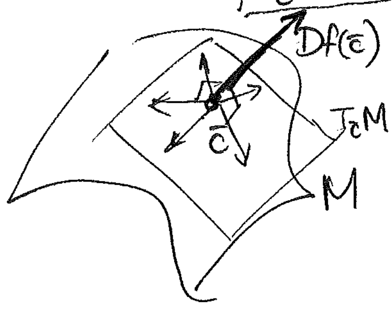
(24) 2/7/2017 >

But just as usual (unconstrained) extrema occur for $\mathbb{R}^n \xrightarrow{f} \mathbb{R}$ at points $\bar{a} \in \mathbb{R}^n$ where $Df(\bar{a}) = \bar{0}$, i.e. all directional derivatives for f vanish, one would expect the extrema for f on M to occur at points $\bar{c} \in M$ where <u>f does not ~~change~~ change in the directions tangent to M at $\bar{c}$</u>,

i.e. $\boxed{T_{\bar{c}}M \subset \ker Df(\bar{c})}$ ← called a <u>constrained critical point of f on M</u> (DEF'N 3.7.2)



let's check this...

<u>THM 3.7.1</u>: For a manifold $M \subset \mathbb{R}^n$ and $\underset{\mathbb{R}^n}{\overset{\cap}{U^{open}}} \xrightarrow{f} \mathbb{R}$ in $C^1(U)$, if a point $\bar{c} \in U \cap M$ is a local extremum for f on M, then $T_{\bar{c}}M \subset \ker Df(\bar{c})$

<u>proof</u>: If M is k-dimensional, we can parametrize it locally near $\bar{c} = \begin{pmatrix} \bar{a} \\ \bar{b} \end{pmatrix}$ as the image of some γ like this

$$\begin{array}{ccccc}
V^{open} & \xrightarrow{\gamma} & U^{open} & \xrightarrow{f} & \mathbb{R} \\
\cap & & \cap & & \\
\mathbb{R}^k & & \mathbb{R}^n & & \\
\bar{x} & \longmapsto & \begin{pmatrix} \bar{x} \\ \bar{g}(\bar{x}) \end{pmatrix} & & \\
 & & \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} & \longmapsto & f\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} \\
\bar{a} & \longmapsto & \begin{pmatrix} \bar{a} \\ \bar{b} \end{pmatrix} = \bar{c} = \gamma(\bar{a}) & &
\end{array}$$

Then, as before, $\bar{a}$ must be a critical point for the composite map $f \circ \gamma : \mathbb{R}^k \longrightarrow \mathbb{R}$, that is

$$D(f \circ \gamma)(\bar{a}) = [0]$$

$$\overset{\text{chain rule}}{=} \underbrace{[Df(\gamma(\bar{a}))]}_{[Df(\bar{c})]}[D\gamma(\bar{a})].$$

In other words, $\operatorname{im}[D\gamma(\bar{a})] \subset \ker[Df(\bar{c})]$.

$\operatorname{im}[D\gamma(\bar{a})] = \text{graph } \left\{ \begin{pmatrix} \bar{x} \\ [D\bar{g}(\bar{a})]\bar{x} \end{pmatrix} : \bar{x} \in \mathbb{R}^k \right\} \text{ of } D\bar{g}(\bar{a}) \overset{\text{by definition}}{=} T_{\bar{c}}M$

∎

When M is not parametrized, but cut out implicitly, there is a handy way to rephrase this, called the <u>Method of Lagrange Multipliers</u>:

<u>THM-DEF'N 3.7.5:</u> Given $\underset{\mathbb{R}^n}{\overset{\wedge}{U^{\text{open}}}} \xrightarrow{\bar{F}} \mathbb{R}^{n-k}$, $\bar{z} \longmapsto \bar{F}(\bar{z}) = \begin{pmatrix} F_1(\bar{z}) \\ \vdots \\ F_{n-k}(\bar{z}) \end{pmatrix}$ cutting out our k-dimensional manifold M as $\bar{F}(\bar{z}) = \bar{0}$ near $\bar{c} \in M$, and some $\mathbb{R}^n \xrightarrow{f} \mathbb{R}$, with $\bar{F}, f \in C^1$, then $\bar{c}$ is a critical point for f <u>constrained to M</u>

$$\Updownarrow \text{ by DEF'N}$$

$$T_{\bar{c}}M \subset \ker Df(\bar{c})$$

$$\Updownarrow \text{ since } T_{\bar{c}}M = \ker D\bar{F}(\bar{c})$$

$\ker D\bar{F}(\bar{c}) \supseteq \ker \begin{bmatrix} D\bar{F}(\bar{c}) \\ \hdashline Df(\bar{c}) \end{bmatrix}$ has equality $\Longleftrightarrow$ $\ker D\bar{F}(\bar{c}) \subset \ker Df(\bar{c})$

