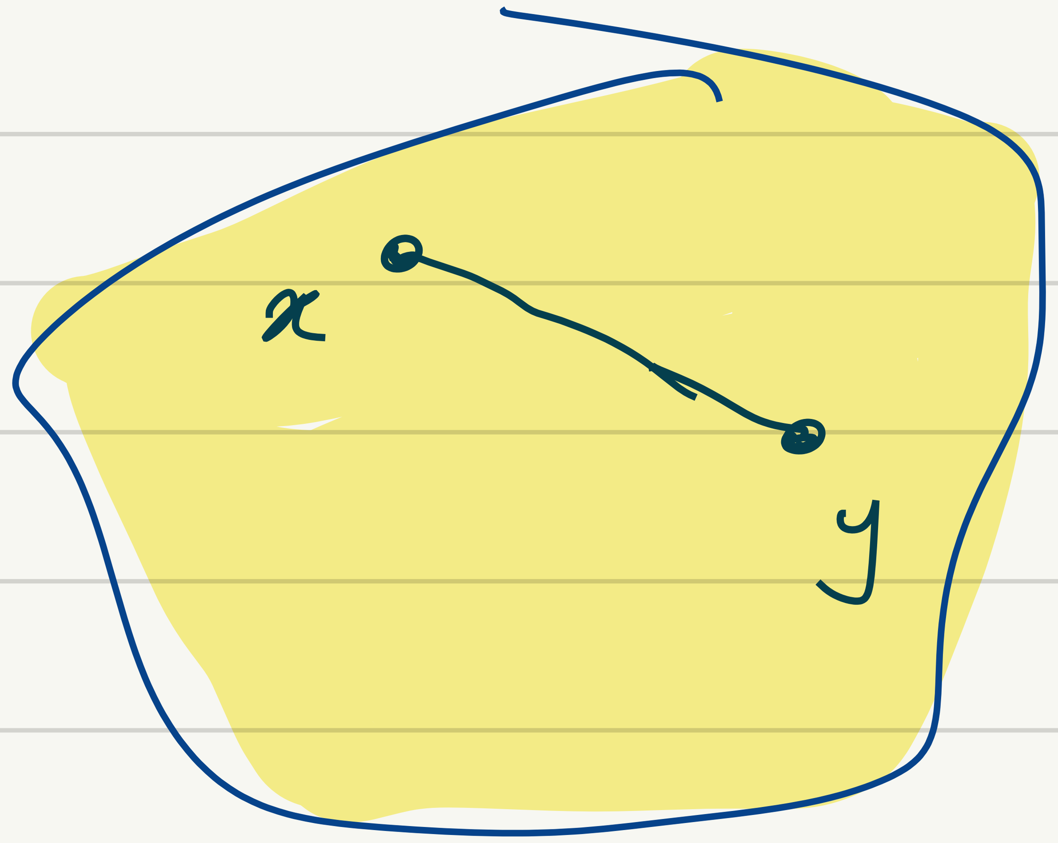
COL726
Optimization

Convex set: $S \subseteq \mathbb{R}^n$

s.t. $\vec{x}, \vec{y} \in S$, $\theta \in [0,1]$

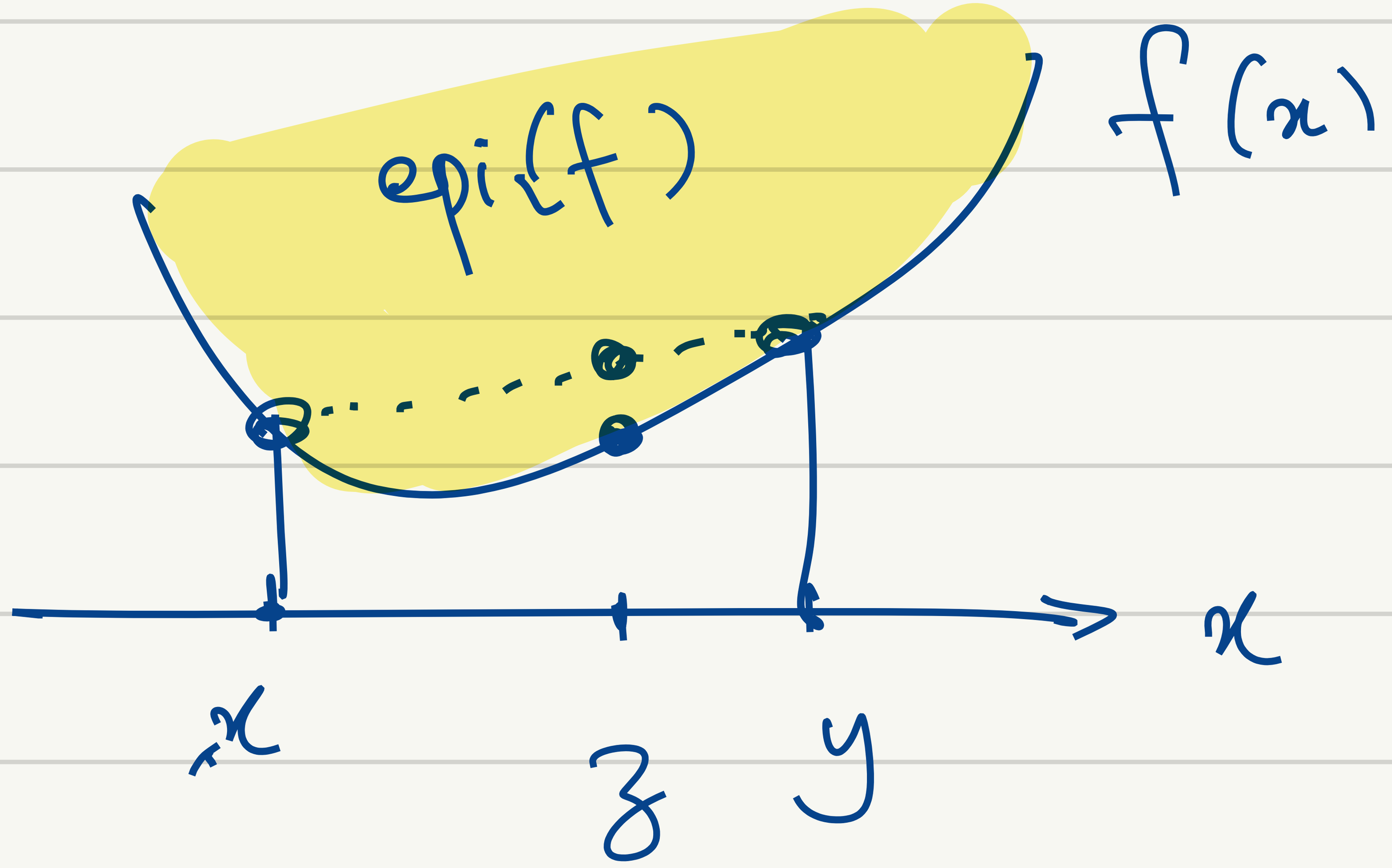
$\Rightarrow \quad \theta \vec{x} + (1-\theta)\vec{y} \in S$



