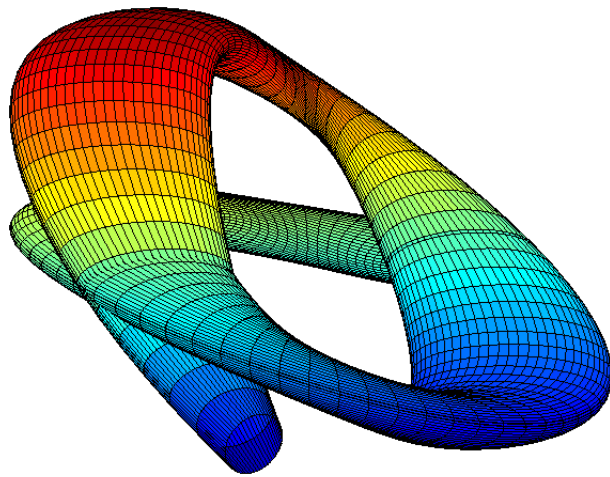


Skriptum zur Vorlesung

# Optimierung

Private Mitschrift



gelesen von

Prof. Dr. Stefan Volkwein

Martin Gubisch

Konstanz, Sommersemester 2009

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>3</b>
<b>2</b>	<b>Optimalitätskriterien</b>	<b>4</b>
2.1	Bedingungen erster und zweiter Ordnung . . . . .	4
2.2	Konvexe Funktionen . . . . .	6
<b>3</b>	<b>Abstiegsverfahren und Schrittweitenstrategien</b>	<b>8</b>
3.1	Abstiegsverfahren . . . . .	9
3.2	Schrittweitenstrategien . . . . .	10
3.2.1	Die Armijo-Regel . . . . .	10
3.2.2	Die Wolfe-Powell-Regel . . . . .	12
3.2.3	Die strenge Wolfe-Powell-Regel . . . . .	14
3.3	Praktische Aspekte . . . . .	14
3.4	Gradientenverfahren . . . . .	14
<b>4</b>	<b>Newton-Verfahren</b>	<b>16</b>
4.1	Das lokale Newton-Verfahren . . . . .	16
4.2	Das inexakte Newton-Verfahren . . . . .	18
4.2.1	Das Newton-CG-Verfahren . . . . .	19
4.3	Das globalisierte Newton-Verfahren . . . . .	20
4.3.1	Die Trust-Region-Methode . . . . .	20
4.3.2	Globale Konvergenz des Trust-Region-Verfahrens . . . . .	21
4.4	Quasi-Newton-Verfahren . . . . .	22
	<b>Übungsaufgaben</b>	<b>24</b>
	<b>Index</b>	<b>31</b>
	<b>Literaturverzeichnis</b>	<b>32</b>

## 1. Einleitung

Unter einem **endlichdimensionalen Optimierungsproblem** verstehen wir folgende Aufgabe:

$$\begin{cases} \text{Gegeben sei eine Menge } X \subseteq \mathbb{R}^n \text{ und eine stetige Funktion } f : X \rightarrow \mathbb{R}. \\ \text{Gesucht wird ein } x^* \in X \text{ mit } \forall x \in X : f(x^*) \leq f(x) \end{cases}. \quad (1.1)$$

In Kurznotation:

$$\min_{x \in X} f(x) \quad \text{u.d.N. } x \in X \quad \text{bzw.} \quad \min_{x \in X} f(x). \quad (1.2)$$

Ist  $X = \mathbb{R}^n$ , so heißt (1.1) bzw. (1.2) **unrestringiert**, andernfalls **restringiert**.

Im Allgemeinen nennt man  $X$  den **Zulässigkeitsbereich** und  $f$  die **Zielfunktion**.

### Bemerkung 1.1.

Soll  $f$  maximiert werden für  $x \in X$ , so ist dies gleichbedeutend damit, dass  $-f$  minimiert wird u.d.N.  $x \in X$ .  $\diamond$

Die Aufgabenstellung (1.1) erhält ihre Bedeutung dadurch, dass sie ein mathematisches Modell für viele Probleme zum Beispiel aus der Physik, Medizin, Ökonomie und den Ingenieurwissenschaften ist.

