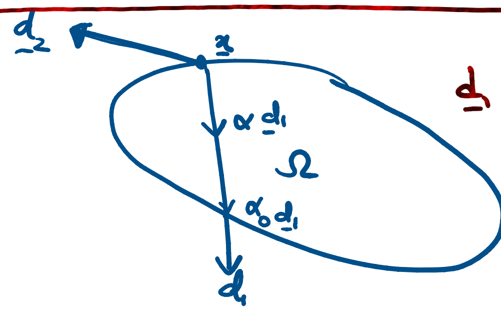
Lecture 2

Thursday, 6 January 2022 1:58 PM

→ More on feasible direction [Exs]
→ Directional derivatives → Read Chapter 5
→ Necessary condⁿ for minima

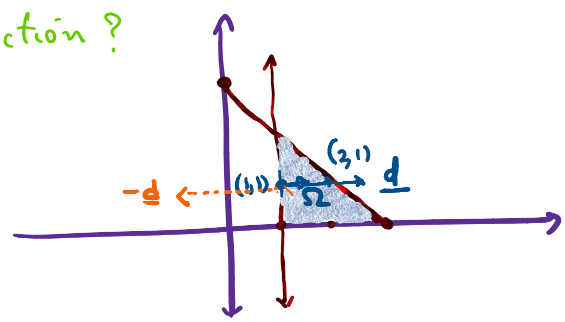
Recall: A vector $\boxed{\underline{d} \in \mathbb{R}^n}$ $\boxed{(\underline{d} \neq 0)}$ is a feasible direction at $\boxed{\underline{x} \in \Omega}$ if $\exists$ exists $\boxed{\alpha_0 > 0}$ such that $\boxed{\underline{x} + \alpha \underline{d} \in \Omega}$ for all $\boxed{\alpha \in [0, \alpha_0]}$.



d_1 is a feasible direction, d_2 is not.

Examples 1. Let $\underline{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\Omega = \{ \underline{x} \in \mathbb{R}^2 : x_1 + x_2 \leq 3,\ x_1 \geq 1,\ x_2 \geq 0 \}$.

Is $\underline{d} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ a feasible direction?



Soln. $\underline{x} + \alpha \underline{d} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$= \begin{pmatrix} 1+\alpha \\ 1 \end{pmatrix} \in \Omega$

$1 + \alpha + 1 = 2 + \alpha \leq 3$ ✓ $\alpha \in [0, \overset{\alpha_0}{1}]$

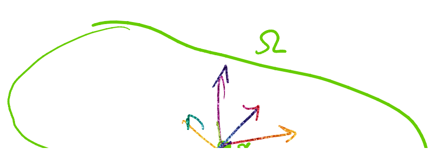
$1 + \alpha \geq 1$

$1 \geq 0$

} $\underline{d}$ is a feasible direction.

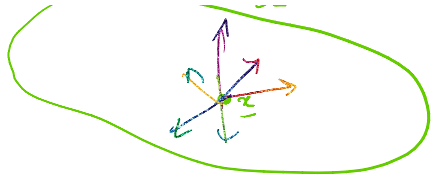
Ex. S.T. $-\underline{d}$ is not feasible.

2.



$\underline{x}$ is an interior point, all directions $\underline{d} \in \mathbb{R}^n$

2.



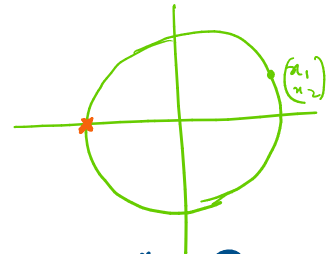
all directions $\underline{d} \in \mathbb{R}^n$ are feasible directions.

$\mapsto$ nonlinear constraint

3. For $\Omega = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : \boxed{x_1^2 + x_2^2 = 1} \right\}$,

(Feasible set) Points on bdry of the circle.

$\begin{pmatrix} x_1 + \alpha d_1 \\ x_2 + \alpha d_2 \end{pmatrix}$



$f = 2x_1$

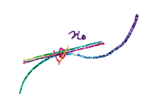
$\min\limits_{x \in \Omega} f(x)$

there are NO feasible directions at any $\underline{x}^* \in \Omega$.

$\underline{x}^* + \alpha \underline{d} \in S \quad \forall \alpha \in [0, \alpha_0]$ happens only for $\boxed{\underline{d} = \underline{0}}$ not allowed!

4. $\Omega = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_2 = -x_1^3 \right\} \to$ Feasible directions $\to$ empty set.