Convex function: $f : \mathbb{R}^n \to \mathbb{R}$

s.t. $\vec{x}, \vec{y} \in \mathbb{R}^n$, $\theta \in [0,1]$

$\Rightarrow f(\theta x + (1-\theta) y) \le \theta f(x) + (1-\theta) f(y)$

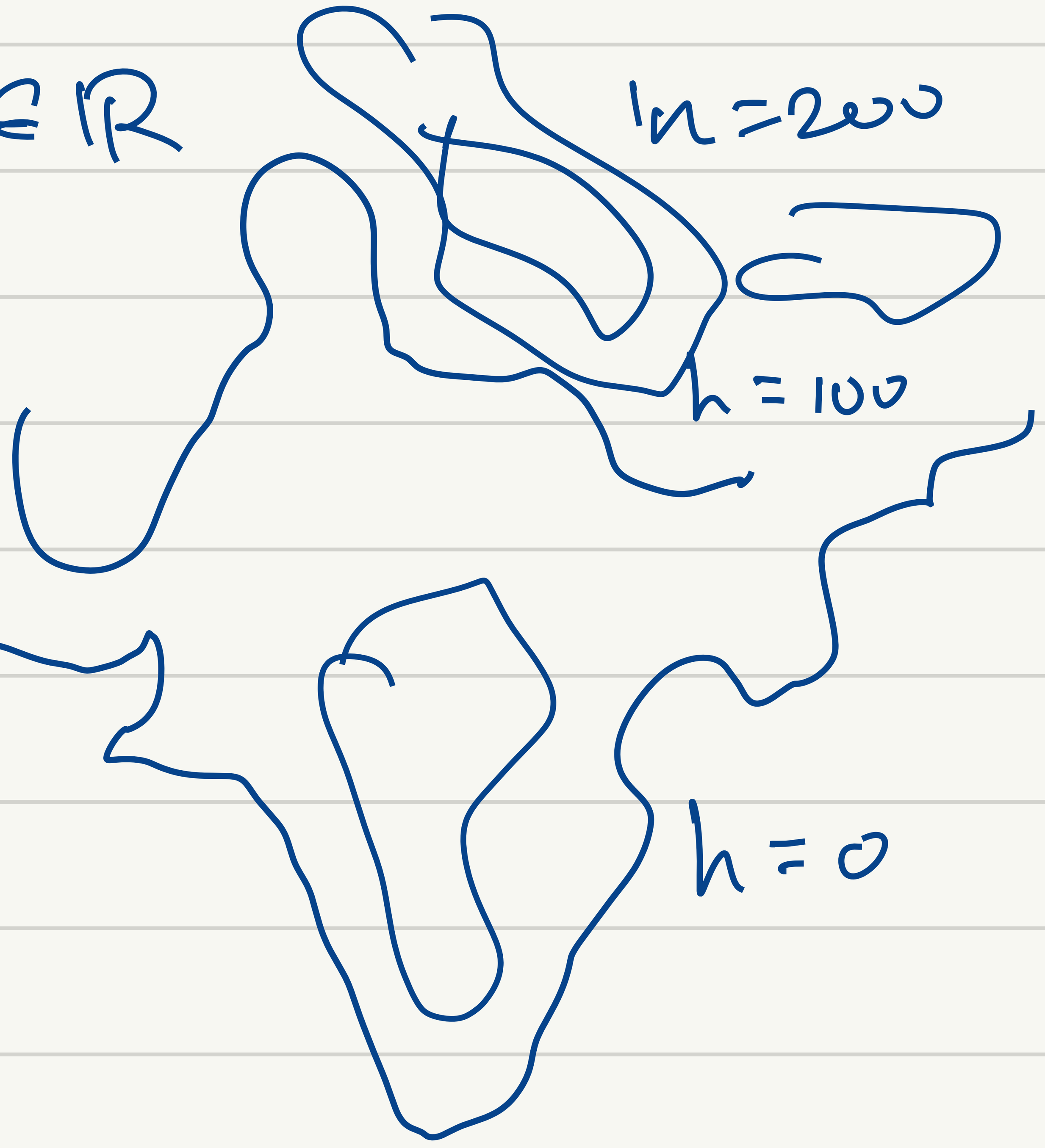


$z = \theta x + (1-\theta) y$

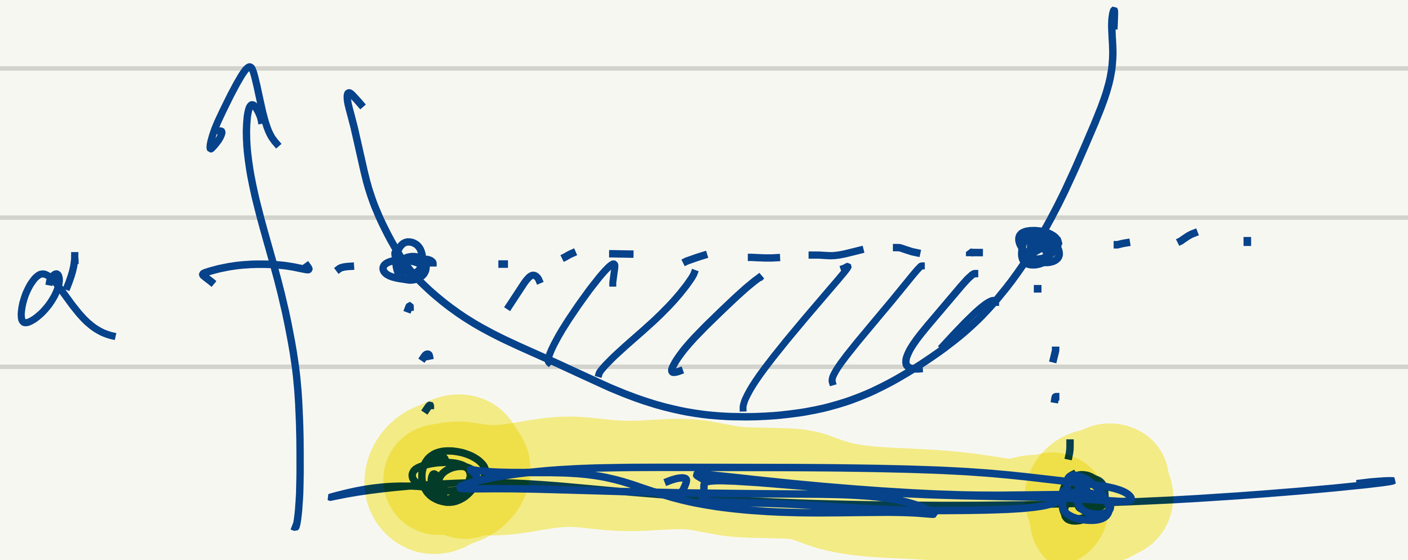
$\mathrm{epi}(f) \subseteq \mathbb{R}^{n+1}$