$\Updownarrow$ rank-nullity formula $\qquad\qquad \Updownarrow \leftarrow$ needs proof! (see ~~right~~ side)

row $D\bar{F}(\bar{c}) \subseteq$ row $\begin{bmatrix} D\bar{F}(\bar{c}) \\ \hdashline Df(\bar{c}) \end{bmatrix}$ has equality $\Longleftrightarrow$ $\underset{\substack{\text{row vectors} \\ 1\times n}}{Df(\bar{c})} \overset{(*)}{=} \lambda_1 DF_1(\bar{c}) + \dots + \lambda_{n-k} DF_{n-k}(\bar{c})$ for some $\lambda_1, \dots, \lambda_{n-k}$ called <u>Lagrange multipliers</u>

$$\Updownarrow \text{ transpose vectors!}$$

$$\left(\underset{\substack{\text{column vectors} \\ n\times 1}}{\nabla f(\bar{c})} = \lambda_1 \nabla F_1(\bar{c}) + \dots + \lambda_{n-k} \nabla F_{n-k}(\bar{c}) \right)$$

$$\Updownarrow \text{ rephrasing}$$

not mentioned in our book

$(\bar{c}, \bar{\lambda})$ is a <u>critical point</u> for the <u>Lagrangian function</u>

$$\mathcal{L}(\bar{z}, \bar{\lambda}) := f(\bar{z}) \overset{\text{(plus in some books)}}{-} \left(\lambda_1 F_1(\bar{z}) + \dots + \lambda_{n-k} F_{n-k}(\bar{z}) \right) \text{ in variables } (\bar{z}, \bar{\lambda})$$

i.e. $(\bar{c}, \bar{\lambda})$ solves the system

$$\begin{cases} \dfrac{\partial L}{\partial z_i} = 0 & \text{for } i = 1, 2, \dots, n \\[2ex] \dfrac{\partial L}{\partial \lambda_j} = 0 & \text{for } j = 1, 2, \dots, n-k \end{cases}$$

$\leftarrow$ these just say $(*)$ holds at $(\bar{c}, \bar{\lambda})$

$\leftarrow$ these just say $\bar{F}(\bar{c}) = 0$, i.e. $\bar{c}$ lies on M

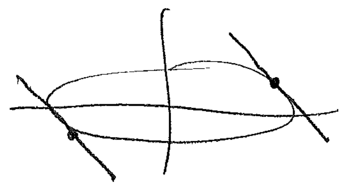
<u>EXAMPLES</u>:

① To find points on ellipse $0 = F\begin{pmatrix} x \\ y \end{pmatrix} = \frac{x^2}{9} + \frac{y^2}{4} - 1$ minimizing/maximizing $f\begin{pmatrix} x \\ y \end{pmatrix} = x + y$, look for critical points $\begin{pmatrix} x \\ y \\ \lambda \end{pmatrix}$ of

$$\mathcal{L}(x, y, \lambda) = f\begin{pmatrix} x \\ y \end{pmatrix} - \lambda F\begin{pmatrix} x \\ y \end{pmatrix}$$
$$= x + y - \lambda \left(\frac{x^2}{9} + \frac{y^2}{4} - 1 \right)$$

(2b) So we solve
$$\begin{cases} 0 = \frac{\partial \mathcal{L}}{\partial x} = 1 - \lambda \frac{2x}{9} & \Rightarrow x = \frac{9}{2\lambda} \\ 0 = \frac{\partial \mathcal{L}}{\partial y} = 1 - \lambda \frac{2y}{4} & \Rightarrow y = \frac{2}{\lambda} \\ 0 = \frac{\partial \mathcal{L}}{\partial \lambda} = \frac{x^2}{9} + \frac{y^2}{4} - 1 & \Rightarrow \frac{9^2}{9 \cdot 4\lambda^2} + \frac{4}{4\lambda^2} = 1 \end{cases}$$

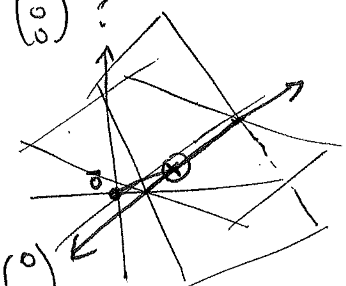
$$\frac{9}{4} + 1 = \lambda^2$$
$$\lambda = \pm\sqrt{\frac{13}{4}} = \frac{\pm\sqrt{13}}{2}$$



$$\begin{pmatrix} x \\ y \end{pmatrix} = \pm \begin{pmatrix} 9/\sqrt{13} \\ 4/\sqrt{13} \end{pmatrix}$$

---

(2) What point on the line in $\mathbb{R}^3$ defined by $\left\{ \begin{matrix} x+y+z=1 \\ x+2y+3z=4 \end{matrix} \right\}$ lies closest to $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$?



