VORL. 7: NICHTLINEARE OPTIMIERUNG MIT GLEICHUNGS- UND UNGLEICHUNGS-BESCHRÄNKUNGEN $\boxed{7}$ 1

ALLGEMEINES NLP:

$$(1) \quad \boxed{\begin{array}{ll} \min\limits_{x \in \mathbb{R}^n} & f(x) \\ \text{s.t.} & c_E(x) = 0 \\ & c_I(x) \geq 0 \end{array}}$$

$c_E : \mathbb{R}^n \to \mathbb{R}^{m_E}$ E = "EQUALITIES"

$c_I : \mathbb{R}^n \to \mathbb{R}^{m_I}$ I = "INEQUALITI

NLP IST KONVEX, FALLS c_E AFFIN, f KONVEX, $(c_I)_i$ KONKAV

7.1 NICHTLINEARE INNERE PUNKTE (NIP) METHODE

TRANSFORMATION AUF STANDARDFORM DURCH EINFÜHRUNG VON SCHLUPFVARIABLEN $\xi \in \mathbb{R}^{m_I}$, $x^+ \in \mathbb{R}^n$, $x^- \in \mathbb{R}^n$

$$\boxed{\begin{array}{lll} \min\limits_{x^+, x^-, \xi} & f(x^+ - x^-) & \\ \text{s.t.} & c_E(x^+ - x^-) & = 0 \\ & c_I(x^+ - x^-) - \xi & = 0 \\ & x^+ & \geq 0 \\ & x^- & \geq 0 \\ & \xi & \geq 0 \end{array}} \iff \boxed{\begin{array}{ll} \min\limits_{x \in \mathbb{R}^n} & f(x) \\ \text{s.t.} & c(x) = 0 \\ & x \geq 0 \end{array}} \quad (2$$

$$\begin{bmatrix} x_+ \\ x_- \\ \xi \end{bmatrix} =: x$$

$$\begin{bmatrix} c_E(x_+ - x_-) \\ c_I(x_+ - x_-) - \xi \end{bmatrix} =: c(x) \qquad c : \mathbb{R}^n \to \mathbb{R}^m$$

BETRACHTE AB JETZT NLP (2).

LAGRANGE FUNKTION (MIT UNGLEICHUNGS-MULTIPLIKATOR $y \in \mathbb{R}^n$)

$$\mathcal{L}(x, \lambda, y) = f(x) - c(x)^T \lambda - y^T x$$

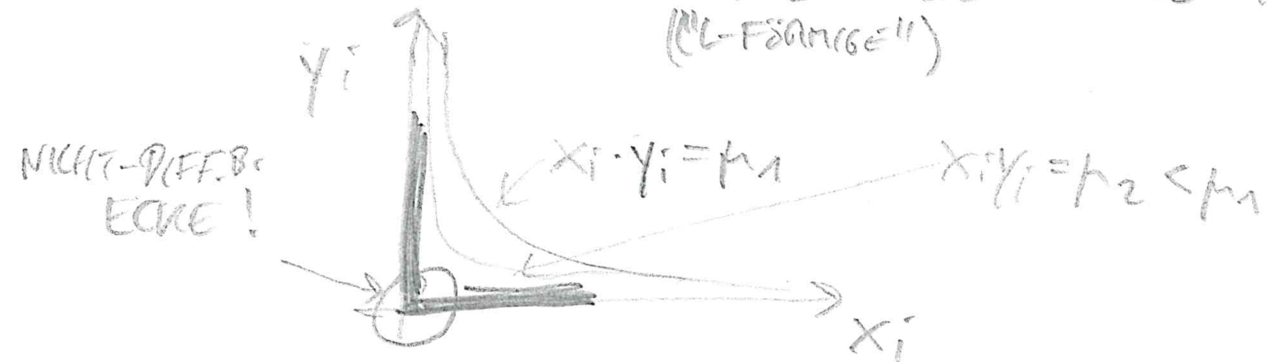
KKT-BEDINGUNGEN

$$\nabla_x \mathcal{L}(x,\lambda,y) = 0$$

$$c(x) = 0$$

$$\left.\begin{aligned} x_i &\geq 0 \\ y_i &\geq 0 \\ x_i y_i &= 0 \end{aligned}\right\} \text{FÜR } i = 1,\dots,n \Bigg\} \text{"KOMPLEMENTARITÄTS-BEDINGUNGEN"}$$