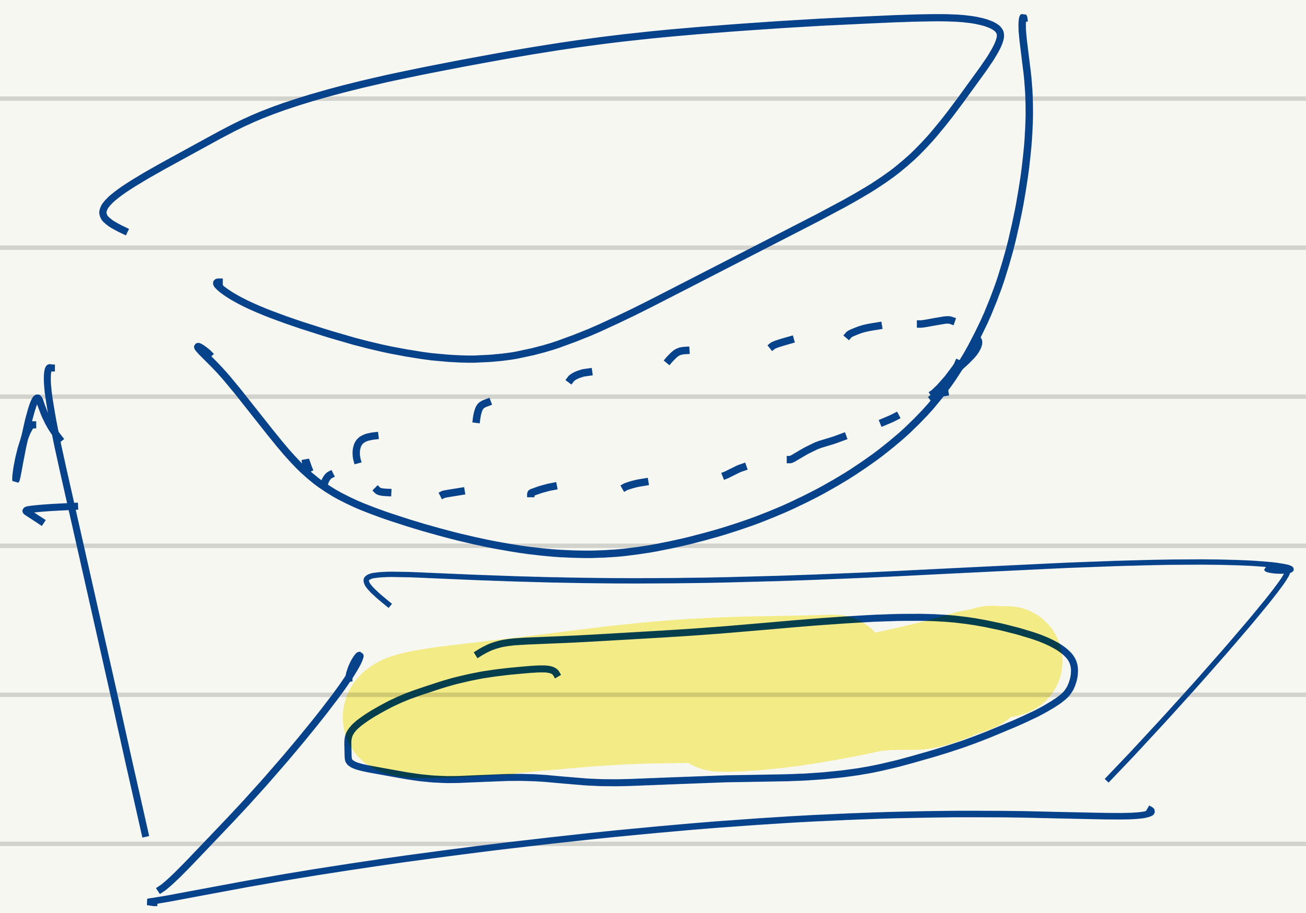
Sublevel set

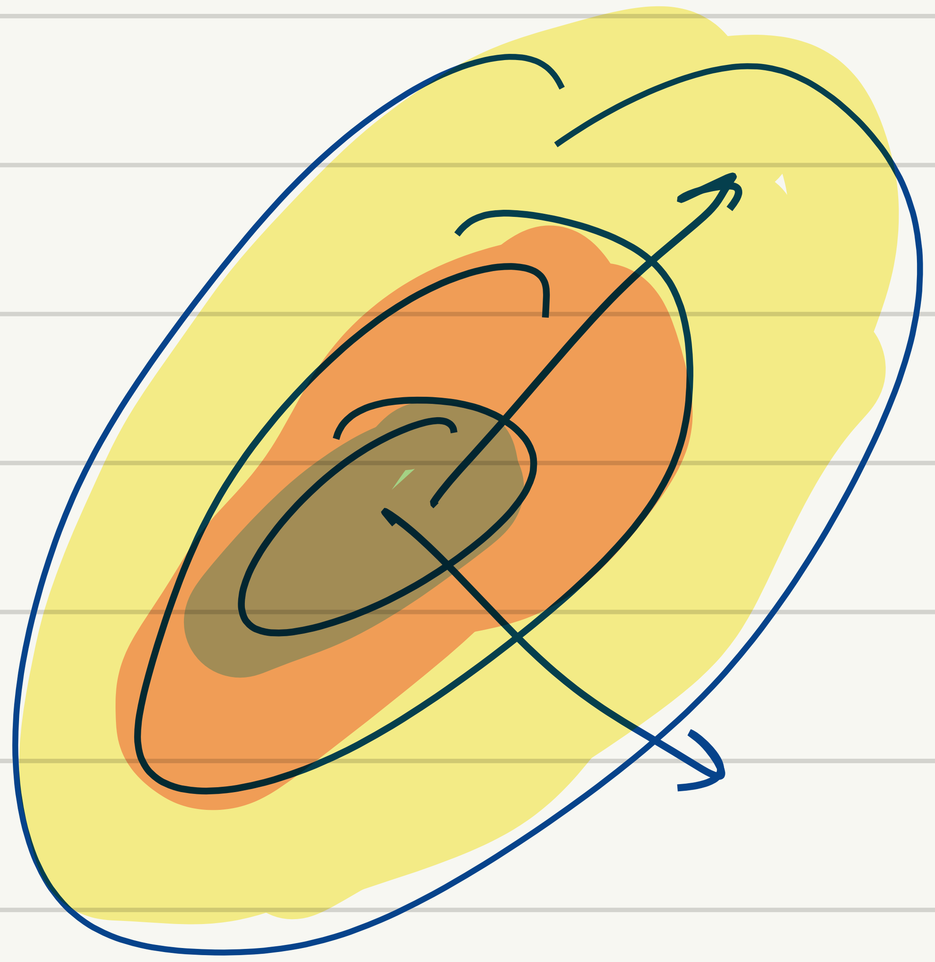
level set at $\alpha \in \mathbb{R}$
$= \{x : f(x) = \alpha\}$



Sublevel set $C_\alpha = \{x : f(x) \le \alpha\}$

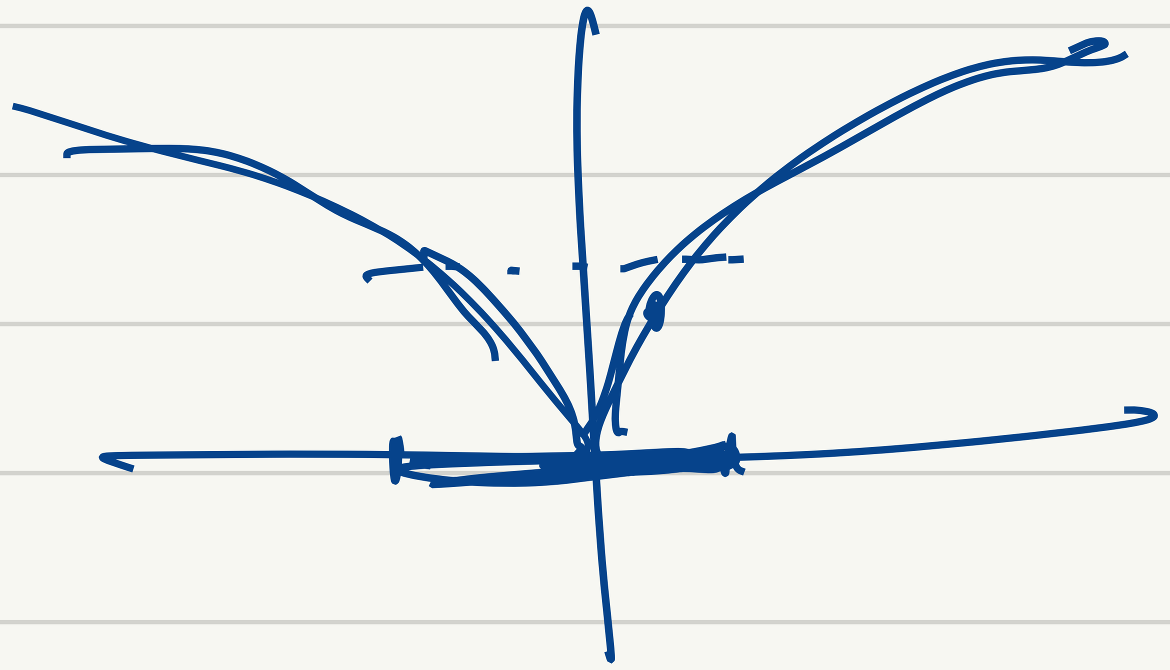






f is convex $\Rightarrow$

all sublevel sets are convex



$\sqrt{|x|}$

quasiconvex : all sublevel sets convex

Optimization problems

objective function

standard form: minimize $f_0(\vec{x})$

constraint functions

subject to $f_i(\vec{x}) \le 0 \quad \forall\, i = 1, \dots, m$

$h_i(\vec{x}) = 0 \quad \forall\, i = 1, \dots, p$

minimize $f_0(x)$

subject to $f_i(x) \leq 0 \quad \forall\, i = 1, \ldots m$

$h_i(x) = 0 \quad \forall\, i = 1 \ldots p$

$f(x) = y \rightarrow \underbrace{f(x) - y}= 0$

$f(x) \geq 0 \rightarrow -f(x) \leq 0$

$\max f(x) \rightarrow \min -f(x)$

$\vec{x}$ is $\boxed{\text{feasible}}$ if it satisfies constraints

$\{\vec{x} : \vec{x} \text{ is feasible}\} = \boxed{\text{feasible set}}$ or $\boxed{\text{constraint set}}$

$f_i(x) \leq 0 \Rightarrow \vec{x} \in$ 0-sublevel set of f_i

$h_i(x) = 0 \Rightarrow \vec{x} \in$ 0-level set of h_i

If no constraints, $\boxed{\text{unconstrained}}$ problem. $\Rightarrow$ feasible set $= \mathbb{R}^n$

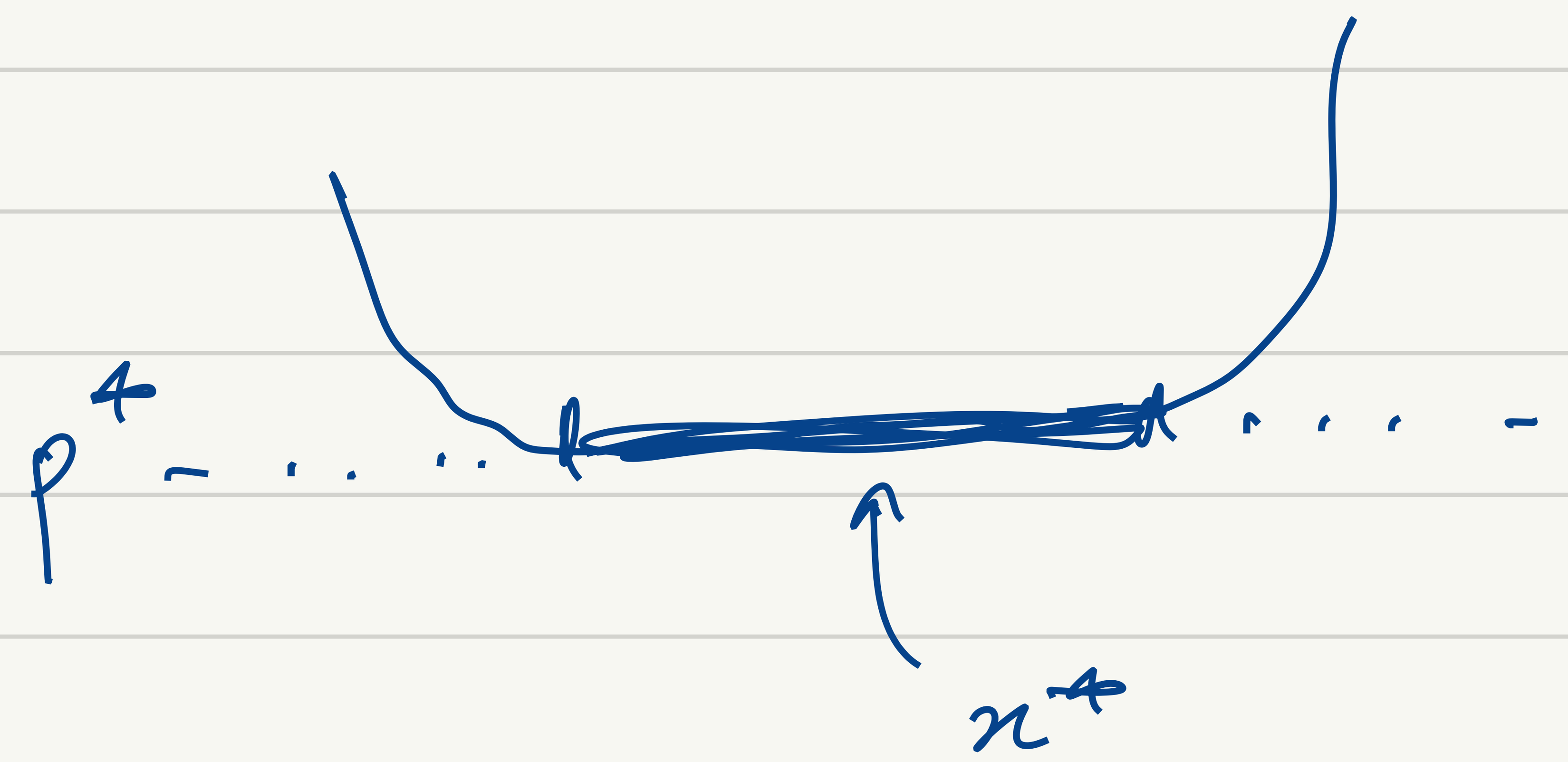
Goal: find $\inf_{\text{feasible } \vec{x}} f_0(\vec{x}) = p^* \in \mathbb{R}$: $\boxed{\text{optimal value}}$

