5. Constraint Optimization

Buy a BMW

Optimization: buy the car that maximizes your "utility"

Constraint Opt: buy the car that maximizes your utility under a (budget/time/etc.) constraint

Math formulation

$$\begin{cases} \text{maximize} & f(x_1, x_2, \dots, x_n) \\ \text{subject to} & g(x_1, x_2, \dots, x_n) = 0 \end{cases}$$

add cond

Rk: You could have more than one constraint ($g_1 = 0$, $g_2 = 0$, $\dots$, $g_k = 0$)

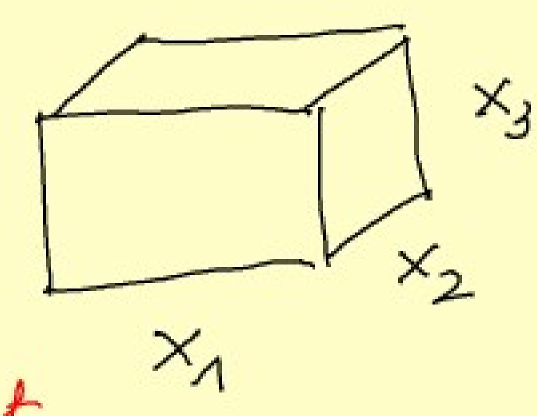
Simple but highly nontrivial example (see Exercises):
Box with minimal surface area and fixed volume (=1)