Für den Fall, dass  $X \neq \mathbb{R}^n$  ist, lässt sich der Zulässigkeitsbereich sehr häufig in der Form  $X = \Omega_1 \cap \Omega_2 \cap \Omega_3$  schreiben mit

$$\Omega_1 = \{x \in \mathbb{R}^n \mid c_i(x) = 0, i \in I_1\}, \quad \Omega_2 = \{x \in \mathbb{R}^n \mid c_i(x) \leq 0, i \in I_2\}, \quad \Omega_3 = \{x \in \mathbb{R}^n \mid x_i \in \mathbb{Z}, i \in I_3\},$$

gewisse Indexmengen  $I_1, I_2, I_3$  und Abbildungen  $c_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i \in I_1, I_2$ . Die Mengen  $\Omega_1, \Omega_2, \Omega_3$  werden als **Gleichungs-**, **Ungleichungs-** bzw. **Ganzzahligkeitsrestriktionen** bezeichnet.

Ist  $X$  eine Menge von diskreten Punkten, so spricht man von einem **diskreten** oder **kombinatorischen** Optimierungsproblem, andernfalls von einem **stetigen** Optimierungsproblem.

Ist  $f$  nicht differenzierbar, so nennt man (1.1) ein **nicht-differenzierbaren** Optimierungsproblem.

### Definition 1.2.

Sei  $f : X \rightarrow \mathbb{R}$  mit  $X \subseteq \mathbb{R}^n$ . Ein Punkt  $x^* \in X$  heißt ...

1. **globale Minimalstelle** von  $f$  (auf  $X$ ), wenn  $f(x^*) \leq f(x)$  für alle  $x \in X$ .  $f(x^*)$  heißt dann **globales Minimum**.
2. **strikte globale Minimalstelle** von  $f$  (auf  $X$ ), wenn  $f(x^*) < f(x)$  für alle  $x \in X$ .  $f(x^*)$  heißt dann **striktes globales Minimum**.
3. **lokale Minimalstelle** von  $f$  (auf  $X$ ), wenn es eine Umgebung  $U$  von  $x^*$  gibt, so dass  $f(x^*) \leq f(x)$  für alle  $x \in X \cap U$ ;  $f(x^*)$  heißt dann **lokales Minimum**.
4. **strikte lokale Minimalstelle** von  $f$  (auf  $X$ ), wenn es eine Umgebung  $U$  von  $x^*$  gibt mit  $f(x^*) < f(x)$  für alle  $x \in U \cap X$ .  $f(x^*)$  heißt dann **striktes lokales Minimum**.

### Bemerkung 1.3.

Ein Punkt  $x^* \in X$  ist genau dann (**globale, strikte globale, lokale, strikte lokale**) **Maximalstelle** von  $f$  auf  $X$ , wenn  $x^*$  (**globale, strikte globale, lokale, strikte lokale**) **Minimalstelle** von  $-f$  auf  $X$  ist.  $\diamond$

Im Folgenden bezeichne  $\nabla f(x) = (\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x))^T \in \mathbb{R}^n$  den Gradienten von  $f$  in  $x$ .

### Definition 1.4.

Seien  $X \subseteq \mathbb{R}^n$  eine offene Menge und  $f : X \rightarrow \mathbb{R}$  eine stetig differenzierbare Funktion.

Ein Punkt  $x^* \in X$  heißt **stationärer Punkt** von  $f$ , wenn  $\nabla f(x^*) = 0$  gilt.

## 2. Optimalitätskriterien

### 2.1. Bedingungen erster und zweiter Ordnung

Wir behandeln unter geeigneten Differenzierbarkeitsannahmen notwendige und hinreichende Bedingungen für lokale Minimalstellen.

#### Satz 2.1. (Notwendige Optimalitätsbedingung erster Ordnung)

Seien  $X \subseteq \mathbb{R}^n$  offen und  $f : X \rightarrow \mathbb{R}$  stetig differenzierbar. Ist  $x^* \in X$  eine lokale Minimalstelle von  $f$  (auf  $X$ ), so gilt  $\nabla f(x^*) = 0$ , d.h.  $x^*$  ist ein stationärer Punkt.