I. Directional derivative of $f : \mathbb{R}^n \to \mathbb{R}$ in the direction $\underline{d}$ is denoted by $\frac{\partial f}{\partial \underline{d}}$ and is defined by



$$\frac{\partial f}{\partial \underline{d}}(\underline{x}) = \lim_{\alpha \to 0} \frac{f(\underline{x} + \alpha \underline{d}) - f(\underline{x})}{\alpha}$$

Computing directional derivatives:
$$\frac{\partial f}{\partial \underline{d}}(\underline{x}) = \frac{d}{d\alpha} f(\underline{x} + \alpha \underline{d})\Big|_{\alpha = 0}$$
$$= \nabla f(\underline{x})^T \underline{d}$$
$$= \underline{d}^T \nabla f(\underline{x})$$
$$= \langle \nabla f(\underline{x}), \underline{d} \rangle$$

Sections 5.2, 5.3, 5.4, 5.5, 5.6

$\nearrow$ (DD)

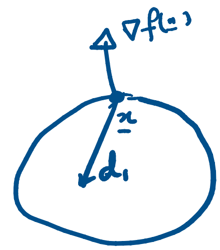
If $\| \underline{d} \| = 1$, this is the rate of increase of f in the direction $\underline{d}$.

Ex. Let $f : \mathbb{R}^3 \to \mathbb{R}$ $\quad f(x) = x_1 x_2 x_3 \quad \underline{d} = \left[\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}} \right]^T$

Ex: Let $f : \mathbb{R}^3 \to \mathbb{R}$ $\quad f(x) = x_1 x_2 x_3 \quad \underline{d} = [\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}]$

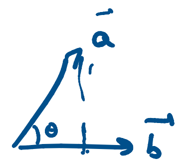
$$\frac{\partial f}{\partial \underline{d}}(\underline{x}) = \langle \nabla f(\underline{x}), \underline{d} \rangle = \left\langle \begin{pmatrix} x_2 x_3 \\ x_1 x_3 \\ x_1 x_2 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \right\rangle$$

$$= \frac{x_2 x_3 + x_1 x_3}{2} + \frac{x_1 x_2}{\sqrt{2}}$$

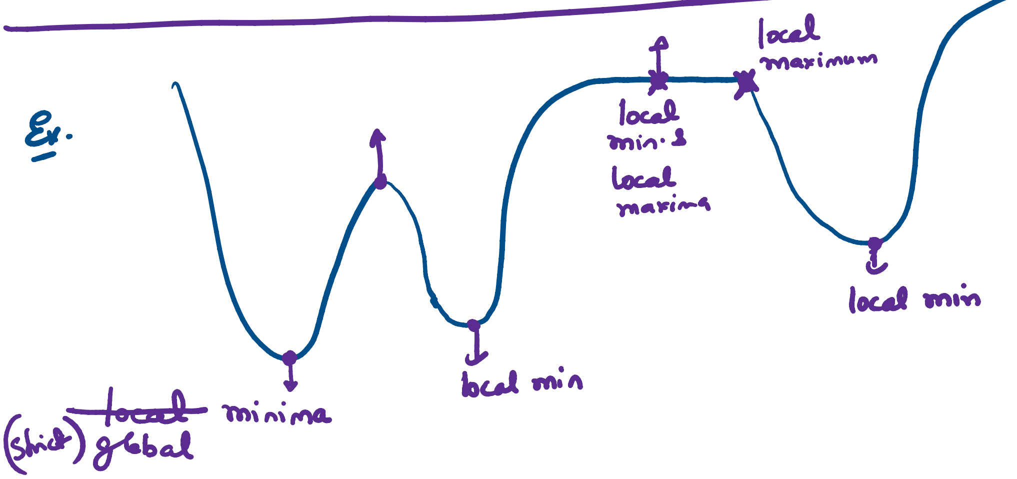


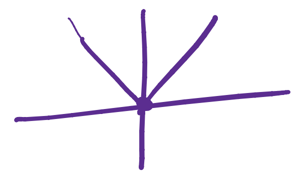
$$\|\underline{d}\| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \frac{1}{2}} = \sqrt{\frac{1}{4} + \frac{1}{4} + \frac{1}{2}} = 1$$

Dot product in $\mathbb{R}^2$: $\quad \underline{a} \cdot \underline{b} = |a|\,|b| \cos\theta$



Ex.