$$2x_1x_2 + 2x_2x_3 + 2x_3x_1 \longrightarrow \text{min}$$

surface

$$x_1x_2x_3 = 1$$

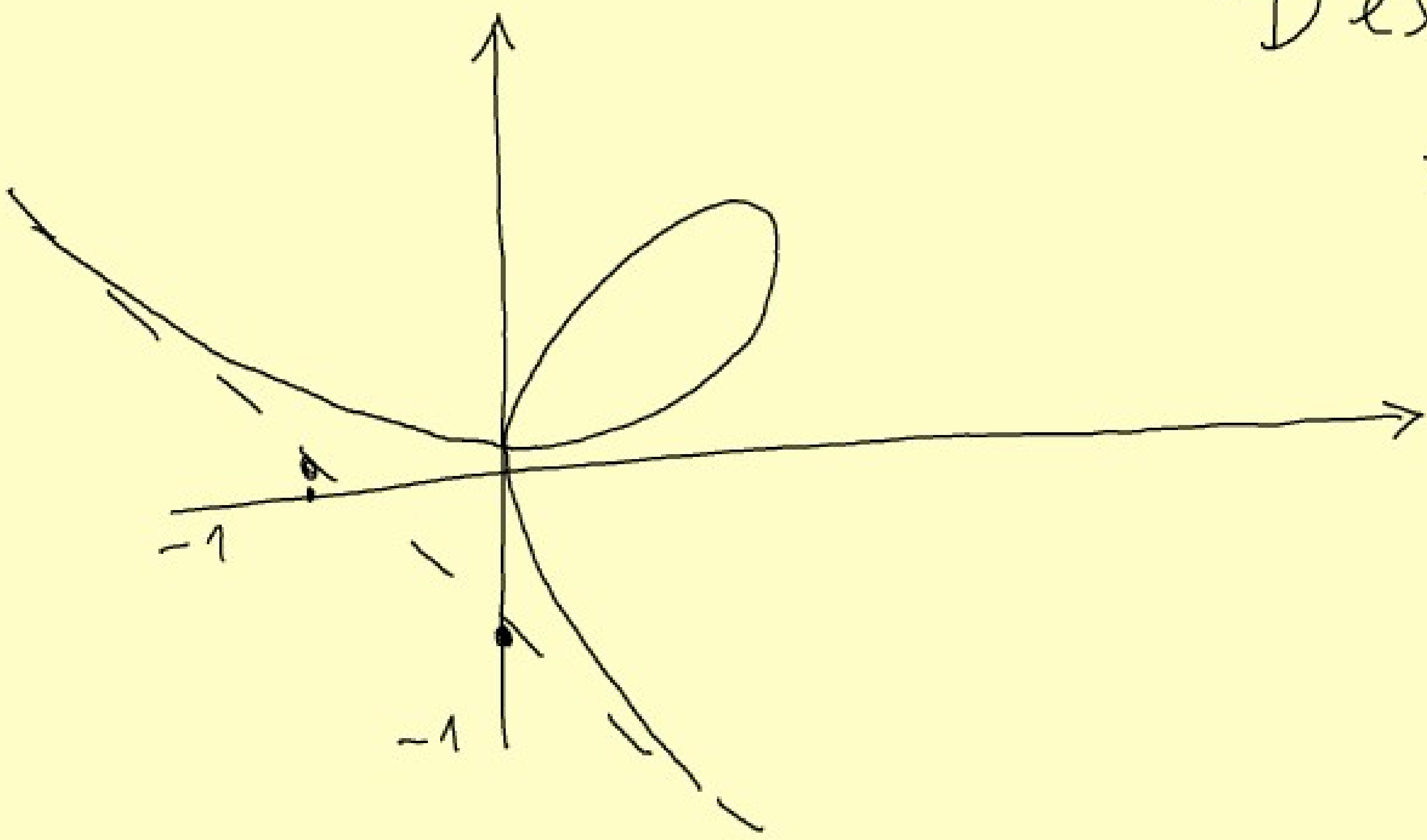
volume



Constraint Opt. probl reduce to unconstr Opt.
LAGRANGE Multiplier Meth.

§5.1. Planar curves and the Implicit Function Theorem.

"Functions" vs. "Curves"



Descarte's Folium (Leaf)
= the locus of all (x_1, x_2) with
$$x_1^3 + x_2^3 - 3x_1x_2 = 0$$
1638

Descartes challenged Fermat to find the tangent to this curve (Fermat recently had discovered the "method of tangents" = find tangent using 1st derivative)

Fermat solves the tricky problem using implicit differentiation

we'll get there!

But first: How should we "represent" curves?

3 Ways for describing curves

(C1) <u>implicit form</u> $\quad f(x_1, x_2) = 0$

Ex. Circle (unit) $\quad x_1^2 + x_2^2 = 1$

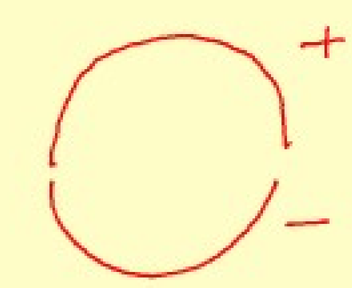
Folium $\quad x_1^3 + x_2^3 - 3x_1x_2 = 0$

(C2) <u>explicit form</u> (solve the implicit equation)

$$x_2 = \varphi_f(x_1)$$

Circle $\quad x_2 = \pm\sqrt{1 - x_1^2}$

two branches



Folium $\quad$?

(C3) <u>parametric form</u> (add a "parameter")

$$\begin{cases} x_1 = x_1(t) \\ \cancel{y}\,x_2 = \cancel{y}\,x_2(t) \end{cases} \qquad t \in [0, T)$$

param

Circle $\quad \begin{cases} x_1 = \cos t \\ x_2 = \sin t \end{cases} \qquad t \in [0, 2\pi)$

obviously $\quad x_1^2 + x_2^2 = (\cos t)^2 + (\sin t)^2 = 1$

Folium $\quad x_1 = \dfrac{3t}{1+t^3} \qquad x_2 = \dfrac{3t^2}{1+t^3}$

Rk: explicit form = automatically parametric form $\begin{cases} x_1 = x_1 \\ x_2 = \varphi(x_1) \end{cases}$

$x_1 = $ param

Thm 1. (Implicit Function Thm in $\mathbb{R}^2$)

Let $f : \mathbb{R}^2 \to \mathbb{R}$, $x^* = (x_1^*, x_2^*)$ with

(i) $f(x^*) = f(x_1^*, x_2^*) = 0$

(ii) f cont. diff-able

(iii) $\frac{\partial f}{\partial x_2}(x_1^*, x_2^*) \neq 0$

Then there exist $I_\varepsilon = (x_1^* - \varepsilon, x_1^* + \varepsilon) \subset \mathbb{R}$, $\varepsilon > 0$ and a function $\varphi : I_\varepsilon \to \mathbb{R}$ such that