**Beweis.** (Analysis II)

Sei  $x^* \in X$  eine lokale Minimalstelle von  $f$ , aber  $\nabla f(x^*) \neq 0$ . Dann existiert  $d \in \mathbb{R}^n$  mit  $\nabla f(x^*)^T d < 0$  (z.B.  $d = -\nabla f(x^*)$ ). Da nach Voraussetzung  $f$  stetig differenzierbar ist, existiert die Richtungsableitung  $f'(x^*; d)$  von  $f$  in  $x^*$  in Richtung  $d$ . Es gilt

$$f'(x^*; d) = \lim_{t \searrow 0} \frac{f(x^* + td) - f(x^*)}{t} = \nabla f(x^*)^T d < 0.$$

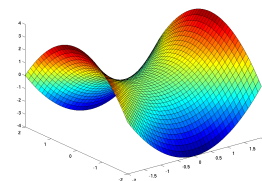
Folglich existiert ein  $t_0 > 0$  mit  $x^* + td \in X$  und  $\frac{1}{t}(f(x^* + td) - f(x^*)) < 0$  für alle  $t \in (0, t_0]$ . Somit ist auch  $f(x^* + td) < f(x^*)$  für alle  $t \in (0, t_0]$ , was einen Widerspruch zur Voraussetzung ergibt.  $\square$

#### Bemerkung 2.2.

Die Bedingungen aus Satz 2.4 sind nicht hinreichend dafür, dass  $x^*$  eine lokale Minimalstelle ist.

Betrachte zum Beispiel die Funktion  $f(x) = x_1^2 - x_2^2$  auf  $\mathbb{R}^2$  und den Punkt  $x^* = (0, 0)$ , dann gilt  $\nabla f(x_1, x_2) = (2x_1, -2x_2)^T$ , also ist  $x^* = (0, 0)$  ein stationärer Punkt von  $f$ , aber für beliebig kleines  $\epsilon > 0$  ist  $f(x_\epsilon) = -\epsilon^2 < 0 = f(x^*)$  mit  $x_\epsilon = (0, \epsilon)$ , d.h. in jeder Umgebung von  $x^*$  existiert ein Punkt mit kleinerem Funktionswert.

Da Satz 2.1 nur Ableitungen bis zur ersten Ordnung verwendet, gibt er eine notwendige Bedingung erster Ordnung an.  $\diamond$



**Fig. 1:** Eine Funktion mit einem **Sattelpunkt**, d.h. einem kritischen Punkt, der kein Optimum ist.

Wir zitieren folgende Abschätzung aus der Spektraltheorie symmetrischer Matrizen:

#### Lemma 2.3.

Sei  $\mathcal{S}_n$  der Vektorraum der symmetrischen  $(n \times n)$ -Matrizen. Für  $A \in \mathcal{S}_n$  sei  $\lambda(A)$  der kleinste Eigenwert von  $A$ . Dann gilt

$$|\lambda(A) - \lambda(B)| \leq \|A - B\|_2 \quad \text{für alle } A, B \in \mathcal{S}_n,$$

wobei  $\|\cdot\|_2$  hier die **Spektralnorm**  $\|A\|_2 = \max\{|\lambda| \mid \lambda \text{ ist Eigenwert von } A\}$  bezeichnet.

Mit Hilfe von Lemma 2.3 folgt aus der Stetigkeit der Hessematrix  $\nabla^2 f(x) = \left(\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f(x)\right)_{i,j=1,\dots,n} \in \mathbb{R}^{n \times n}$  von  $f$ : Ist  $\nabla^2 f(x^*)$  positiv definit, dann ist dass  $\nabla^2 f(x)$  positiv definit in einer Umgebung von  $x^*$ . Eine analoge Folgerung gilt für den Fall, dass  $\nabla^2 f(x^*)$  negativ definit ist.

#### Satz 2.4. (Notwendige Optimalitätsbedingung zweiter Ordnung)

Seien  $X \subseteq \mathbb{R}^n$  offen und  $f : X \rightarrow \mathbb{R}$  zweimal stetig differenzierbar. Ist  $x^* \in X$  eine lokale Minimalstelle von  $f$  auf  $X$ , so ist  $\nabla f(x^*) = 0$  und  $\nabla^2 f(x^*)$  ist positiv semidefinit.

**Beweis.** (Analysis II)

Die Bedingung  $\nabla f(x^*) = 0$  folgt aus Satz 2.1. Sei  $x^*$  eine lokale Minimalstelle von  $f$ , jedoch  $\nabla^2 f(x^*)$  nicht positiv semidefinit. Dann existiert ein  $d \in \mathbb{R}^n$  mit  $d^T \nabla^2 f(x^*) d < 0$ .