Theorem 1 [First order necessary condition] (FONC)

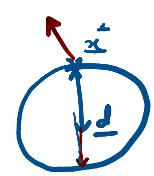
Let Ω be a subset of $\mathbb{R}^n$ and $f \in \mathcal{C}^{(1)}$ be a real-valued ($f : \mathbb{R}^n \to \mathbb{R}$) function on Ω. If $\underline{x}^*$ is a local minimizer of f over Ω, then for any feasible direction $\underline{d}$ at $\underline{x}^*$,

$$\underline{d}^T \nabla f(\underline{x}^*) \geq 0 \qquad \left(\frac{\partial f}{\partial \underline{d}}(\underline{x}^*) \geq 0 \right).$$

[$\mathcal{C}^1 \to$ first derivative exists & is continuous]

Pf. Define $\underline{x}(\alpha) = \underline{x}^* + \alpha \underline{d} \in \Omega$

$\underline{x}(0) = \underline{x}^*$



Define the composite function $\phi(\alpha) = f(\underline{x}(\alpha))$

Define the composite function $\phi(\alpha) = f(\ldots)$

$$f(\underline{x}^* + \alpha \underline{d}) - f(\underline{x}^*) = \phi(\alpha) - \phi(0)$$
$$= \phi'(0)\alpha + \boxed{o(\alpha)}$$

$\left[\lim_{\alpha \to 0} \frac{|o(\alpha)|}{\alpha} = 0 \right]$

Given that $\underline{x}^*$ is a local minimum for f

$\Rightarrow f(\underline{x}^* + \alpha \underline{d}) - f(\underline{x}^*) \geq 0$ for sufficient small values of 'α'.

$$\phi'(\alpha) = \nabla f(\underline{x}(\alpha)) \cdot \underline{d}$$

$$f(\underline{x}^* + \alpha \underline{d}) - f(\underline{x}^*) = \alpha \underbrace{\underline{d}^T \nabla f(\underline{x}^*)}_{\geq 0 \ ?} + o(\alpha) \cdot \geq 0$$

(for suff. small values of α)

$$\overbrace{h(\alpha)}^{\geq 0} = \alpha \, \underline{d}^T \nabla f(\underline{x}^*) + o(\alpha)$$

$$\Rightarrow -\alpha \, \underline{d}^T \nabla f(\underline{x}^*) = -h(\alpha) + o(\alpha)$$

$$\lim_{\alpha \to 0} \frac{o(\alpha)}{\alpha} = 0 \Rightarrow \forall |\alpha| < \delta, \quad \left| \frac{o(\alpha)}{\alpha} \right| \leq \boxed{\frac{h(\alpha)}{2\alpha}}$$

$$\Rightarrow |o(\alpha)| \leq \frac{h(\alpha)}{2}$$

$$\Rightarrow -\alpha \, \underline{d}^T \nabla f(\underline{x}^*) = -h(\alpha) + \frac{h(\alpha)}{2} = -\frac{\overline{h(\alpha)}}{2} \leq 0$$

$$\Rightarrow \boxed{\underline{d}^T \nabla f(\underline{x}^*) \geq 0}$$

$\boxed{\text{Interior pt } x^* \in \Omega}$ For any $\underline{d} \in \mathbb{R}^n$,

$$\underline{d}^T \nabla f(\underline{x}^*) \geq 0$$
$$-\underline{d}^T \nabla f(\underline{x}^*) \geq 0$$

$$0 = \langle \underline{d}, \nabla f(\underline{x}^*) \rangle \Leftarrow \Rightarrow \underline{d}^T \nabla f(\underline{x}^*) = 0$$

$$\Rightarrow \boxed{\nabla f(\underline{x}^*) = 0}$$

$\ldots$ order $\Big] \quad f(x) = o(g(x)) \Big]$ examples

Exercise: Small 'o' order] $f(x) = o(g(x))$] → Examples

Capital 'O'] $f(x) = O(g(x))$]

Taylor's Theorem.

Alternate: $f(\underline{x}^*) \le f(\underline{x}^* + \alpha d)$

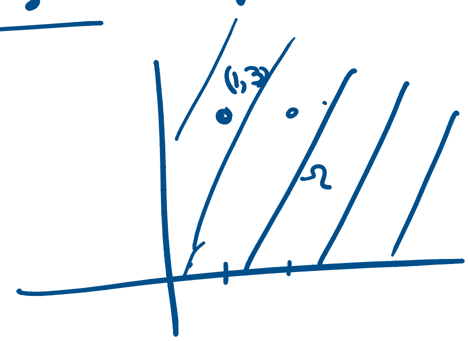
$$\frac{f(\underline{x}^* + \alpha \underline{d}) - f(\underline{x}^*)}{\alpha} \geq 0$$

Take limit as $\alpha \to 0$ $\Rightarrow \frac{\partial f}{\partial \underline{d}}(\underline{x}^*) \geq 0$

Ex Consider

$$\min \quad x_1^2 + \frac{1}{2}x_2^2 + 3x_2 + 4.5$$
$$\text{s.t.} \quad x_1, x_2 \geq 0$$

Is the FONC for a local minimizer satisfied at $\underline{x}^* = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.



$$\nabla f(\underline{x}) = \begin{bmatrix} 2x_1 \\ x_2 + 3 \end{bmatrix}$$

$$\nabla f(\underline{x}^*) = \begin{bmatrix} 2 \\ 6 \end{bmatrix} \neq 0.$$

→ $\underline{x}^* = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}$