PROBLEM: KOMPLEMENTARITÄTSBED. BESCHREIBEN FÜR JEDES i NICHT-DIFFERENZIERBARE MENGE IN (x_i, y_i):
("L-FÖRMIGE")



IDEE 1: GLÄTTE ECKE DURCH PARAMETER $\mu > 0$ UND VERLANGE STATT $x_i y_i = 0$ EINFACH $\boxed{x_i y_i = \mu}$

"GEGLÄTTETE KKT BEDINGUNGEN" SIND

$$\boxed{\begin{aligned} \nabla_x \mathcal{L}(x,\lambda,y) &= 0 \\ c(x) &= 0 \\ x_i \cdot y_i - \mu &= 0 \quad i = 1,\dots,n \end{aligned}} \qquad (3)$$

$(x_i \geq 0,\ y_i \geq 0) \longleftarrow$ WIRD AUTOMATISCH ERFÜLLT, WENN $x_i y_i = \mu > 0$

FASSE WIEDER ALLES ALS EIN NULLSTELLENPROBLEM AUF:

$$z := \begin{bmatrix} x \\ \lambda \\ y \end{bmatrix} \in \mathbb{R}^{(n+m+n)}$$

$$F(z,\mu) := \begin{bmatrix} \nabla_x \mathcal{L}(x,\lambda,y) \\ \hline c(x) \\ \hline x_1 y_1 - \mu \\ \vdots \\ x_n y_n - \mu \end{bmatrix} \in \mathbb{R}^{(n+m+n)}$$

LÖSE $\boxed{F(z,\mu) = 0}$ FÜR $\mu > 0$ DURCH NEWTON-VERFAHREN, (SIEHE 6.3)

IDEE 2: STARTE MIT GROSSEM μ, VERRINGERE μ SCHRITTWEISE, [illegible]

NIP-ALGORITHMUS:

1) WÄHLE $z_{[0]}$, $\mu_1 \gg 0$, SETZE $k = 0$

2) CHECKE OB $\|F(z_{[k]}, 0)\| \leq \varepsilon$ [illegible]

3) LÖSE $\boxed{F(z, \mu_{k+1}) = 0}$, MIT NEWTON-VERFAHREN, GESTARTET IN $z_{[k]}$.
ERGEBNIS: $z_{[k+1]}$

4) SETZE $k \leftarrow k+1$

5) VERRINGERE $\mu_{k+1} = \alpha \mu_k$, $\alpha \in (0,1)$

THEOREM: DAS PROBLEM (3)

$$\boxed{F(z,\mu)=0}$$

IST ÄQUIVALENT

ZUM "BARRIERE-OPTIMIERUNGSPROBLEM"

$$\boxed{\begin{aligned} &\min_{x\in\mathbb{R}^n} \quad f(x)-\mu\sum_{i=1}^{n}\operatorname{Log} x_i \\ &\text{s.t.} \quad c(x)=0 \end{aligned}} \qquad (4)$$

BEWEIS: (3) IST GLEICH

$$\left.\boxed{\begin{aligned} \nabla f(x)-\frac{\partial c}{\partial x}(x)^T\cdot\lambda-y &= 0 \\ c(x) &= 0 \\ x_i y_i-\mu &= 0 \\ \text{FÜR } i=1,\dots,n & \end{aligned}}\right\} \Longleftrightarrow y_i=\frac{\mu}{x_i}$$

ELIMINATION VON $y_1,\dots,y_n$ ERGIBT KKT VON (4)

$$\boxed{\begin{aligned} \nabla f(x)-\mu\begin{bmatrix} x_1^{-1} \\ \vdots \\ x_n^{-1}\end{bmatrix}-\frac{\partial c}{\partial x}(x)^T\cdot\lambda &= 0 \\ c(x) &= 0 \end{aligned}}$$

