Definition: Sei $\min\{f(x) \mid x \in X\}$ OP gegeben. Ein $x \in X$ heißt

1) Lokales Minimum, falls $\exists \varepsilon > 0: f(x) \leq f(x') \ \forall x' \in B_\varepsilon(x) \cap X$

   wobei $B_\varepsilon(x) = \{y \in \mathbb{R}^n \mid \|x-y\| < \varepsilon\}$ und $\|z\| = \sqrt{z^T z} = \left(\sum_{n=1}^{N} |z_n|^2\right)^{\frac{1}{2}}$

2) striktes (oder isoliertes) Minimum, falls $\exists \varepsilon > 0$ mit $f(x) < f(x')$

   $\forall x' \in (B_\varepsilon(x) \cap X) \setminus \{x\}$

3) globales Minimum, falls $f(x) \leq f(x') \quad \forall x' \in X$

4) striktes globales Minimum, falls $f(x) < f(x') \ \forall x' \in X \setminus \{x\}$

Satz 1.4 Sei OP gegeben, wobei f stetig ist. Ferner sei $x_0 \in X$ mit

$N_f(x_0) = \{x \in X : f(x) \leq f(x_0)\}$ „Niveaumenge zum Niveau x_0"

kompakt (abgeschlossen und beschränkt). Dann hat das OP eine Lösung.

Beweis: Aus Analysis, mit Satz von Weierstrass: g stetig, Y kompakt

$\Rightarrow \exists y \in Y$ mit $g(y) \leq g(y') \ \forall y' \in Y$. Nun Satz für f und $N_f(x_0)$

verwenden. □

Frage: Wie findet man / charakterisiert man solche Minima?

$\rightsquigarrow$ Darum geht es in dieser Vorlesung!

Arten von OP

Uneingeschränktes OP: $\min\{f(x) : x \in \mathbb{R}^n\}$

Eingeschränktes OP: üblicherweise wird die zulässige Menge über Nebenbedingungen (Gleichungen und Ungleichungen) definiert.

Sei $h: \mathbb{R}^n \to \mathbb{R}^p$, $g: \mathbb{R}^n \to \mathbb{R}^m$, dann $X = \{x \in \mathbb{R}^n : h(x) = 0, \ g(x) \leq 0\}$

(Un-)Gleichungen sind komponentenweise zu verstehen.

Für das OP schreiben wir dann
$$\min f(x) \quad \text{s.t. } h(x) = 0, \ g(x) \leq 0$$
„Nicht-lineares OP"

18.10.18

Lineares OP: f, h, g linear.
$$\begin{aligned} &\min c^T x && c \in \mathbb{R}^n \\ &\text{s.t. } Bx - d = 0 && B \in \mathbb{R}^{p \times n}, \ d \in \mathbb{R}^p \\ &\phantom{\text{s.t. }} Ax - b \leq 0 && A \in \mathbb{R}^{m \times n}, \ b \in \mathbb{R}^m \end{aligned}$$

Quadratische OP: Zielfunktion $f(x) = x^T Q x + c^T x$, $Q \in \mathbb{R}^{n \times n}$, $c \in \mathbb{R}^n$

und h, g linear

Konvexe OP: Hier sind f, g konvex und h linear

Wichtige Eigenschaft (später): x lokales Min $\Rightarrow$ x globales Min.

Bsp. 1.5: Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch, betrachte das OP

$$\min_{\text{st. } x^T x = 1} x^T A x = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \qquad \left(x^T x = \sum_{i=1}^n x_i^2 = 1\right)$$

Beh: Der Wert dieses OP ist λ_1, wobei $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$

die Eigenwerte von A sind.

Beweis: Da A symmetrisch, gibt es eine Orthonormalbasis des $\mathbb{R}^n$ aus

Eigenvektoren von A, d.h. es gibt Vektoren $v_1, \ldots, v_n \in \mathbb{R}^n$ mit

1) $v_i^T v_j = 1$ wenn $i = j$ und $= 0$ sonst für $1 \leq i, j \leq n$

2) $A v_i = \lambda_i v_i$ $1 \leq i \leq n$

3) Basis des $\mathbb{R}^n$

Sei $x \in \mathbb{R}^n$ mit $x^T x = 1$ und schreibe eindeutig $x = \sum_{i=1}^n \alpha_i v_i$, $\alpha_i \in \mathbb{R}$

Dann $1 = x^T x = \left(\sum_{i=1}^n \alpha_i v_i\right)^T \left(\sum_{i=1}^n \alpha_i v_i\right) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i v_i^T v_j \alpha_j = \sum_{i=1}^n \alpha_i^2$

Damit erhalten wir $x^T A x = x^T A \left(\sum_{i=1}^n \alpha_i v_i\right) = x^T \sum_{i=1}^n \alpha_i \lambda_i v_i =$

$= \sum_{j=1}^n \sum_{i=1}^n \alpha_j \alpha_i \lambda_i v_j^T v_i = \sum_{i=1}^n \alpha_i^2 \lambda_i \geq \lambda_1 \sum_{i=1}^n \alpha_i^2 = \lambda_1$