i.e. minimize $f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + y^2 + z^2$

subject to constraints $\bar{F}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \overbrace{x+y+z-1}^{F_1\begin{pmatrix} x \\ y \\ z \end{pmatrix}} \\ \underbrace{x+2y+3z-4}_{F_2\begin{pmatrix} x \\ y \\ z \end{pmatrix}} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Lagrangian $\mathcal{L}\begin{pmatrix} x \\ y \\ z \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = f\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \lambda_1 F_1\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \lambda_2 F_2\begin{pmatrix} x \\ y \\ z \end{pmatrix}$

$$= x^2 + y^2 + z^2 - \lambda_1(x+y+z-1) - \lambda_2(x+2y+3z-4)$$

has critical points $\begin{pmatrix} \bar{x} \\ \bar{\lambda} \end{pmatrix}$ solving the system

$$\nabla f = \lambda_1 \nabla F_1 + \lambda_2 F_2 \begin{cases} 0 = \frac{\partial \mathcal{L}}{\partial x} = 2x - \lambda_1 - \lambda_2 \\ 0 = \frac{\partial \mathcal{L}}{\partial y} = 2y - \lambda_1 - 2\lambda_2 \\ 0 = \frac{\partial \mathcal{L}}{\partial z} = 2z - \lambda_1 - 3\lambda_2 \end{cases}$$

$$\bar{F} = \bar{0} \begin{cases} 0 = \frac{\partial \mathcal{L}}{\partial \lambda_1} = x+y+z-1 \\ 0 = \frac{\partial \mathcal{L}}{\partial \lambda_2} = x+2y+3z-4 \end{cases}$$

a linear system, easy to solve (if you're a computer) $\Longrightarrow$ $\begin{pmatrix} x \\ y \\ z \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} -2/3 \\ 1/3 \\ 4/3 \\ -10/3 \\ 2 \end{pmatrix}$

i.e. $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2/3 \\ 1/3 \\ 4/3 \end{pmatrix}$

REMARKS:

① There is meaning attached to the Lagrange multipliers $\lambda_1, \ldots, \lambda_{n-k}$.
If one imagines changing the constraint values $\bar{F}(\bar{z}) = \bar{0}$ slightly to $\bar{F}(\bar{z}) = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{n-k} \end{pmatrix}$, so $\mathcal{L}\begin{pmatrix} \bar{z} \\ \bar{\lambda} \end{pmatrix} = f(\bar{z}) + \lambda_1(F_1(\bar{z}) - \epsilon_1) + \ldots + \lambda_{n-k}(F_{n-k}(\bar{z}) - \epsilon_{n-k})$

one sees that $\lambda_i = \frac{\partial \mathcal{L}}{\partial \epsilon_i}$, and since $\mathcal{L}\begin{pmatrix} \bar{z} \\ \bar{\lambda} \end{pmatrix} = f(\bar{z})$ on the manifold $\bar{F}(\bar{z}) = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{n-k} \end{pmatrix}$,

λ_i gives the approximate change in the extremal value of $f(\bar{z})$ as one perturbs the constraint $F_i(\bar{z}) = 0$ to $F_i(\bar{z}) = \epsilon_i$ (called the shadow price for F_i in economics)

② There is a more complicated version of the Hessian/2nd derivative test, stated as THM 3.7.12 (proven in appendix A.14).

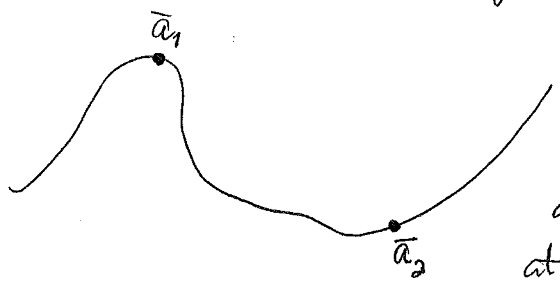
2/8/17

§3.8 Geometry of curves & surfaces

(— a whirlwind tour of some of Math 5378 "Differential geometry"!)

2nd derivatives and Hessians help us quantify curvature on curves and surfaces at a given point.

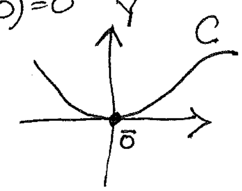
Given a curve $C \subset \mathbb{R}^2$, to quantify greater curvature at $\bar{a}_1$ than $\bar{a}_2$,



we put C into best coordinates near $\bar{a} \in C$ by translating $\bar{a}$ to $\bar{0}$ and rotating to make the tangent line at $\bar{a}$ horizonal.

i.e. locally C near $\bar{a} = \bar{0}$ is the graph of $Y = g(X)$ with $g(0) = 0$, $g'(0) = 0$

so $Y = \frac{g''(0)}{2} X^2 + o(|X|^2)$



DEF'N 3.8.1: The curvature $\kappa(\bar{a})$ at $\bar{a}$ on C is $\kappa(\bar{a}) := |g''(0)|$ (Why absolute value of $g''(0)$?)
and the radius of curvature is $r = \frac{1}{\kappa(\bar{a})}$.