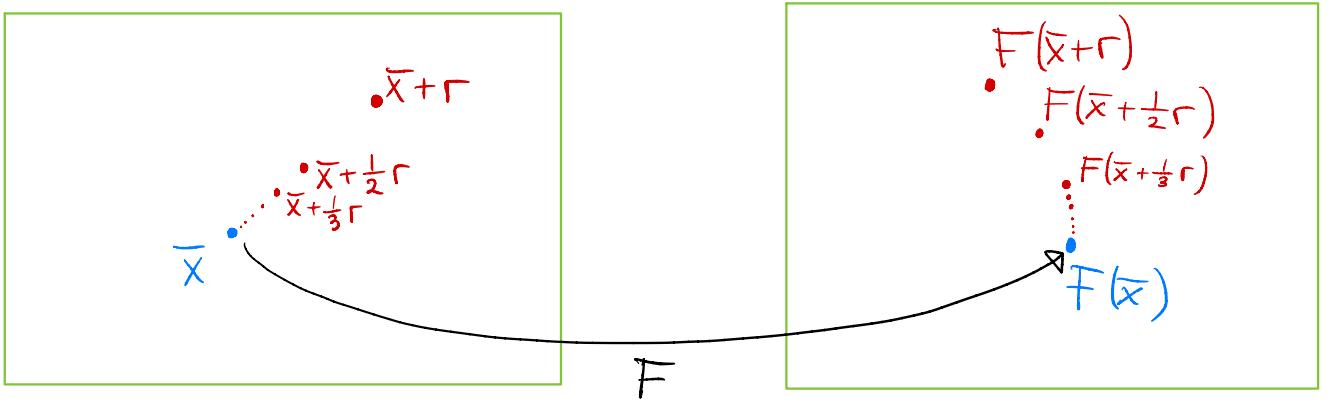
$F'(\bar{x}; r)$ is the rate of change in F when walking in direction r.

Definition

Let $F: O \to Y$ be a function and $\bar{x} \in O$.

The Gateaux derivative (or directional derivative) of F at $\bar{x}$ in direction $r \in X$ is

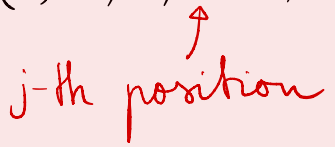
$$F'(\bar{x}; r) = \lim_{\substack{t \to 0 \\ t \in \mathbb{R}}} \frac{F(\bar{x} + tr) - F(\bar{x})}{t}$$



Recall: The partial derivatives of $F: \mathbb{R}^m \to \mathbb{R}^n$ are

$$\frac{\partial F}{\partial x_j}(\bar{x}) = \lim_{t \to 0} \frac{F(\bar{x} + t e_j) - F(\bar{x})}{t} = F'(\bar{x}; e_j)$$

where $e_j = (0, \dots, 0, 1, 0, \dots, 0) = (\delta_{j,k})_{k=1}^{m}$



Proposition

If F is Fréchet differentiable at $\bar{x}$, then

$$F'(\bar{x}; r) = F'(\bar{x})(r) \qquad \text{for all } r \in X.$$

Proof: Let $h = tr$ (for $t \in \mathbb{R}$) in the definition of Fréchet differentiability:

$$\frac{\| F(\bar{x}+tr) - F(\bar{x}) - F'(\bar{x})(tr) \|_Y}{\|tr\|_X} = \frac{1}{\|r\|_X} \left\| \frac{F(\bar{x}+tr)-F(\bar{x})}{t} - F'(\bar{x})(r) \right\|_Y \xrightarrow{t \to 0} 0$$

Hence, $F'(\bar{x})(r) = \lim_{t \to 0} \frac{F(\bar{x}+tr)-F(\bar{x})}{t} = F'(\bar{x}; r)$. ▨

Proposition

If F is Fréchet differentiable at $\bar{x}$, then

$$F'(\bar{x}; r) = F'(\bar{x})(r) \qquad \text{for all } r \in X.$$

Note:

- The Gateaux derivative $F'(\bar{x}; r)$ is easier to compute than the Fréchet derivative $F'(\bar{x})$, so the above is a tool to compute $F'(\bar{x})$:
  (i) let $A(r) = F(\bar{x}; r)$ for all r
  (ii) check that $A \in \mathcal{L}(X, Y)$.
  (iii) check that $\dfrac{\|F(\bar{x}+h) - F(\bar{x}) - A(h)\|}{\|h\|} \xrightarrow[h \to 0]{} 0$

Then $F'(\bar{x}) = A$.