and feasible $\vec{x}^*$ s.t. $f(\vec{x}^*) = p^*$: $\boxed{\text{optimal point}}$

globally

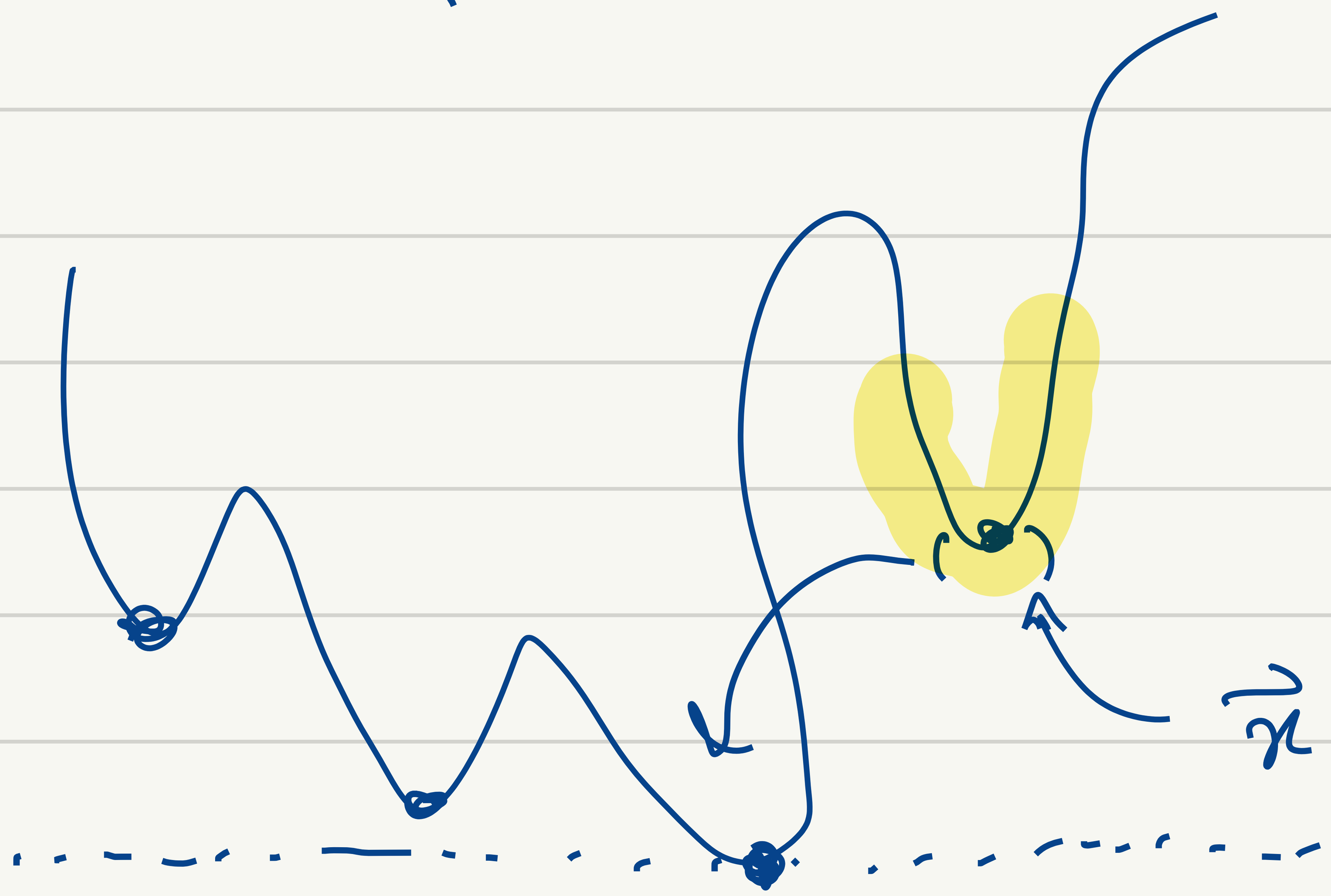
p^* may not exist: $\min x$ over $\mathbb{R}$

$\vec{x}^*$ may not exist: $\min e^x$ over $\mathbb{R}$



If f_0 is $\boxed{\text{coercive}}$ on feasible set : $f_0(\vec{x}) \to \infty$ if $\|\vec{x}\| \to \infty$

then p^*, $\vec{x}^*$ exist



$\vec{x} \stackrel{?}{=} \vec{x}^*$

If $f(x^*) \leq f(x)$ for all nearby feasible x then $\boxed{\text{locally optimal point}}$