(1) $\varphi(x_1^*) = x_2^*$

(2) $f(x_1, \varphi(x_1)) = 0 \quad \forall\, x_1 \in I$
$(\Leftrightarrow |x_1 - x_1^*| < \varepsilon)$

(3) φ is diff-able on I_ε and

$$\varphi'(x) = - \frac{\frac{\partial f}{\partial x_1}(x, \varphi(x))}{\frac{\partial f}{\partial x_2}(x, \varphi(x))}$$

"implicit differentiation"

Rk For the general $d > 2$ setting see [D. Popa, Thm 6.9.1]

HW: use above Thm for Descarte's Folium.

§ 5.2. Level sets (Level curves)

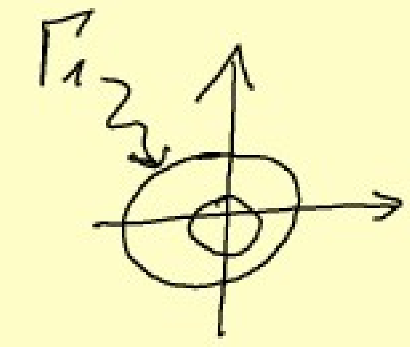
Def: $\quad f: \mathbb{R}^d \to \mathbb{R}, \quad c \in \mathbb{R}$

$$\Gamma_c = \{ (x_1, \dots, x_d) \in \mathbb{R}^d : f(x_1, \dots, x_d) = c \}$$

(may be empty!)

↑ c-level set of f

Ex. $\quad f(x_1, x_2) = x_1^2 + x_2^2$



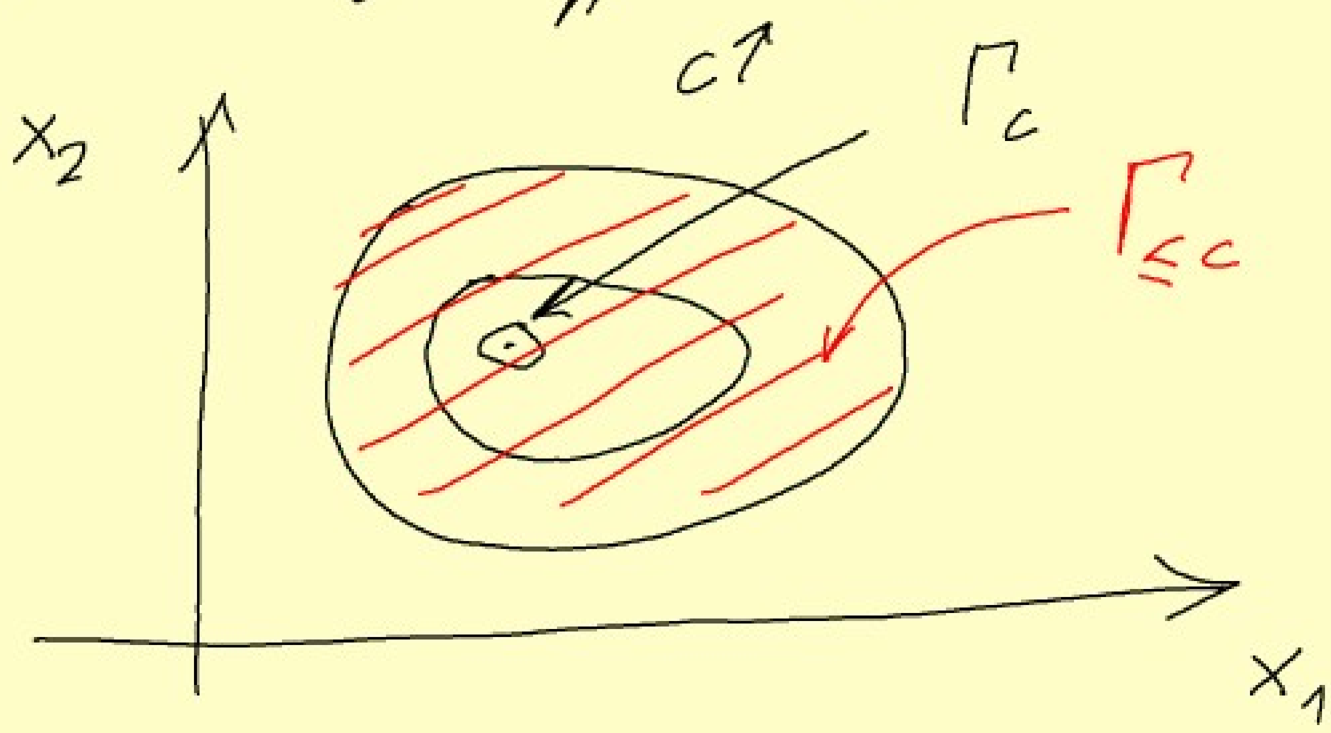
$c = 1 \qquad x_1^2 + x_2^2 = 1 \qquad$ unit circle