Proposition

If F is Fréchet differentiable at $\bar{x}$, then

$$F'(\bar{x}; r) = F'(\bar{x})(r) \quad \text{for all } r \in X.$$

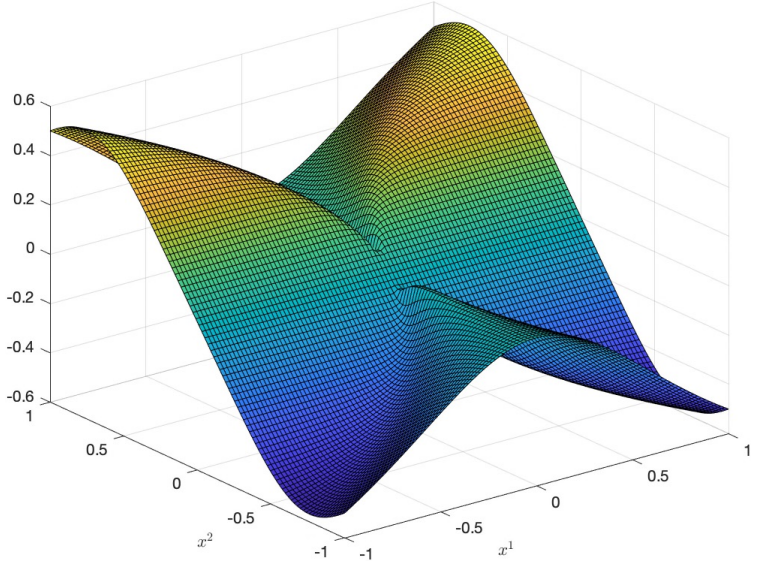
Note:

- The opposite is <u>not</u> true! $F'(\bar{x}; r)$ might exist for all r, but not $F'(\bar{x})$.

Example

Let $F: \mathbb{R}^2 \to \mathbb{R}$ be given by $F(x_1, x_2) = \begin{cases} \dfrac{x_1^2 x_2}{x_1^2 + x_2^4} & (x \neq 0) \\ 0 & (x = 0). \end{cases}$

Then F is Gateaux, but not Fréchet differentiable at $x = 0$.



Example

Let $F: \mathbb{R}^2 \to \mathbb{R}$ be given by $F(x_1, x_2) = \begin{cases} \dfrac{x_1^2 x_2}{x_1^2 + x_2^4} & (x \neq 0) \\ 0 & (x = 0). \end{cases}$

Then F is Gateaux, but not Fréchet differentiable at $\overline{x} = 0$.

The Gateaux derivative is computed in the usual way:

$$F'(0; r) = \lim_{t \to 0} \frac{F(tr_1, tr_2)}{t} = \lim_{t \to 0} \frac{t^3 r_1^2 r_2}{t(t^2 r_1^2 + t^4 r_2^4)} = \lim_{t \to 0} \frac{r_1^2 r_2}{r_1^2 + t^2 r_2^4}$$

$$= \begin{cases} \dfrac{r_1^2 r_2}{r_1^2} = r_2 & (\text{if } r_1 \neq 0) \\ 0 & (\text{if } r_1 = 0). \end{cases}$$

Example

Let $F: \mathbb{R}^2 \to \mathbb{R}$ be given by $F(x_1, x_2) = \begin{cases} \dfrac{x_1^2 x_2}{x_1^2 + x_2^4} & (x \neq 0) \\ 0 & (x = 0). \end{cases}$

Then F is Gateaux, but not Fréchet differentiable at $x = 0$.

Thus, if $F'(0)$ exists then $F'(0)(r) = F'(0; r) = \begin{cases} r_2 & (r_1 \neq 0) \\ 0 & (r_1 = 0). \end{cases}$

But this function is not linear:

$$F'(0; (1,1)) + F'(0; (-1,1)) = 1 + 1 \neq 0 = F'\big(0; (1,1) + (-1,1)\big).$$

Example

Let $F: \mathbb{R}^2 \to \mathbb{R}$ be given by $F(x_1, x_2) = \begin{cases} \dfrac{x_1^2 x_2}{x_1^2 + x_2^4} & (x \neq 0) \\ 0 & (x = 0). \end{cases}$

Then F is Gateaux, but not Fréchet differentiable at $x = 0$.