Convex optimization

$$\min \; f_0(x)$$

$$\text{s.t. } f_i(x) \le 0 \;, \quad i = 1 \dots m$$

$$h_i(x) = 0 \;, \quad i = 1 \dots p$$

Problem is $\boxed{\text{convex}}$ if

f_0, f_i are convex,

h_i are affine : $h_i(x) = \vec{a}_i^T \vec{x} + b_i$

$\rightarrow$ sublevel sets of f_i are convex

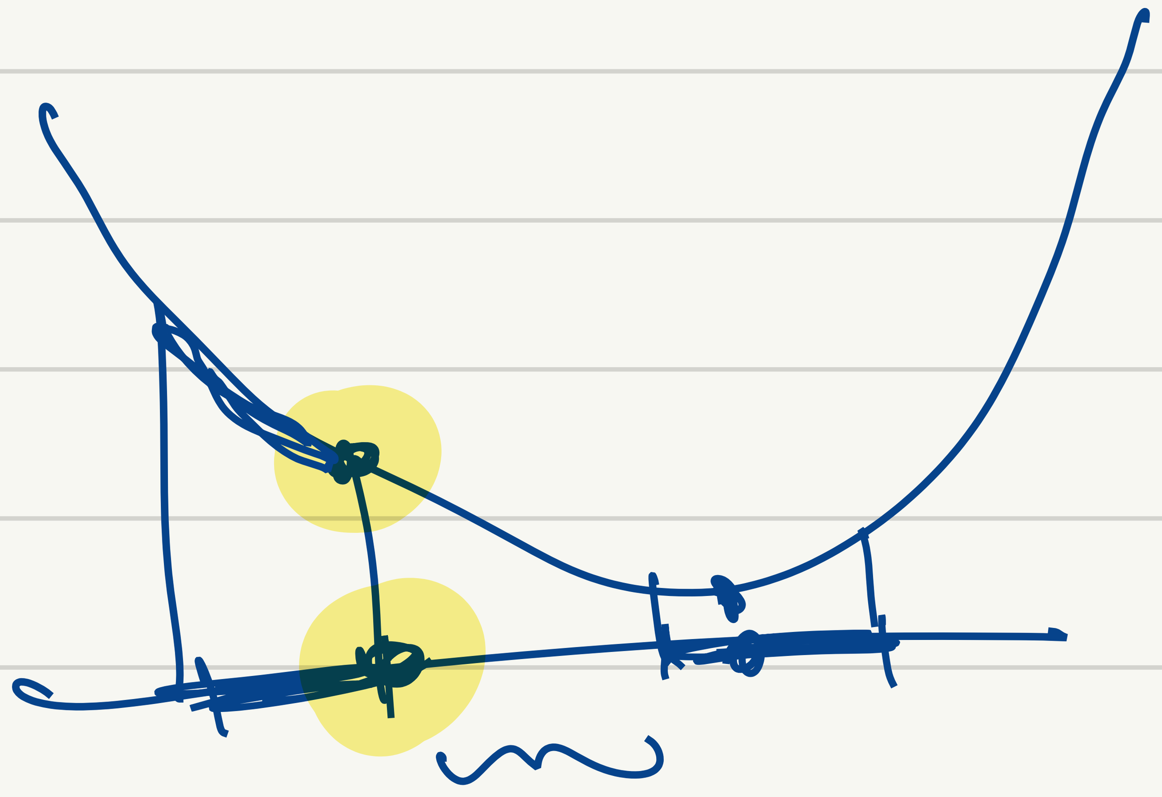
level sets of h_i are hyperplanes

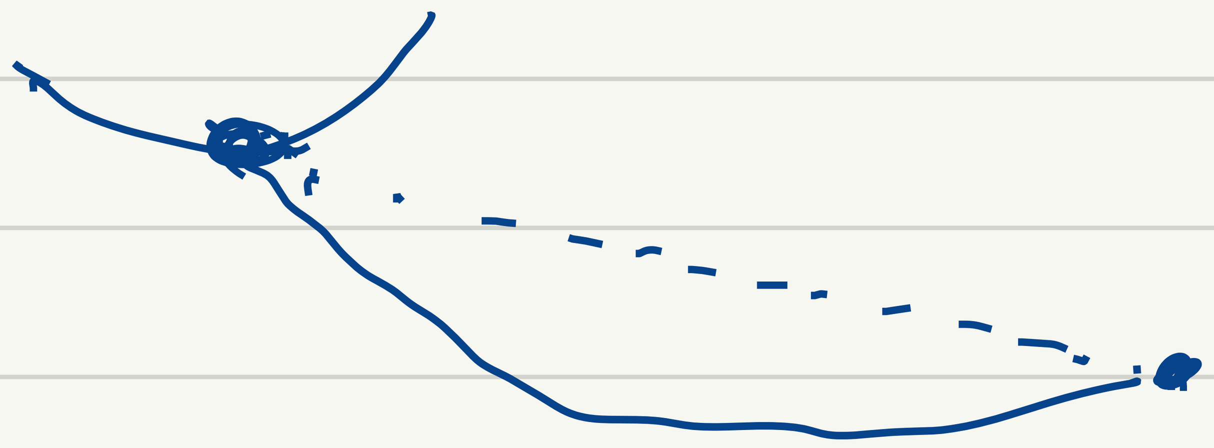
$\Rightarrow$ feasible set is convex.

$\boxed{\text{local optima are global optima!}}$

obj. fn. is convex,

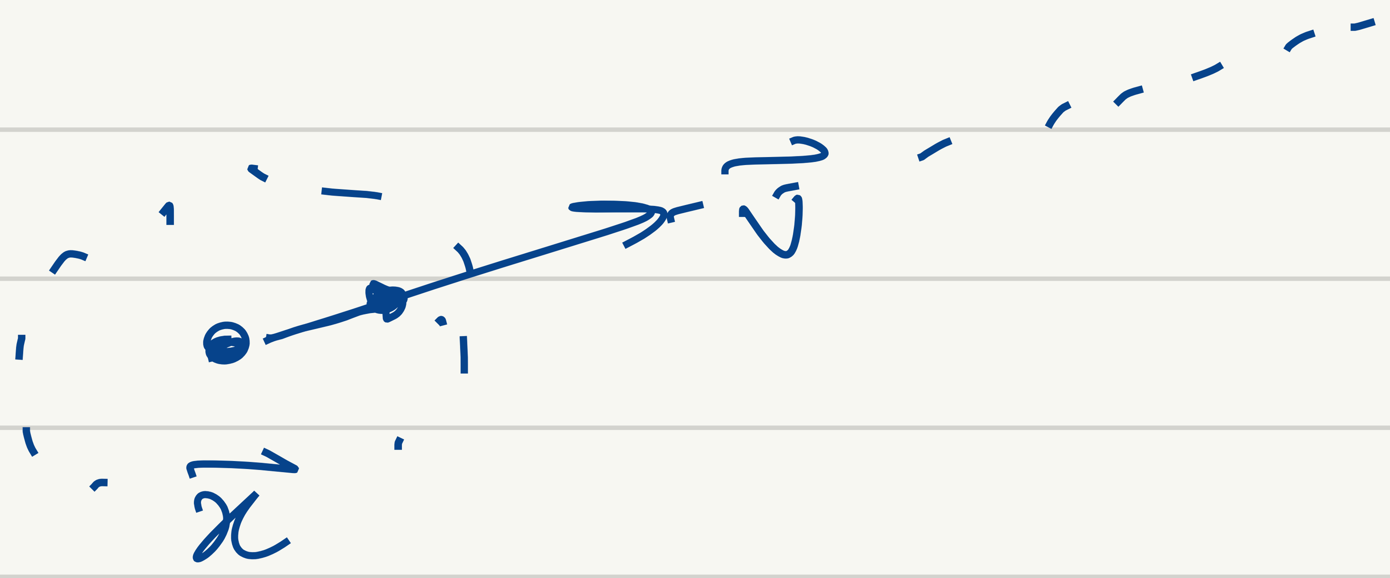
feasible set is convex

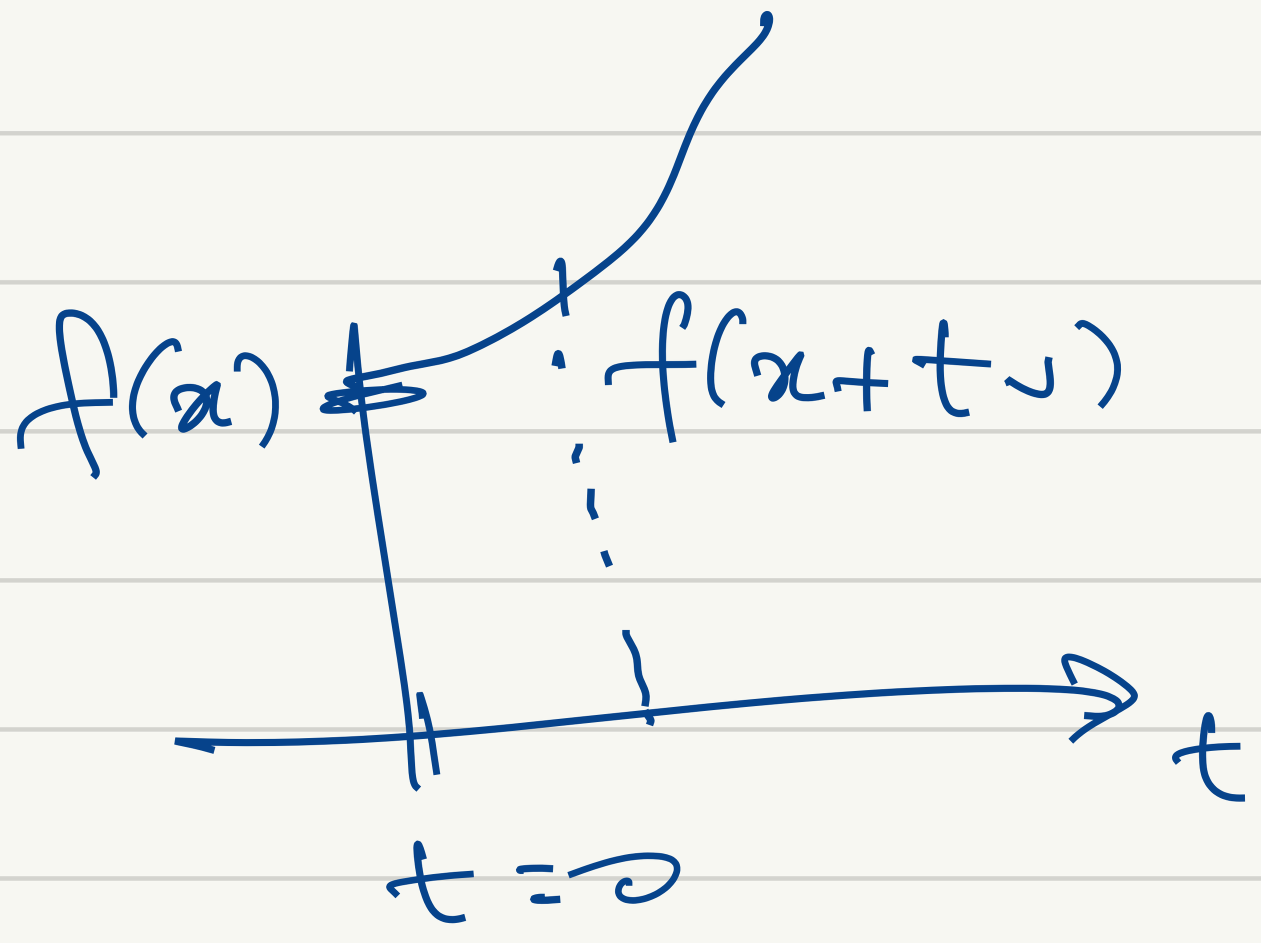




Assume f is differentiable

$\vec{x}$ local min?