$c = \frac{1}{4} \qquad x_1^2 + x_2^2 = \frac{1}{4}$

$c = -1 \qquad x_1^2 + x_2^2 = -1 \ \lightning \qquad (\Gamma_{-1} = \emptyset)$

Why are level sets important?

↳ They offer geometric intuition to Optimiz.



Contour plot!

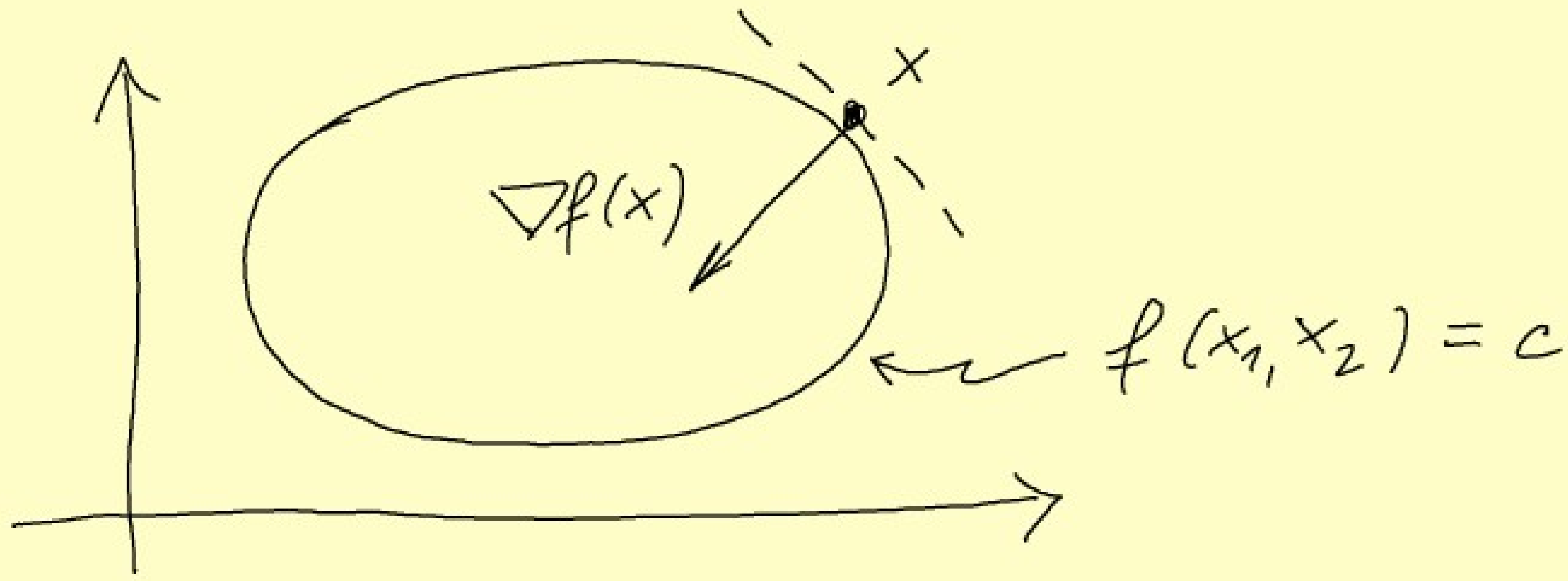
Level curves = also called "Contour lines".