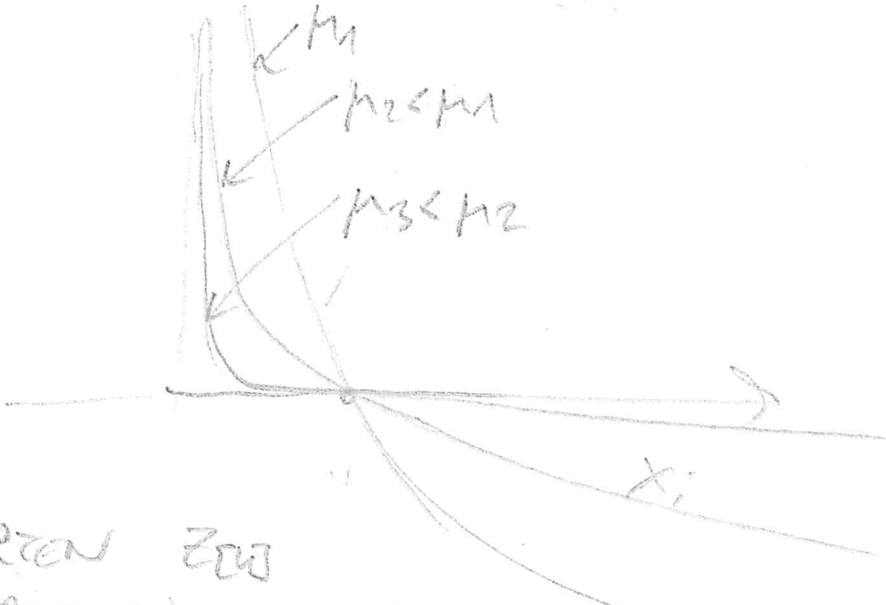
$\square$

INTERPRETATION:

UNGLEICHUNG $x_i \geq 0$ WIRD DURCH "BARRIERE"

$\mu \log x_i$

ERZWUNGEN.



ALLE ITERIERTEN ZEIJ
SIND "INNERE PUNKTE" (IP) DER ZULÄSSIGEN MENGE.

IP-METHODEN KÖNNEN (AUCH SEHR ZUVERLÄSSIG & EFFIZIENT) ZUR LÖSUNG VON

- LINEAREN PROGRAMMEN (LP) UND
- QUADRATISCHEN PROGRAMMEN (QP)

ANGEWENDET WERDEN. WG. DER KONVEXITÄT SIND DIE LÖSUNGEN DANN GLOBAL OPTIMAL.

7.2 Sequentielle Quadratische Programmierung (SQP)

Betrachte wieder (1)

$$\boxed{\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & c_E(x) = 0 \\ & c_I(x) \geq 0 \end{aligned}}$$

Falls wir einen QP-Löser verfügbar haben (z.B. IP-Methode), dann können wir iterativ QPs lösen, die durch Linearisierung an Stelle $x[k]$ gewonnen werden:

$$QP(x[k], Q_k) \qquad \boxed{\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \nabla f(x[k])^T (x - x[k]) + \tfrac{1}{2} (x - x[k])^T Q_k (x - x[k]) \\ \text{s.t.} \quad & c_E(x[k]) + \frac{\partial c_E}{\partial x}(x[k])(x - x[k]) = 0 \\ & c_I(x[k]) + \frac{\partial c_I}{\partial x}(x[k])(x - x[k]) \geq 0 \end{aligned}}$$

Wahl der "Hessematrix Approximation" sollte

a) positiv semi-definit sein (Konvexität des QP!)

b) ähnlich zu $\nabla_x^2 \mathcal{L}(x, \lambda_E, \lambda_I)$ sein (quadratische Konvergenz)

7.2.1 BESCHRÄNKTES GAUSS-NEWTON (CGN) VERFAHREN

FALLS $f(x) = \frac{1}{2}\|r(x)\|_2^2$ UND $J(x) = \frac{\partial r}{\partial x}(x)$,

DANN IST $Q_k := J(x_{[k]})^T J(x_{[k]})$

EINE GUTE WAHL. DAS QP IST DANN ÄQUIVALENT ZU

$$\boxed{\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \underbrace{\frac{1}{2}\|r(x_{[k]}) + J(x_{[k]})(x - x_{[k]})\|_2^2}_{=: f_{GN}(x; x_{[k]})} \\ \text{s.t.} \quad & c_E(x_{[k]}) + \frac{\partial c_E}{\partial x}(x_{[k]})(x - x_{[k]}) = 0 \\ & c_I(x_{[k]}) + \quad '' \quad \geq 0 \end{aligned}}$$

BEWEIS: $\nabla f(x_{[k]}) = J(x_{[k]})^T r(x_{[k]}) = \nabla f_{GN}(x_{[k]}; x_{[k]})$

UND $Q_k = J(x_{[k]})^T J(x_{[k]}) = \nabla^2 f_{GN}(x_{[k]}, x_{[k]})$.

VORTEIL VON GN-SQP: KONVEXE STRUKTUREN BLEIBEN IN QP ERHALTEN, ZUVERLÄSSIGE & ROBUSTE KONVERGENZ.