Next: Gradient.

For $f: U \to \mathbb{R}$ ($U \subset \mathbb{R}^n$ open) differentiable, we have $(Df)_i = (\nabla f)_i$, where

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$ is called the **gradient of f, or "nabla f"**.

Note: Often we write $\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}$, a differential operator.

Two results:

- If $(Df)(x) \neq 0$, then f has greatest directional derivative in direction $\frac{\nabla f(x)}{\|\nabla f(x)\|}$.

  ($D_u f = Df \cdot u = \nabla f \cdot u = \|\nabla f\| \underbrace{\|u\|}_{=1} \cos\varphi$ is maximal for $\varphi = 0$.)

  angle between ∇f and u

- If f has a local extremum at x, then $\nabla f(x) = 0$.

  (If $\nabla f(x) \neq 0$, then f increases in at least one direction and decreases in the opposite direction (from 1-dim. calculus), thus it cannot have a local extremum.)

2.2 Higher Order Derivatives

We showed: $f: U \to \mathbb{R}^m$ ($U \subset \mathbb{R}^n$ open) continuously differentiable.

$\iff$

All $\frac{\partial f_i}{\partial x_j}$ exist and are continuous.

We call this class of functions C^1, or $C^1(U)$.

Second partial derivatives are defined as $\frac{\partial}{\partial x_i} \frac{\partial f_\ell}{\partial x_j} = \frac{\partial^2 f_\ell}{\partial x_i \partial x_j}$.

We say $\boxed{f \text{ is of class } C^k}$ (or $C^k(U)$) if all k-th partial derivatives exist in all components and are continuous.

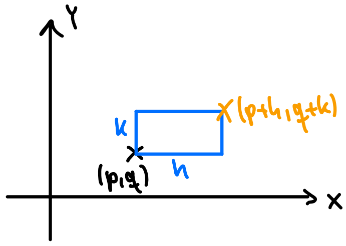
Generally, $\frac{\partial f_\ell}{\partial x_i \partial x_j} \neq \frac{\partial f_\ell}{\partial x_j \partial x_i}$ is possible, see homework.

But, we have:

<u>Theorem</u> (Clairaut's thm., or Schwarz's thm.):
If $f: U \to \mathbb{R}^m$ ($U \subset \mathbb{R}^n$ open) is of class C^2, then $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$ $\forall i, j$.

This implies the more general result: If f is of class C^k, then all partial derivatives up to order k $\underbrace{\text{commute}}_{\text{can be interchanged}}$.

Sketch of proof:



We apply the mean-value thm. twice: $\exists$ point (x,y) in $\square$ s.t.

- $\dfrac{f(p+h,q)-f(p,q)}{h} = \dfrac{\partial f}{\partial x}(x,y)$
- $\dfrac{1}{k}\left[\dfrac{f(p+h,q+k)-f(p,q+k)}{h} - \left(\dfrac{f(p+h,q)-f(p,q)}{h}\right)\right] = \dfrac{\partial}{\partial y}\dfrac{\partial f}{\partial x}(x,y)$

$$\Rightarrow \frac{f(p+h,q+k)-f(p+h,q)-f(p,q+k)+f(p,q)}{hk} = \frac{\partial^2 f}{\partial y \partial x}(x,y)$$

$$\text{LHS} \longrightarrow \frac{\frac{\partial f}{\partial y}(p+h) - \frac{\partial f}{\partial y}(p)}{h} \longrightarrow \frac{\partial}{\partial x}\frac{\partial f}{\partial y}(p,q), \quad \text{RHS} \longrightarrow \frac{\partial^2 f}{\partial y \partial x}(p,q) \text{ by continuity.} \quad \square$$

(A more detailed proof is, e.g., in Kantorovitz: Theorem 2.2.2, or in Rudin: Theorem 9.41.)

Note: The matrix H with $(H_f(x))_{ij} := \dfrac{\partial^2 f(x)}{\partial x_i \partial x_j}$ is called $\boxed{\text{Hessian matrix of } f.}$

Due to Schwarz, H_f is symmetric (for $f \in C^2$), i.e., $(H_f)_{ij} = (H_f)_{ji}$.