Rk $\quad$ You can also talk about sub-level sets

$$\Gamma_{\leq c} = \{ (x_1, \dots, x_d) \in \mathbb{R}^d : f(x_1, \dots, x_d) \leq c \}$$

Th 2. The gradient is orthogonal to level sets!

$f : \mathbb{R}^2 \to \mathbb{R}$, ∇f cont., $\Gamma_c \neq \phi$, $c \in \mathbb{R}$



Rk If $\Gamma \begin{cases} x_1 = x_1(t) \\ x_2 = x_2(t) \end{cases}$ $t \in [0, T)$ is a diff-able parametric curve

Then the tangent to Γ is given by

$$\frac{d}{dt}(x_1(t), x_2(t)) = (x_1'(t), x_2'(t)) \qquad (T)$$

$(x_1, x_2 : [0, T) \to \mathbb{R}$, diff-able functions of $t)$

Proof of Th 2: If $\nabla f(x) = 0_{\mathbb{R}^2}$ ($0_{\mathbb{R}^2} \perp$ any direction) trivially ✓

If $\nabla f(x) \neq 0_{\mathbb{R}^2}$ (ass $\frac{\partial f}{\partial x_2}(x) \neq 0$ (*))

Apply Implicit Fct. Thm to $f(x_1, x_2) - c$

$\exists$ (locally around x) φ : $x_2 = \varphi(x_1)$

$$f(x_1, \varphi(x_1)) - c = 0 \quad (*)$$

Now use the CHAIN rule

$$\frac{d}{dx_1} f(x_1, \varphi(x_1)) = \nabla f(x_1, \varphi(x_1)) \cdot \frac{d}{dx_1}(x_1, \varphi(x_1))$$

Implicit diff of φ

$$= \nabla f(x_1, \varphi(x_1)) \cdot \left(1, -\frac{\partial f / \partial x_1}{\partial f / \partial x_2}\right) = 0$$

§ 5.3. The Lagrange Multiplier Method

Constr. Optimiz $\left.\begin{array}{r} f(x_1,x_2) \to \min \\ \text{subj to } g(x_1,x_2) = 0 \end{array}\right\}$

Lagr. Multipl. Meth (Idea): reduce this to unconstrained optimization (with an additional variable λ called Lagrange multiplier)