$\gamma(t) = \vec{x} + t\vec{v}$

$\hat{f}(t) = f(\gamma(t))$

$\hat{f}(t) \overset{?}{\geq} \hat{f}(0)$

$\hat{f}'(t) = \nabla f(\gamma(t))^T \gamma'(t)$

$= \underbrace{\nabla f(\vec{x})^T \vec{v}}$

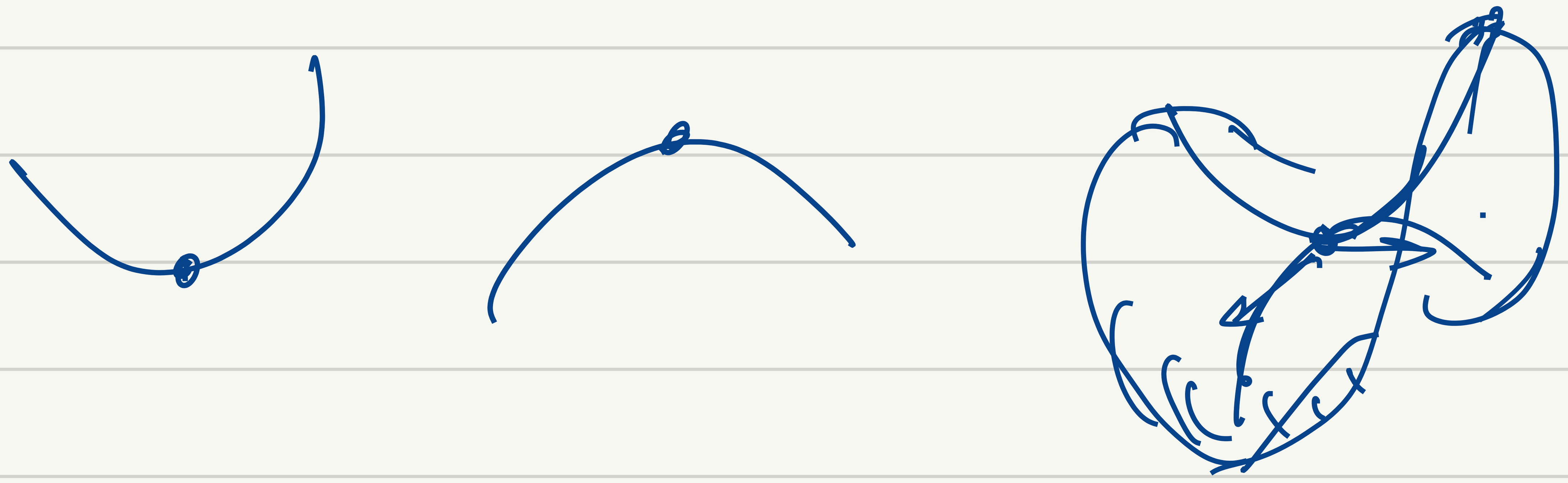
If $\hat{f}'(0) < 0$ then $\hat{f}(t) < \hat{f}(0)$

$\Rightarrow$ x is not local min

$\vec{x}$ is local min $\Rightarrow$ $\hat{f}'(0) \geq 0$ $\Rightarrow$ $\nabla f(x)^T \vec{v} \geq 0$ for all $\vec{v} \in R^n$

$\Rightarrow$ $\nabla f(\vec{x}) = 0$ $\boxed{\text{first-order necessary condition}}$

If nonconvex f, $\nabla f(x) = 0 \Rightarrow$ could be local min, local max, saddle point, . .



$$\hat{f}(t) = f(\vec{x} + t\vec{v})$$

$$\hat{f}'(0) = 0$$

for all $\vec{v} \in \mathbb{R}^n$

$$\hat{f}''(0) = \vec{v}^T \nabla^2 f(\vec{x}) \vec{v} \geq 0 \Rightarrow \nabla^2 f(\vec{x}) \succeq 0$$

$\boxed{\text{Second-order necessary cond.}}$

$\nabla^2 f(\vec{x}) \succ 0$: $\boxed{\text{second order sufficient cond.}}$

$$f(x) = x^4$$

$$f(x) = -x^4$$

Equivalent problems

$A\vec{x} = \vec{b} \iff \underbrace{MA}\vec{x} = M\vec{b} \iff \overbrace{AN}\vec{y} = b$

$\vec{x} = N\vec{y}$

$\min f(x) \iff$ **change of var.** $\phi : \mathbb{R}^n \to \mathbb{R}^n$

$\vec{x} = \phi(\vec{z})$, $\min f(\phi(\vec{z}))$

preserves convexity if ϕ is affine: $\vec{x} = A\vec{z} + \vec{b}$

$\iff$ **transformation of fn**: $\psi : \mathbb{R} \to \mathbb{R}$ monotonically increasing on range(f)

$\min \psi(f(x))$

$\min \|A\vec{x} - \vec{b}\| \iff \min \|A\vec{x} - \vec{b}\|^2$

preserves convexity if ψ is convex

$\iff \min_{x,y} f(x,y) = \min_x \underbrace{\left(\min_y f(x,y)\right)}$

$g(y)$

$\min_{x,y} f(x) + g(y) + \underbrace{\|x - y\|^2}$

$$\min_{x,y} \; f(x) + \underbrace{q(y) + \|x - y\|^2} \qquad q(y) = \tfrac{1}{2} y^T P y + q^T y + r$$

$$\tfrac{1}{2} y^T (P + 2I) y + (q - 2x)^T y + x^T x$$

$$\min_{x} \; f(x) + q_{1/2}(x) \qquad y^*(x) = (P + 2I)^{-1}(2x - q)$$

Unconstrained minimization algorithms

min $f(\vec{x})$ over all $\vec{x} \in \mathbb{R}^n$

Assume f is twice continuously differentiable $\Rightarrow$ $\nabla f(\vec{x})$, $\nabla^2 f(\vec{x})$ continuous in $\vec{x}$

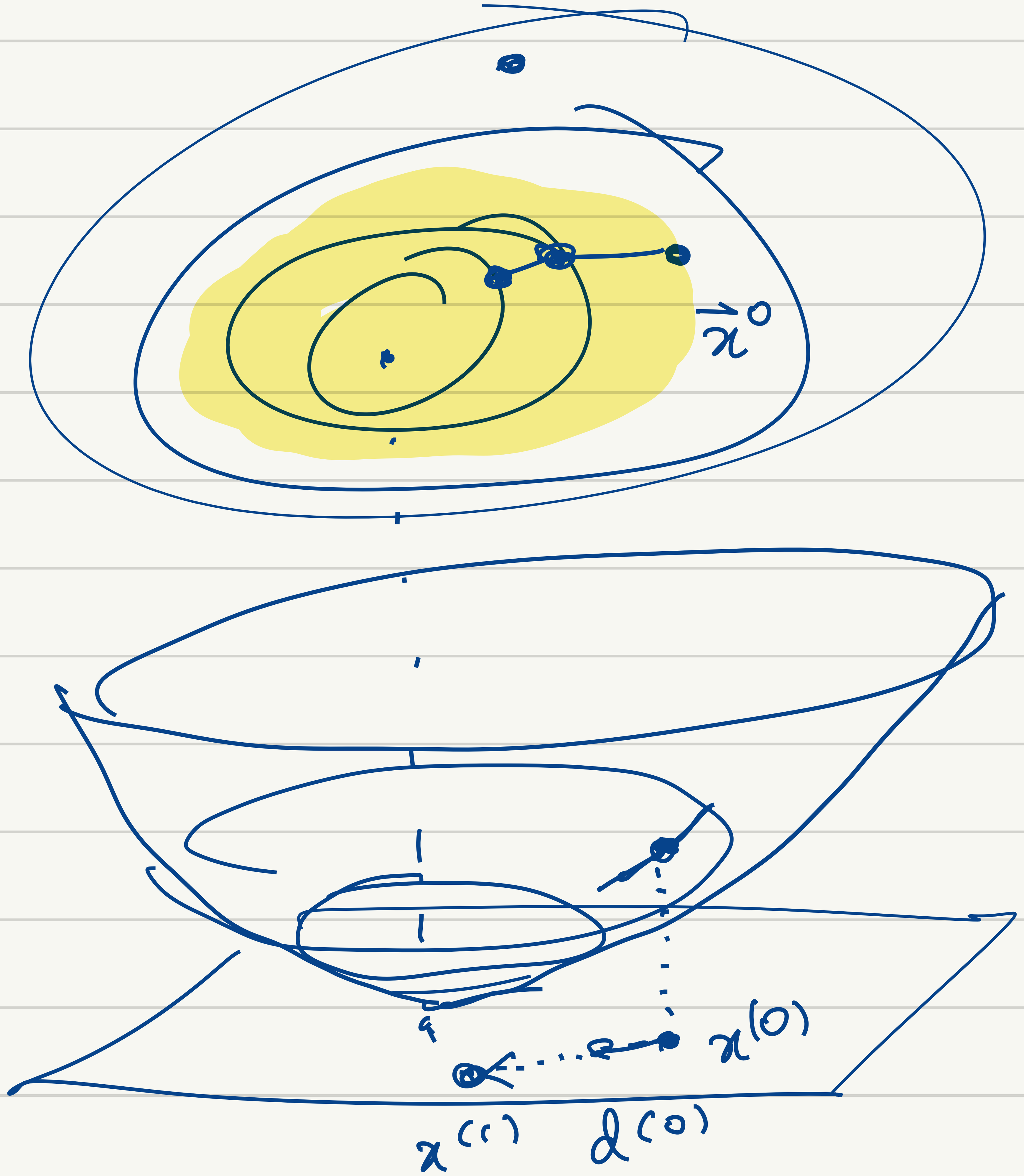
Only looking for local min: find $\vec{x}^*$ with $\nabla f(\vec{x}^*) = 0$,