We can also note that if F were Fréchet differentiable then $F'(0)(r) \overset{!}{=} \underbrace{\nabla F(0)}_{=0} r = 0$, but on the other hand,

$$0 = F'(0)((1,1)) \overset{!}{=} F'(0; (1,1)) = 1,$$ a contradiction.

Example

Let $F : \ell^2(\mathbb{R}) \longrightarrow \ell^1(\mathbb{R})$ be given by $F(x) = (x(1)^2, x(2)^2, x(3)^2, \ldots)$.

Compute $F'(\bar{x})$ for any $\bar{x} \in \ell^2(\mathbb{R})$.

Pause now and try to solve this!

Example

Let $F : \ell^2(\mathbb{R}) \longrightarrow \ell^1(\mathbb{R})$ be given by $F(x) = (x(1)^2, x(2)^2, x(3)^2, \dots)$.

Compute $F'(\bar{x})$ for any $\bar{x} \in \ell^2(\mathbb{R})$.

Note first that F is well-defined: If $x \in \ell^2$ then

$$\|F(x)\|_{\ell^1} = \sum_{n=1}^{\infty} |F(x)(n)| = \sum_{n=1}^{\infty} |x(n)^2| = \|x\|_{\ell^2}^2 < \infty,$$

so $F(x) \in \ell^1$ for every $x \in \ell^2$.

We compute first the Gateaux derivative. For every $n \in \mathbb{N}$, $r \in \ell^2$ and $t \neq 0$,

$$\left(\frac{F(\bar{x}+tr) - F(\bar{x})}{t}\right)(n) = \frac{(\bar{x}(n)+tr(n))^2 - \bar{x}(n)^2}{t}$$

$$= \frac{\bar{x}(n)^2 + 2t\bar{x}(n)r(n) + t^2 r(n)^2 - \bar{x}(n)^2}{t}$$

$$= 2\bar{x}(n)r(n) + tr(n)^2.$$

Example

Let $F : \ell^2(\mathbb{R}) \to \ell^1(\mathbb{R})$ be given by $F(x) = (x(1)^2, x(2)^2, x(3)^2, \dots)$.

Compute $F'(\bar{x})$ for any $\bar{x} \in \ell^2(\mathbb{R})$.

Hence, $\dfrac{F(\bar{x}+tr) - F(\bar{x})}{t} = 2\bar{x}r + tr^2$, where $r^2(n) = (r(n)^2)_{n\in\mathbb{N}}$ and $(\bar{x}r)(n) = (\bar{x}(n)r(n))_{n\in\mathbb{N}}$. Note that $r^2 \in \ell^1$ for every $r \in \ell^2$, and $\bar{x}r \in \ell^1$ since $\|\bar{x}r\|_{\ell^1} \le \|\bar{x}\|_{\ell^2}\|r\|_{\ell^2}$, by Hölder's inequality.

We get $\|tr^2\|_{\ell^1} = |t|\,\|r^2\|_{\ell^1} \xrightarrow{t\to 0} 0$

($\|r^2\|_{\ell^1}$: finite and independent of t)

$$\Longrightarrow \quad F'(\bar{x}; r) = \lim_{t\to 0} \left(2\bar{x}r + tr^2\right) = 2\bar{x}r.$$

Example

Let $F : \ell^2(\mathbb{R}) \longrightarrow \ell^1(\mathbb{R})$ be given by $F(x) = \left(x(1)^2, x(2)^2, x(3)^2, \dots\right)$.

Compute $F'(\bar{x})$ for any $\bar{x} \in \ell^2(\mathbb{R})$.

Claim: $F'(\bar{x})$ exists, and $F'(\bar{x})(h) = F'(\bar{x}; h) = 2\bar{x}h$.