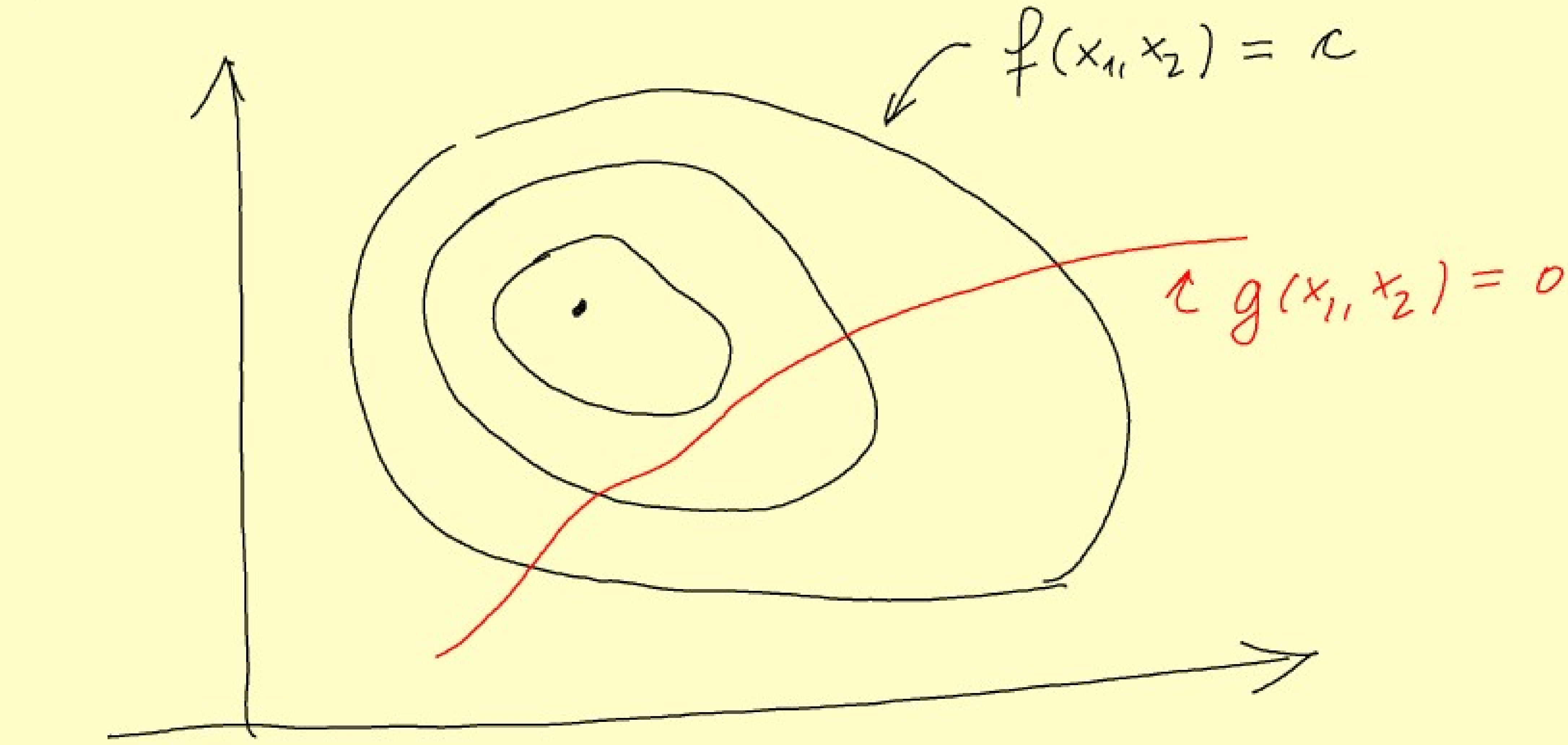
$$L(x_1,x_2,\lambda) := f(x_1,x_2) - \lambda\, g(x_1,x_2) \to \min$$

Def: $f, g : \mathbb{R}^2 \to \mathbb{R}$ cont. diff-able
$x^* = (x_1^*, x_2^*)$ is called local conditional min if $g(x^*) = 0$ and $f(x^*) < f(x)$ $\forall x$ with $g(x) = 0$.

$\boxed{\text{T}}$ 3 (Lagrange Multiplier Method)
$f, g : \mathbb{R}^2 \to \mathbb{R}$ cont. diff-able, x^* cond. min.
Then there exists $\lambda^* \in \mathbb{R}$ such that $(x_1^*, x_2^*, \lambda^*)$ is a local (unconditional) min for $L(x_1, x_2, \lambda) = f(x_1,x_2) - \lambda g(x_1,x_2)$,
that is $\nabla L(x_1,x_2,\lambda) = 0_{\mathbb{R}^3} \iff \begin{cases} \frac{\partial L}{\partial x_1}(x^*, \lambda^*) = 0 \\ \frac{\partial L}{\partial x_2}(x^*, \lambda^*) = 0 \\ \frac{\partial L}{\partial \lambda}(x^*, \lambda^*) = 0 \iff g = 0 \end{cases}$

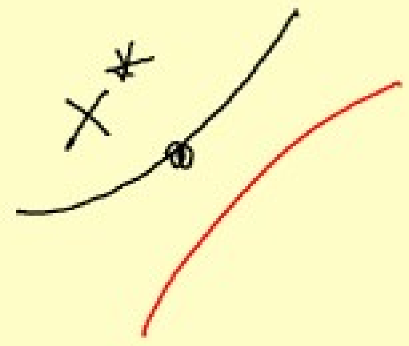
w.r.t. x_1, x_2, λ

Lagrange Multipl. Meth. (Idea of Proof):



Geometric Insight: At the cond. min. point the $g=0$ and $f=c$ contour lines are <u>tangent to eachother</u>.

