O:
A
N

Globalübung 14
Mathe 3 WS21

Lambert Theisen
26.01.2022 ①

O: ·Nächste Woche ≙ Fragestunde mit Ulrika & mir
↳ Fragen bitte vorab per Mail

1) Parametrisierung

N Constraint Optimization

$$\min_{x\in\mathbb{R}^n} f(x) \quad \text{subject to} \begin{cases} g_i(x) \geq 0 \\ h_j(x) = 0 \end{cases}$$

$$\boxed{\begin{aligned} f &: \mathbb{R}^n \to \mathbb{R} \\ g &: \mathbb{R}^n \to \mathbb{R}^m \\ h &: \mathbb{R}^n \to \mathbb{R}^q \end{aligned}}$$

$\min_{(x,y)\in\mathbb{R}^2} f$
mit
$2x \leq y \;\hat{=}\; \tilde{g}(x)$
$\Leftrightarrow 0 \leq \underbrace{y - 2x}_{g(x)}$

LP (Linear Progr.): f, g, h linear

NLP (Nonlinear Prog.): f oder g oder h nicht linear

Feasability Region $\mathcal{X}$ (chi)**:** $\min_{x\in\mathcal{X}} f(x)$ mit $\mathcal{X} = \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, h_j(x) = 0\}$

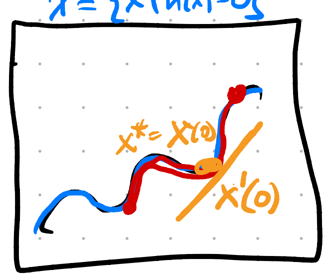
Minimizer: $x^* = \operatorname{argmin}_{x\in\mathcal{X}} f(x)$

KKT-Bedingungen: Sind notwendig für kritischen Punkt (Min, Max, Sattel)
(wenn bestimmte Regularität → LICQ)

→ $\boxed{\text{Aber wo kommen die her?}}$

Equality Constraints: Betrachte $\begin{cases} \min f(x) \\ \text{s.t. } h(x) = 0 \end{cases}$

$\mathcal{X} : (-1,1) \to \mathcal{X}$
$\mathcal{X} = \{x \mid h(x) = 0\}$



→ Wir wollen natürlich ein x^* mit: $\boxed{h(x^*) = 0}$ $\boxed{A}$

→ Weiterhin: $\nabla h(x^*)$ muss parallel zu $\nabla f(x^*)$.

Warum?: Sei $\mathcal{X} : (-1,1) \to \mathcal{X} = \{x \mid h(x) = 0\}$ Param. mit
$\mathcal{X}(0) = x^*$

Nun gehe entlang $F(t) := f(\mathcal{X}(t))$. Wir haben $f(\mathcal{X}(0)) = f(x^*) = F(0)$ ist kritischer Pkt.

kritischer Punkt heißt: $F'(0)=0 \Rightarrow \frac{d}{dt} f(x(t))\Big|_{t=0} = 0$

Kettenregel

$$\Rightarrow \nabla f(\underbrace{x(0)}_{x^*})^T \cdot x'(0) = 0 \Leftrightarrow \boxed{\nabla f(x^*) \perp x'(0)}$$

Weiterhin ist ja per Definition: $h(x(t))=0 \;\; \forall t \Rightarrow \frac{d}{dt} h(x(t))\Big|_{t=0} = 0 \Rightarrow \boxed{\nabla h(x^*) \perp x'(0)}$

Daher im kritischen Punkt x^*: $\boxed{\nabla f(x^*) \parallel \nabla h(x^*) \Leftrightarrow \nabla f(x^*) = \mu^* \nabla h(x^*)}$ $\boxed{B}$

$\mu^* \in \mathbb{R}$

Jetzt clevere Idee

Definiere Lagrange function: $\mathcal{L}(x,\mu) := f(x) - \mu h(x)$ why?!

und suche ein $\mu^* \in \mathbb{R}$

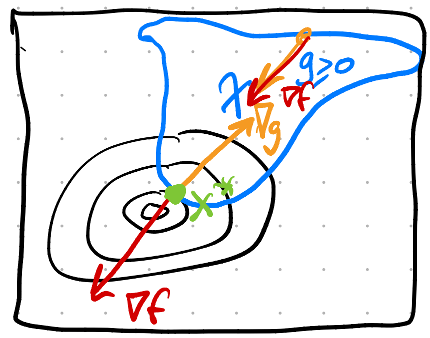
$$\nabla \mathcal{L}(x^*, \mu^*) = \begin{bmatrix} \nabla_x \mathcal{L}(x^*, \mu^*) \\ \nabla_\mu \mathcal{L}(x^*, \mu^*) \end{bmatrix} = \begin{bmatrix} \nabla f(x^*) - \mu^* \nabla h(x^*) \\ -h(x^*) \end{bmatrix} \stackrel{!}{=} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{array}{l} \boxed{B} \\ \boxed{A} \end{array}$$

→ Das kombiniert $\boxed{A}$ & $\boxed{B}$!

Inequality Constraint

$\boxed{A}$ $\boxed{g(x^*) \geq 0}$ (Def)

I) Inactive ($g(x^*) > 0$) $\Rightarrow$ $\boxed{\nabla f(x^*) = 0}$ $\boxed{B1}$ wie in Güna



II) Active ($g(x^*) = 0$) $\Rightarrow$ $\boxed{\nabla f(x^*) = \lambda^* \nabla g(x^*)}$ $\boxed{B2}$ (Wie oben bei Eq.-Const h)

↳ siehe oben

$\boxed{\lambda^* \geq 0}$ $\boxed{C}$

→ $\lambda^* \geq 0$ muss sein sodass $\nabla f(x^*)$ und $\nabla g(x^*)$ in entgegengesetzte Richtg. für Minimum.