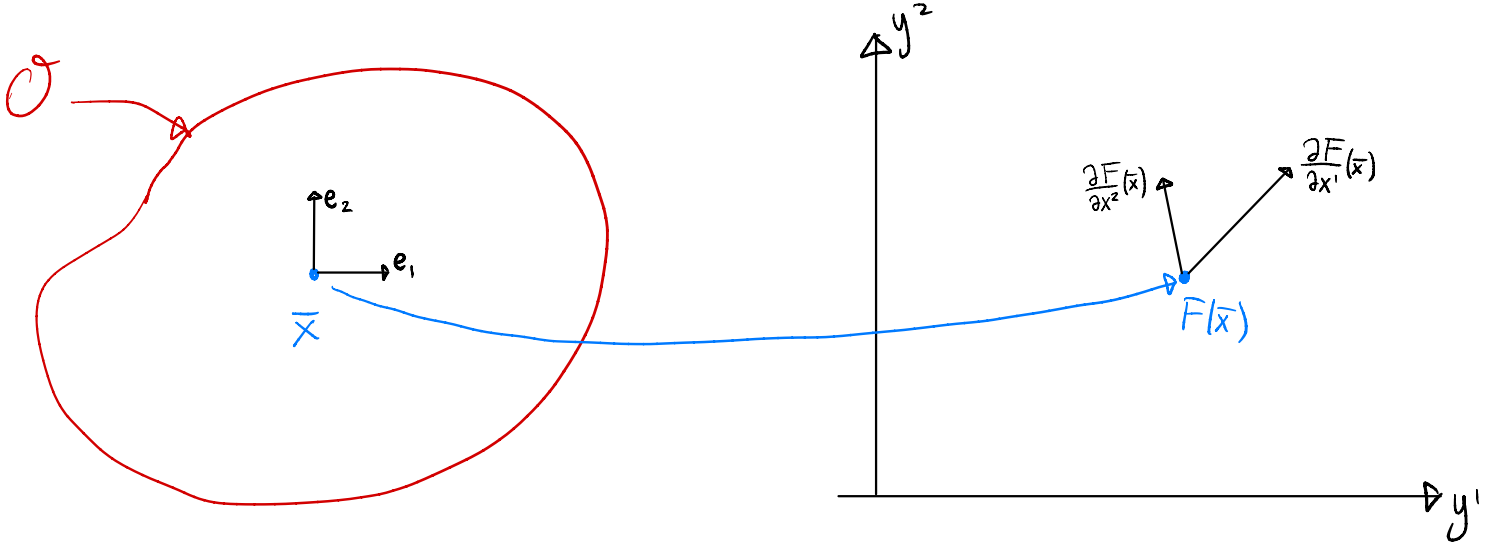
Indeed, $$\frac{\|F(\bar{x}+h) - F(\bar{x}) - 2\bar{x}h\|_{\ell^1}}{\|h\|_{\ell^2}} = \frac{\|h^2\|_{\ell^1}}{\|h\|_{\ell^2}}$$

$$= \frac{\|h\|_{\ell^2}^2}{\|h\|_{\ell^2}} = \|h\|_{\ell^2} \longrightarrow 0 \qquad \text{as } h \longrightarrow 0.$$



Proposition

Let $O \subseteq \mathbb{R}^m$ be open, let $\bar{x} \in O$ and assume $F : O \to \mathbb{R}^n$ is such that all partial derivatives $\frac{\partial F}{\partial x^j}(x)$ $(j = 1, \dots, m)$ exist for all $x \in O$ and are continuous at $x = \bar{x}$.
Then F is Frechet differentiable at $\bar{x}$.



Proposition

Let $O \subseteq \mathbb{R}^m$ be open, let $\bar{x} \in O$ and assume $F : O \to \mathbb{R}^n$ is such that all partial derivatives $\frac{\partial F}{\partial x^j}(x)$ $(j = 1, \dots, m)$ exist for all $x \in O$ and <u>are continuous</u> at $x = \bar{x}$.

Then F is Frechet differentiable at $\bar{x}$.

Recall: The vector-valued functions $\frac{\partial F}{\partial x^j} : O \to \mathbb{R}^n$ are continuous <u>iff</u> each component $\frac{\partial F^i}{\partial x^j} : O \to \mathbb{R}$ is continuous.