Mit Hilfe des Satzes von Taylor ergibt sich

$$f(x^* + td) = f(x^*) + \underbrace{t \nabla f(x^*)^T}_{=0} d + \frac{1}{2} t^2 d^T \nabla^2 f(\xi_t) d$$

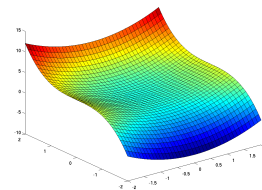
für kleines  $t > 0$ . Dabei ist  $\xi_t = x^* + \vartheta_t td$  für  $\vartheta_t \in (0, 1)$ . Aus Lemma 2.3 und der Stetigkeit der zweiten Ableitung von  $f$  folgt die Existenz von  $t_0 > 0$  mit  $d^T \nabla^2 f(\xi_t) d < 0$  für alle  $t \in (0, t_0]$ , wegen  $\nabla f(x^*) = 0$  also  $f(x^* + td) < f(x^*)$  für alle  $t \in (0, t_0]$ , was einen Widerspruch zur Voraussetzung ergibt.  $\square$

**Bemerkung 2.5.**

Die Bedingungen aus Satz 2.4 sind nicht hinreichend dafür, dass  $x^*$  eine lokale Minimalstelle ist.

Betrachte zum Beispiel die Funktion  $f(x) = x_1^2 + x_2^3$  auf  $\mathbb{R}^2$  und den Punkt  $x^* = (0, 0)$ , dann gelten

$$\nabla f(x^*) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \nabla^2 f(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$



**Fig. 2:** Ein Sattelpunkt mit semidefiniter Hessematrix.

Da Satz 2.4 Ableitungen bis zur zweiten Ordnung verwendet, gibt er eine notwendige Bedingung zweiter Ordnung an.  $\diamond$

Nun kommen wir zu einer hinreichenden Optimalitätsbedingung.

**Satz 2.6. (Hinreichende Optimalitätsbedingung zweiter Ordnung)**

Seien  $X \subseteq \mathbb{R}^n$  offen und  $f : X \rightarrow \mathbb{R}$  zweimal stetig differenzierbar. Gilt  $\nabla f(x^*) = 0$  und ist  $\nabla^2 f(x^*)$  ist positiv definit, dann ist  $x^*$  eine strikte lokale Minimalstelle von  $f$  auf  $X$ .

**Beweis.** (Analysis II)

Aus der positiven Definitheit der Hessematrix folgt die Existenz eines  $\mu > 0$  mit  $d^T \nabla^2 f(x^*) d \geq \mu d^T d$  für alle  $d \in \mathbb{R}^n$ , wähle etwa  $\mu$  als den kleinsten Eigenwert von  $\nabla^2 f(x^*)$ . Da  $x^*$  ein stationärer Punkt ist, gilt nach dem Satz von Taylor für alle  $d \in \mathbb{R}^n$ , die hinreichend nahe bei 0 sind, dass

$$f(x^* + d) = f(x^*) + \underbrace{\nabla f(x^*)^T}_{=0} d + \frac{1}{2} d^T \nabla^2 f(\xi_d) d$$

mit  $\xi_d = x^* + \vartheta_d d$  für  $\vartheta_d \in (0, 1)$  erfüllt ist. Man erhält so mit der Cauchy-Schwarzschen Ungleichung

$$\begin{aligned} f(x^* + d) &= f(x^*) + \frac{1}{2} d^T \nabla^2 f(x^*) d + \frac{1}{2} d^T (\nabla^2 f(\xi_d) - \nabla^2 f(x^*)) d \\ &\geq f(x^*) + \frac{1}{2} (\mu - \|\nabla^2 f(\xi_d) - \nabla^2 f(x^*)\|_2) \|d\|_2^2 > f(x^*), \end{aligned}$$

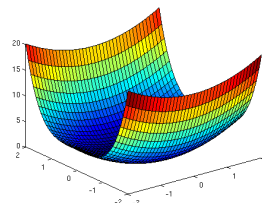
falls  $\|\nabla^2 f(\xi_d) - \nabla^2 f(x^*)\|_2$  kleiner  $\mu$  und  $\|d\|_2$  klein ( $d \neq 0$ ). Damit ist  $x^*$  eine strikte lokale Minimalstelle von  $f$ .  $\square$

**Bemerkung 2.7.**

Die Bedingungen aus Satz 2.6 sind nicht notwendig dafür, dass  $x^*$  eine lokale Minimalstelle ist.