Ferner gilt für $x = v_1$, dass $v_1^T A v_1 = \lambda_1$ $\square$

1.1 Uneingeschränkte OP

$\min\{f(x) : x \in \mathbb{R}^n\}$, $f : \mathbb{R}^n \to \mathbb{R}$

Satz 1.6: Notwendige Bedingung 1. Ordnung

Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ differenzierbar. Dann

$x' \in U$ lokales Minimum $\Rightarrow$ $\nabla f(x') = 0$

Beweis: Sei $d \in \mathbb{R}^n$. Da x' lokales Minimum: $f(x'+td)-f(x') \geq 0$ für $t>0$ hinreichend klein. Damit $(f(x'+td)-f(x')):t \geq 0$ für $t>0$ hinreichend klein. Für $t \to 0$ erhalten wir: $\nabla f(x')^T d \geq 0$. Mit $d=-\nabla f(x') \Rightarrow \nabla f(x')^T \nabla f(x') \leq 0 \Rightarrow \nabla f(x')=0$

Definition: 1) Sei $U \subseteq \mathbb{R}^n$ offen, $x' \in U$, $f: U \to \mathbb{R}$ diff.bar in einer Umgebung von x'. Falls $\nabla f(x')=0$, dann heißt x' stationärer Punkt von f.

2) Ein stationärer Punkt, der weder (lokales) Minimum noch Maximum ist, heißt Sattelpunkt von f.

Satz 1.8: Notwendige Bedingung 2. Ordnung

Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \to \mathbb{R}$ $2\times$ stetig diff.bar. Dann:

$$x' \in U \text{ lokales Min} \Rightarrow \begin{cases} \text{a) } \nabla f(x')=0 \\ \text{b) } \nabla^2 f(x) \text{ ist positiv semidefinit} \end{cases}$$

Bemerkung: $A \in \mathbb{R}^{n \times n}$ pos. semidef. $\Leftrightarrow$ alle EW von A sind ≥ 0

$\Leftrightarrow d^T A d \geq 0 \ \forall d \in \mathbb{R}^n$

Beweis: a) aus Satz 1.6

b) Sei $d \in \mathbb{R}^n \setminus \{0\}$. Dann gibt es $\tau > 0$ s.d., $\forall t \in [0, \tau[$:

$0 \leq f(x'+td)-f(x') = t \cdot \nabla f(x')^T d + \frac{t^2}{2} d^T \nabla^2 f(x') d + \rho(t)$

mit $\rho(t) = o(t^2)$ für $t \to 0$ (Taylor)

Mit a): $0 \leq \frac{t^2}{2} d^T \nabla^2 f(x) d + \rho(t) \Rightarrow d^T \nabla^2 f(x') d \geq -\frac{2\rho(t)}{t^2} \xrightarrow{t \to 0} 0$ $\square$

Satz 1.9 (Hinreichende Bedingung 2. Ordnung)

Sei $U \subseteq \mathbb{R}^n$, $f: U \to \mathbb{R}$ $2\times$ stetig diff.bar, U offen und sei $x' \in U$ mit:

$$\left.\begin{array}{l} \text{a) } \nabla f(x')=0 \\ \text{b) } \nabla^2 f(x') \text{ ist positiv definit} \end{array}\right\} \Rightarrow x' \text{ striktes lokales Minimum}$$

Beweis: Da f 2x stetig diff.bar, gibt es $\varepsilon > 0$ so dass sogar $\nabla^2 f(x)$ $\forall\, x \in B_\varepsilon(x')$ positiv definit ist. Mit Taylor für ein $x \in B_\varepsilon(x') \setminus \{x'\}$: $f(x) - f(x') = \nabla f(x')^T (x - x') + \frac{1}{2} \underbrace{(x-x')^T \nabla^2 f(x')(x-x')}_{>0}$ für ein $\tilde{x} = x + \lambda(x' - x)$, $\lambda \in\,]0,1[$ $\square$

1.2 Das Newton-Verfahren

Iteratives Verfahren zum Lösen von $F(x) = 0$, $F: \mathbb{R}^n \to \mathbb{R}^n$ (z.B. $F = \nabla f$)

Annahme: F stetig diff.bar

Startpunkt: x_0, Folge: $x_1, x_2, x_3, \ldots$

Idee: $0 = F(x) \approx F(x_k) + F'(x_k)(x - x_k)$

$$\Rightarrow x \approx x_k - F'(x_k)^{-1} F(x_k)$$

Newton-Verfahren: $\boxed{x_{k+1} = x_k - F'(x_k)^{-1} F(x_k),\ k \in \mathbb{N}_0}$

Bsp.: $F: \mathbb{R} \to \mathbb{R}$, dann: $\boxed{x_{k+1} = x_k - \frac{F(x_k)}{F'(x_k)},\ \text{stop falls } F'(x_k) = 0}$

Sei $F(x) = 1 - \frac{a}{x^2}$, für ein $a > 0$. Dann: $F(x) = 0 \Leftrightarrow x \in \{\sqrt{a}, -\sqrt{a}\}$

Mit $F'(x) = \frac{2a}{x^3}$ erhalten wir: $x_{k+1} = x_k - \frac{F(x_k)}{F'(x_k)} = \frac{x_k}{2}\left(3 - \frac{x_k^2}{a}\right)$

Betrachten wir mal $x_0 = 1{,}5$, $a = 2$

$\Rightarrow x_1 = 1{,}4$

$x_2 = 1{,}414$ $\quad$ $\hat{=}$ richtige Nachkommastellen

$x_3 = 1{,}414213514$

\# richtiger Stellen vervielfacht sich!

$\vdots$

Tatsächlich gilt: $(x_{k+1} - \sqrt{a}) = \frac{x_k}{2} 3 - \frac{x_k^3}{2a} - \sqrt{a} = (x_k - \sqrt{a})^2 \cdot (x_k + 2\sqrt{a})$

Der Fehler quadriert sich in jedem Schritt!