Proposition

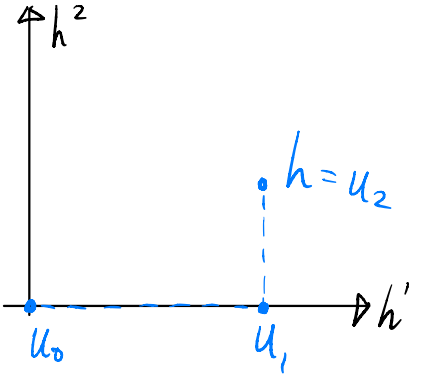
Let $O \subseteq \mathbb{R}^m$ be open, let $\bar{x} \in O$ and assume $F : O \to \mathbb{R}^n$ is such that all partial derivatives $\frac{\partial F}{\partial x^j}(x)$ $(j = 1, \dots, m)$ exist for all $x \in O$ and are continuous at $x = \bar{x}$.

Then F is Frechet differentiable at $\bar{x}$.

Proof: We claim that $F'(\bar{x})(h) = \nabla F(\bar{x})\, h$.

Jacobian of F at $\bar{x}$

Fix $h \in \mathbb{R}^m$ and let $u_j = \sum_{k=1}^{j} h^k e_k$ $(j = 0, \dots, m)$.



Then

$$F(\bar{x}+h) - F(\bar{x}) = F(\bar{x}+u_m) - F(\bar{x}+u_{m-1}) + F(\bar{x}+u_{m-1}) - \dots - F(\bar{x}+u_0)$$

$$= \sum_{j=1}^{m} F(\bar{x}+u_j) - F(\bar{x}+u_{j-1})$$

(a telescoping sum)

Proposition

Let $O \subseteq \mathbb{R}^m$ be open, let $\bar{x} \in O$ and assume $F : O \to \mathbb{R}^n$ is such that all partial derivatives $\frac{\partial F}{\partial x^j}(x)$ $(j = 1, \dots, m)$ exist for all $x \in O$ and are continuous at $x = \bar{x}$.
Then F is Frechet differentiable at $\bar{x}$.

$$F(\bar{x}+h) - F(\bar{x}) = \sum_{j=1}^{m} F(\bar{x}+u_j) - F(\bar{x}+u_{j-1})$$

For each term we can apply the mean value theorem:

$$F(\bar{x}+u_j) - F(\bar{x}+u_{j-1}) = \frac{d}{dt} F\left(\bar{x} + u_{j-1} + t(u_j - u_{j-1})\right)\Big|_{t=\xi} \quad \text{(for some } \xi \in (0,1)\text{)}$$

$$= \nabla F\left(\bar{x} + u_{j-1} + \xi h^j e_j\right) h^j e_j = \frac{\partial F}{\partial x^j}\left(\bar{x} + u_{j-1} + \xi h^j e_j\right) h^j.$$

Proposition

Let $O \subseteq \mathbb{R}^m$ be open, let $\bar{x} \in O$ and assume $F : O \to \mathbb{R}^n$ is such that all partial derivatives $\frac{\partial F}{\partial x^j}(x)$ $(j = 1, \dots, m)$ exist for all $x \in O$ and are continuous at $x = \bar{x}$.
Then F is Frechet differentiable at $\bar{x}$.

$$F(\bar{x}+h) - F(\bar{x}) = \sum_{j=1}^{m} \frac{\partial F}{\partial x^j}\left(\bar{x} + u_{j-1} + \xi h^j e_j\right) h^j.$$

Each $\frac{\partial F}{\partial x^j}$ is continuous at $\bar{x}$, so $\frac{\partial F}{\partial x^j}(\dots) \xrightarrow[h \to 0]{} \frac{\partial F}{\partial x^j}(\bar{x})$. Hence:

$$\frac{\| F(\bar{x}+h) - F(\bar{x}) - \nabla F(\bar{x})\, h \|}{\|h\|} = \left\| \sum_{j=1}^{m} \left(\frac{\partial F}{\partial x^j}\left(\bar{x} + u_{j-1} + \xi h^j e_j\right) - \frac{\partial F}{\partial x^j}(\bar{x}) \right) h^j \right\| / \|h\|$$

$$\leq \sum_{j=1}^{m} \underbrace{\left\| \frac{\partial F}{\partial x^j}(\dots) - \frac{\partial F}{\partial x^j}(\bar{x}) \right\|}_{\to 0} \underbrace{|h^j| / \|h\|}_{\leq 1}$$

$$\xrightarrow[h \to 0]{} 0.$$

∎

QUESTIONS?
COMMENTS?

Next video: The mean value theorem