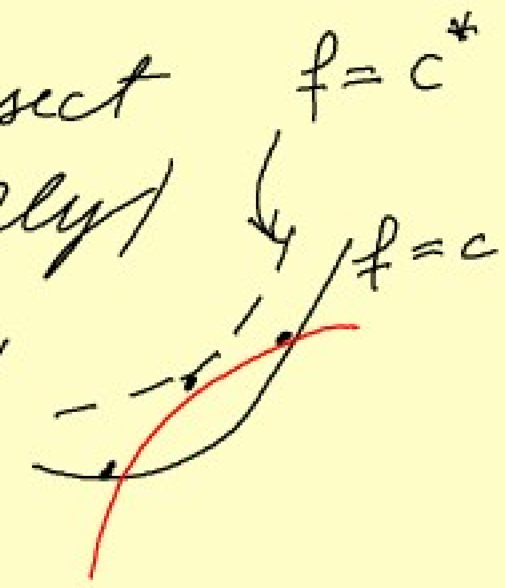
if the two wouldn't touch



Then x^* does <u>not</u> satisf the cond.

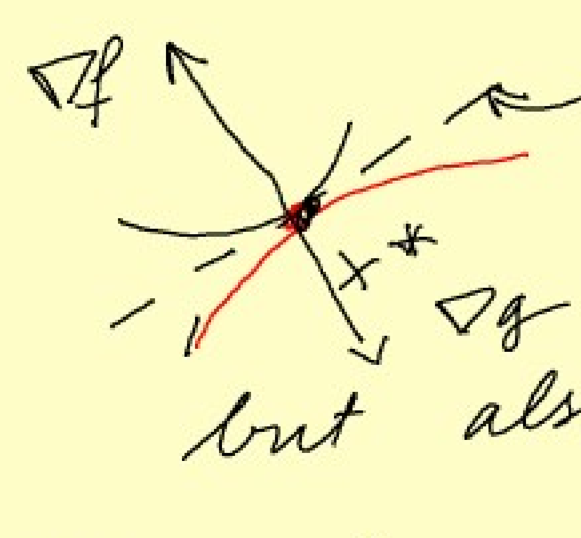
$(g(x^*) \neq 0)$

If they intersect (nontangentially) then $\cap$ points are <u>not</u> optimal



$c^* < c$

The only "good" case is



From '□'3 $\nabla f \perp$ tangent but also $\nabla g \perp$ tangent

(i.e. ∇f and ∇g are colinear)

Mathematically $\underset{\exists \lambda^* \in \mathbb{R}}{}$ $\nabla f(x^*) = \lambda^* g(x^*) \Leftrightarrow$

$$\left.\begin{array}{l} \overbrace{\nabla f(x^*) - \lambda^* g(x^*) = 0}^{\nabla_{x_1,x_2} L = 0} \\ g = 0 \quad \left(\frac{\partial L}{\partial \lambda} = 0\right) \end{array}\right\}$$

$x_1, x_2, \lambda \; \nabla L = 0$

Th 4. ([D. Popa, Th 6.10.1])

General result.

The Lagrange Multiplier Method works in dimension $d > 2$ and with multiple (compatible!) constraints.

Ex. The box of minimal surface and volume $= 1$.

$$f(x_1, x_2, x_3) := \underbrace{2x_1x_2 + 2x_2x_3 + 2x_3x_1}_{\text{surface of box}} \longrightarrow \min$$

$$g(x_1, x_2, x_3) := \overbrace{x_1x_2x_3}^{\text{volume}} - 1 = 0$$

To find (possible) local minima solve the min problem for

$$L(x_1, x_2, x_3, \lambda) = (2x_1x_2 + 2x_2x_3 + 2x_3x_1) - \lambda(x_1x_2x_3 - 1)$$

$$\frac{\partial L}{\partial x_1} = 0 \quad \frac{\partial L}{\partial x_2} = 0 \quad \frac{\partial L}{\partial x_3} = 0 \quad \frac{\partial L}{\partial \lambda} = 0$$

system of 4 eqns & with 4 unknowns (x_1, x_2, x_3, λ).