$f(\vec{x}^*) \le f(\vec{x})$ for nearby $\vec{x}$

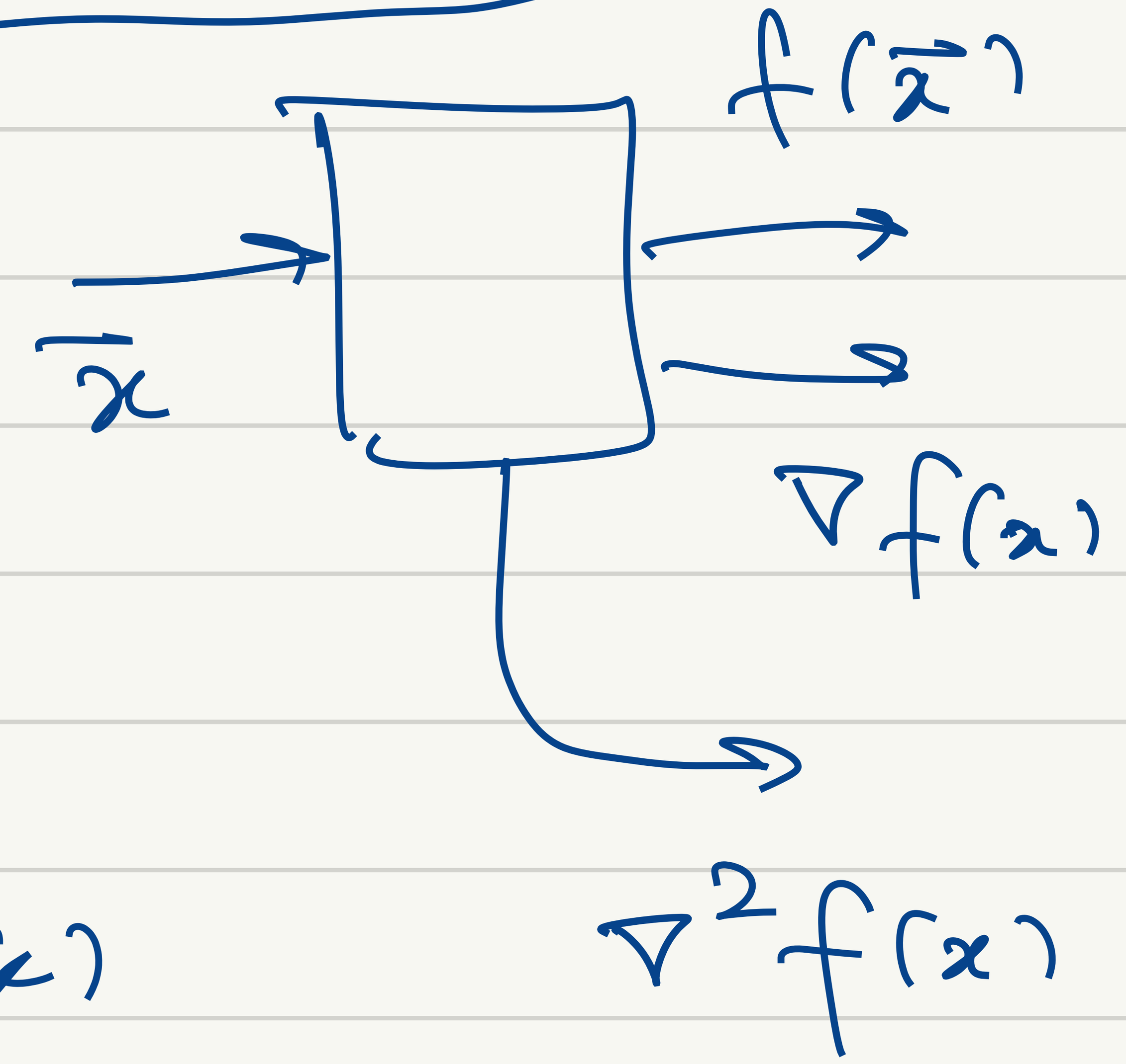
Given $\vec{x}^{(0)}$, find $\boxed{\text{minimizing sequence}}$ $\vec{x}^{(0)}, \vec{x}^{(1)}, \vec{x}^{(2)}, \dots \to \vec{x}^*$

$\boxed{\text{descent}}$: $f(\vec{x}^{(k+1)}) < f(\vec{x}^{(k)})$ unless $\vec{x}^{(k)} = \vec{x}^*$

$x^{(k)} \in S = \{ x : f(x) \le f(x^{(0)}) \}$



Descent methods / line search methods

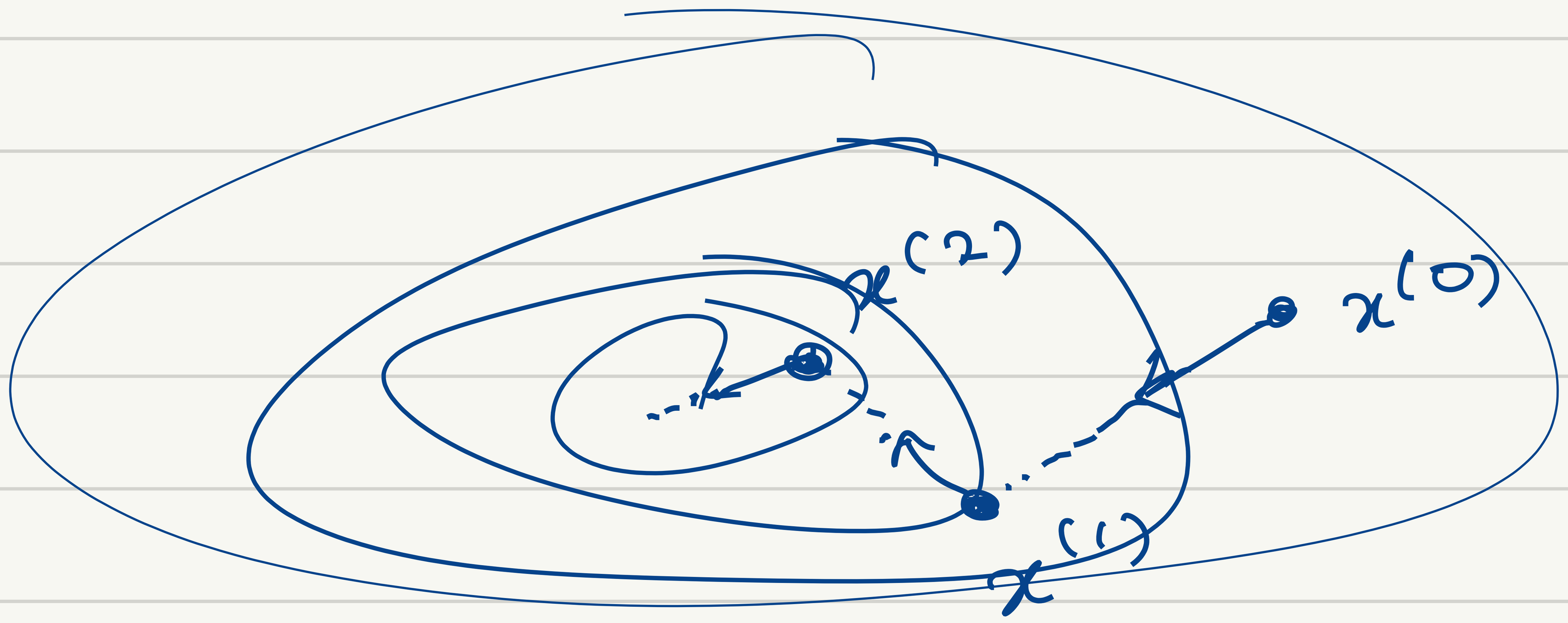


$\vec{x}^{(k)} \xrightarrow{?} \vec{x}^{(k+1)}$

Choose $\boxed{\text{search direction}}$ $\vec{d}^{(k)}$,

set $x^{(k+1)} = x^{(k)} + t^{(k)} \vec{d}^{(k)}$

for some $\boxed{\text{step size}}$ $t^{(k)} \in \mathbb{R}_{++}$



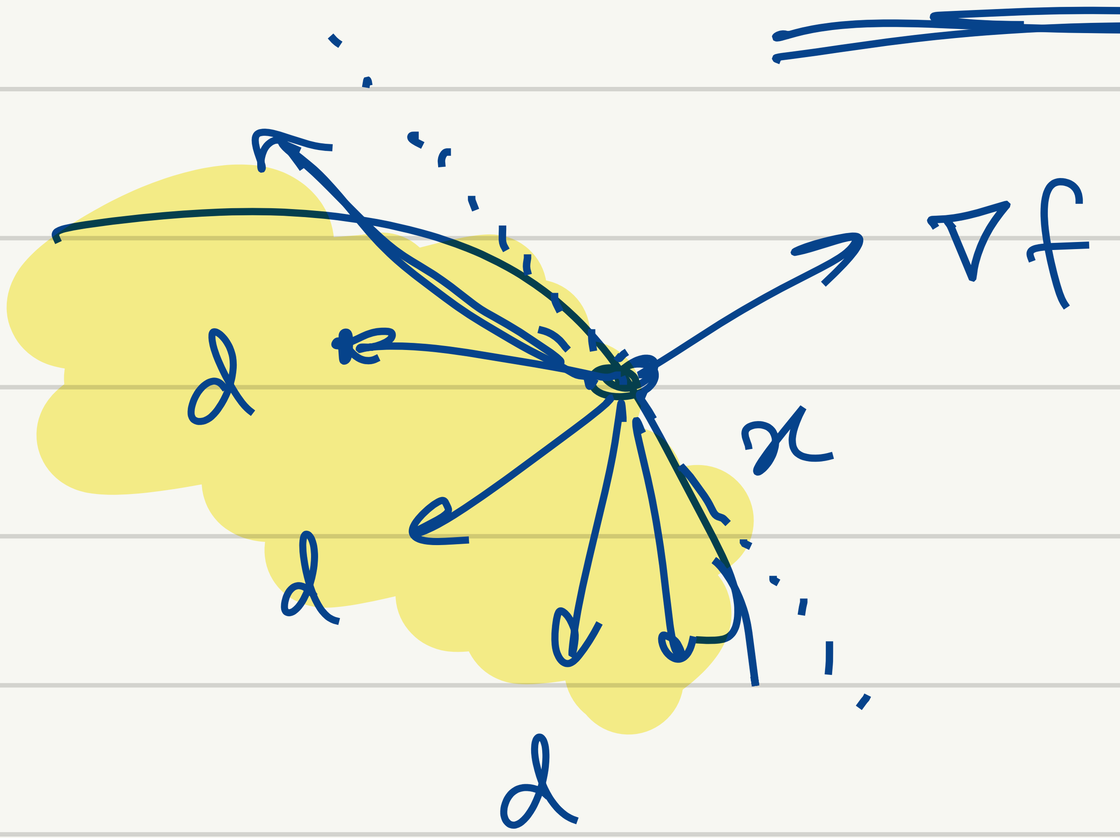
$$\vec{x}^{(k+1)} = \vec{x}^{(k)} + t^{(k)} \vec{d}^{(k)}$$