→ Kombination von $\boxed{A}$, $\boxed{B1}$, $\boxed{B2}$, $\boxed{C}$:

nur aktive...

$$\begin{cases} \nabla f(x^*) - \lambda^* \nabla g_i(x^*) = 0 \\ g_i(x^*) \geq 0 & \boxed{A} \\ \lambda_i^* \geq 0 & \boxed{C} \\ g_i(x^*) \lambda_i^* = 0 & \text{Komplementär} \end{cases}$$

$\boxed{B1} + \boxed{B2}$

wenn $g(x^*)=0$ B1

wenn $g(x^*)>0$ B2

weil $\lambda^* = 0$ dann ✓

Karush-Kuhn-Tucker Bedingungen (kombiniert alle Bdg. für mehrere Constraints)

KKT

5

$$\begin{cases} \nabla \mathcal{L}(x,\lambda,\mu) = \nabla f(x) - \sum_{i=1}^{m} \lambda_i \nabla g_i(x) - \sum_{j=1}^{q} \mu_j \nabla h_j(x) = 0 \\ h_j(x) = 0 \\ g_i(x) \geq 0 \\ \lambda_i \geq 0 \\ g_i(x)\,\lambda_i = 0 \end{cases}$$

mit $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}^q$. $\underbrace{\text{KKT-Point}}_{\text{kritische Punkte}} \mathrel{\hat=} (x^*, \underbrace{\lambda^*, \mu^*}_{\text{Lagrange Multipliers.}})$ wenn Bdg. erfüllt.

Nun: Berechnung der Kandidaten als KKT-Punkte?

Problem: Es gibt NLP für die das Minimum kein KKT Punkt ist.
(Nur wenn bestimmte Glattheits/Regularitäts eigenschaften erfüllt sind)

LICQ: ("Linear Independence Constraint Qualification")

Punkt $x \in \mathcal{X}$ hat LICQ falls: $\left(\{\nabla h_j(x)\}_{j=1}^{q}, \{\nabla g_i(x)\}_{i \in I_g(x)} \right)$ linear unabhängig

mit $I_g(x) = \{ i \in \{1,\dots,m\} \mid g_i(x) = 0 \}$ Menge aller aktiven Bedingungen
z.B. $I_g(x) = \{1, 2, 5, 8\}$

$\Rightarrow$ Wenn LICQ in x^* gilt, dann sind KKT FONC für kritische Punkte x^*
(1st-order necessary condition) (wie $\nabla f = 0$)
2nd $\nabla^2 f > 0 \Rightarrow$ min

In der ~~Praxis~~ HW:
I) Potentielle KKT Punkte finden
II) LICQ dort prüfen
III) Falls LICQ gilt $\Rightarrow$ Hessematrix (z.B.) untersuchen für Min/Max/Sat.

6

Demo: KKT & LICQ

Define Lagrangian

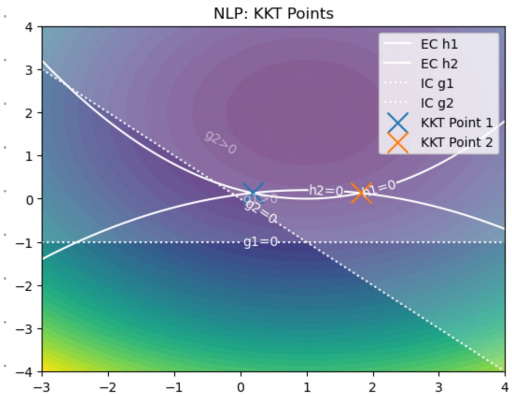
$$\mathcal{L}(x, \lambda, \mu) = f(x) - \sum_{i=1}^{m} \lambda_i g_i(x) - \sum_{j=1}^{q} \lambda_i h_i(x)$$

```
lagrangian (generic function with 1 method)
```

```
function lagrangian(x, f, g, h, λs, μs, Ig)
    # TODO: Include only active gs with Ig
    return (
        f(x)
        # - reduce(+, [g[i](x) * λs[i] for i=1:size(g)[1]], init=0)  # TODO: Include
        - reduce(+, [h[i](x) * μs[i] for i=1:size(h)[1]], init=0)
    )
end
```

$$-\mu\left(-5x_2 + (x_1 - 1)^2\right) + (x_1 - 1)^2 + (x_2 - 2)^2$$

```
lagrangian(x, f1, [], h, lambdas, mus, [])
```



Linear Independence Constraint Quality (LICQ)

Point $x \in \chi$ satisfies LICQ if:

$$\{\nabla h_j(x)\}_{j=1}^{q}, \{\nabla g_i(x)\}_{i \in I_g(x)}$$

are linearly independent. The set of active inequality constraints at point x is labelled with $I_g(x)$.

Index Set of Active Constraints:

```
Ig (generic function with 1 method)
```

```
function Ig(x,g)
    return [i for i=1:size(g)[1] if g[i](x)==0]
end
```

```
LICQ (generic function with 1 method)
```

```
function LICQ(ξ, g, Ig, h)
    set = hcat(
        # [diff(g[i](x), x) for i ∈ Ig(ξ,g)],
        [diff(h[i](x), x) for i ∈ 1:size(h)[1]]
    )
    return set
end
```