$$x^+ = x + t\,d$$

$f(x^+)$ should be as small as possible

$$f(x^+) \approx f(x) + \underbrace{t\, \nabla f(x)^T d} + O(t^2) < f(x)$$

$\Rightarrow$ $\nabla f(x)^T d < 0$: d is a $\boxed{\text{descent direction}}$



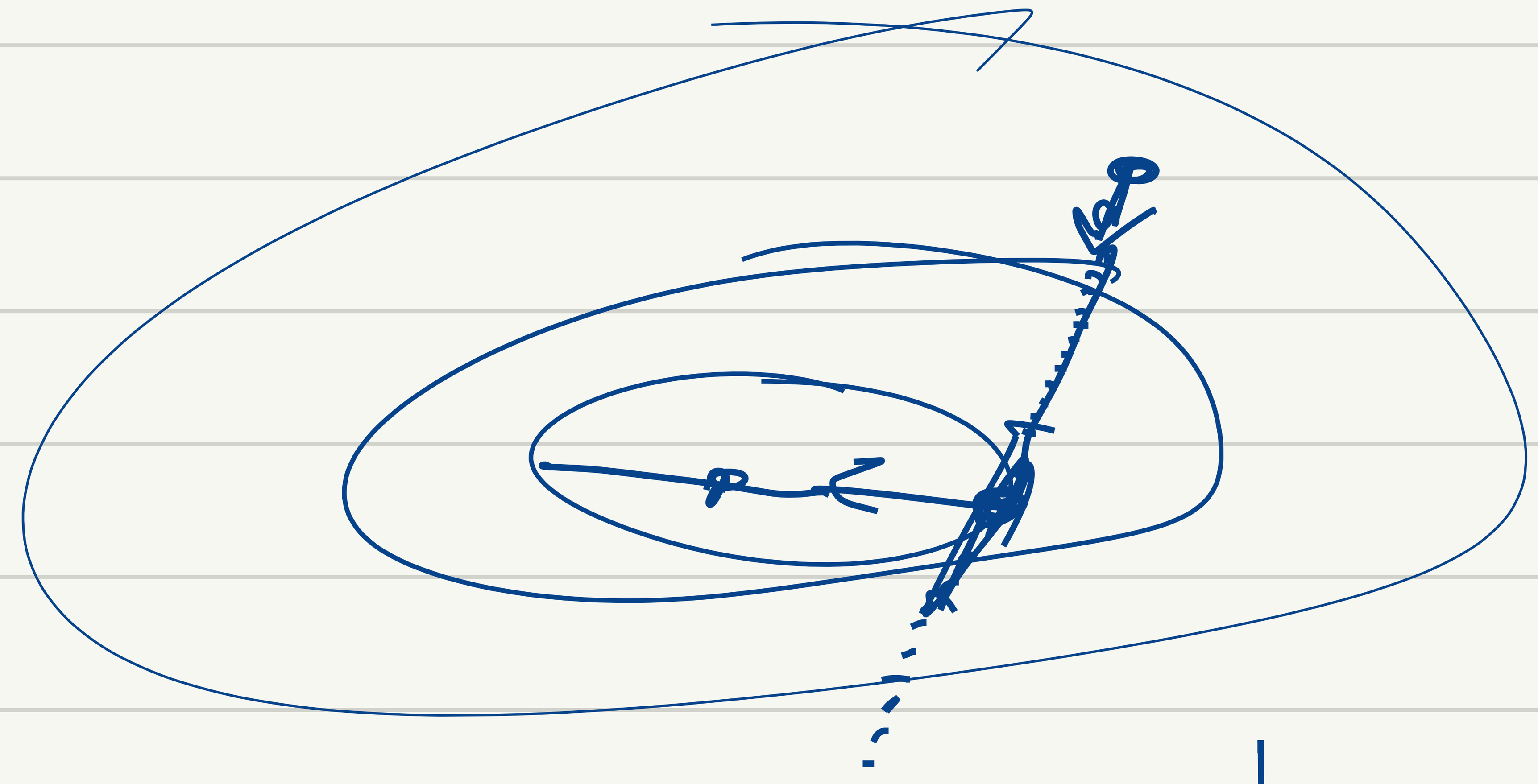
Repeat: until ...

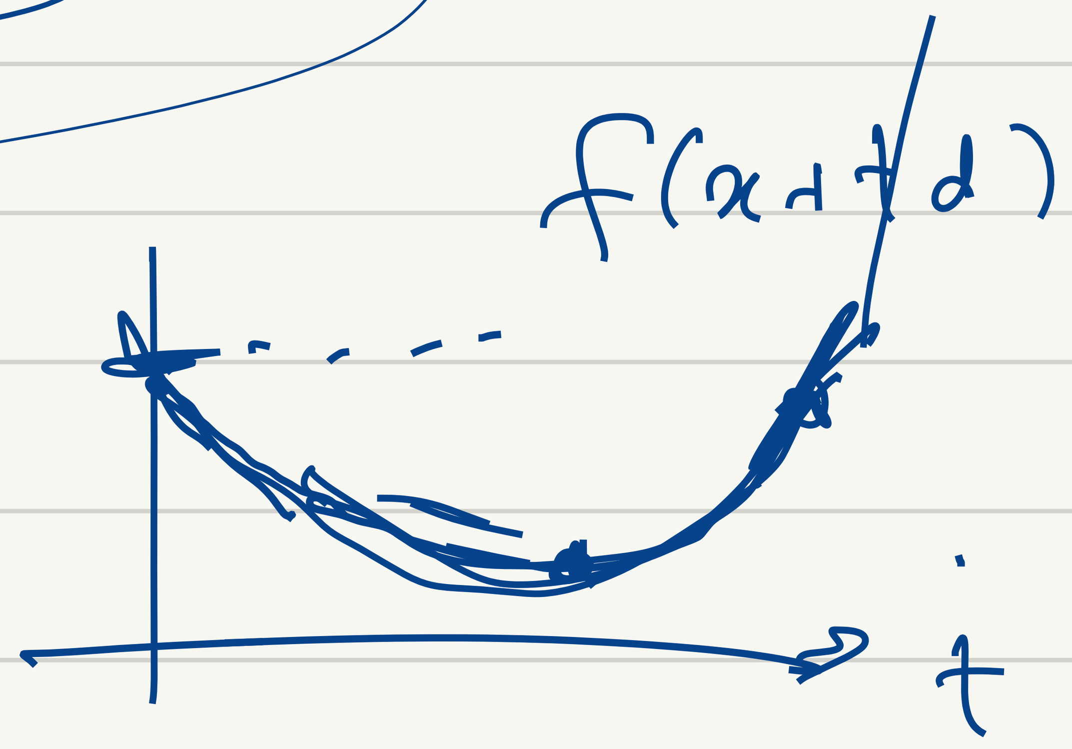
1. Choose $\vec{d}$ s.t. $\nabla f(\vec{x})^T \vec{d} < 0$
2. Choose t s.t. $f(\vec{x} + t\vec{d})$ has sufficient decrease
3. update $\vec{x} \leftarrow \vec{x} + t\vec{d}$

$\boxed{\text{line search}}$

Termination criterion : usually $\|\nabla f(\vec{x})\|$ small enough

<u>line search</u> : find t so $f(x+td)$ is small enough





$\boxed{\text{Exact line search}}$: choose t to minimize $f(\vec{x}+t\vec{d})$

$\boxed{\text{Backtracking line search}}$: pick some initial t,

if $f(\vec{x}+t\vec{d})$ is large, decrease t

$$f(\vec{x}+t\vec{d}) < f(x) + \alpha\, t\, \nabla f(x)^T \vec{d}$$

Pick t, Backtracking line search

while not $f(x + td) \le f(x) + \alpha t \nabla f(x)^T d$

$t \leftarrow \beta t$