Betrachte zum Beispiel die Funktion  $f(x) = x_1^2 + x_2^4$  auf  $\mathbb{R}^2$  und den Punkt  $x^* = (0, 0)$ , dann gelten

$$\nabla f(x^*) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \nabla^2 f(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$



**Fig. 3:** Ein striktes lokales Minimum mit semidefiniter Hessematrix.

Da Satz 2.4 Ableitungen bis zur zweiten Ordnung verwendet, gibt er eine hinreichende Bedingung zweiter Ordnung an.  $\diamond$

## 2.2. Konvexe Funktionen

### Definition 2.8.

Eine Menge  $X \subseteq \mathbb{R}^n$  heißt **konvex**, wenn für alle  $x, y \in X$  und  $\lambda \in (0, 1)$  auch  $\lambda x + (1 - \lambda)y$  in  $X$  liegt.

Seien  $X \subseteq \mathbb{R}^n$  eine konvexe Menge und  $f : X \rightarrow \mathbb{R}$  eine Funktion.

1.  $f$  heißt **strikt konvex** bzw. **konvex**, wenn für alle  $x, y \in X$ ,  $x \neq y$  und alle  $\lambda \in (0, 1)$  gilt

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y) \quad \text{bzw.} \quad f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

2.  $f$  heißt **gleichmäßig konvex** mit **Modul**  $\mu > 0$ , falls für alle  $x, y \in X$ ,  $\lambda \in (0, 1)$  gilt:

$$f(\lambda x + (1 - \lambda)y) + \mu\lambda(1 - \lambda)\|x - y\|^2 \leq \lambda f(x) + (1 - \lambda)f(y).$$

Aus der Definition folgt, dass jede gleichmäßig konvexe Funktion auch strikt konvex ist und jede strikt konvexe Funktion konvex ist. Die Umkehrungen gelten im Allgemeinen nicht.

### Bemerkung 2.9.

$f : \mathbb{R}^n \rightarrow \mathbb{R}$  sei eine quadratische Funktion, d.h. von der Gestalt

$$f(x) = \frac{1}{2}x^T Qx + c^T x + \gamma$$

mit symmetrischer Matrix  $Q \in \mathcal{S}_n$ , Vektor  $c \in \mathbb{R}^n$  und Skalar  $\gamma \in \mathbb{R}$ . Dann gelten:

1.  $f$  ist genau dann konvex, wenn  $Q$  positiv semidefinit ist.
2.  $f$  ist strikt konvex  $\Leftrightarrow f$  ist gleichmäßig konvex  $\Leftrightarrow Q$  ist positiv definit.  $\diamond$

### Satz 2.10. (Charakterisierung konvexer Funktionen)

Seien  $X \subseteq \mathbb{R}^n$  offen und konvex sowie  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  zweimal stetig differenzierbar. Dann gelten:

1.  $f$  ist konvex auf  $X \Leftrightarrow \nabla^2 f(x)$  ist für alle  $x \in X$  positiv semidefinit.
2. Ist  $\nabla^2 f(x)$  für alle  $x \in X$  positiv definit, so ist  $f$  strikt konvex (auf  $X$ ).
3.  $f$  ist genau dann gleichmäßig konvex (auf  $X$ ), wenn  $\nabla^2 f(x)$  gleichmäßig positiv definit ist, d.h. wenn es ein  $\mu > 0$  gibt, so dass für alle  $x \in X$  und alle  $d \in \mathbb{R}^n$  gilt:  $d^T \nabla^2 f(x)d \geq \mu \|d\|^2$ .

### Bemerkung 2.11.

Aussage (2) von Satz 2.10 lässt sich nicht ohne Weiteres umkehren: Die Funktion  $f(x) = x^4$  mit  $X = \mathbb{R}$  beispielsweise ist strikt konvex, aber es gilt  $f''(0) = 0$ .  $\diamond$

Der folgende Hilfssatz beschäftigt sich mit Niveaumengen gleichmäßig konvexer Funktionen:

### Lemma 2.12.

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar,  $x^0 \in \mathbb{R}^n$  beliebig gegeben, die Niveaumenge

$$\mathcal{L}(x^0) = \{x \in \mathbb{R}^n \mid f(x) \leq f(x^0)\}$$

konvex und  $f$  gleichmäßig konvex auf  $\mathcal{L}(x^0)$ . Dann ist  $\mathcal{L}(x^0)$  kompakt.

Da stetige Funktionen auf kompakten Mengen stets ihr Maximum und Minimum annehmen, lässt sich damit folgende Aussage über die Existenz und Eindeutigkeit von Lösungen der Optimierungsaufgabe 1.2

$$\min f(x) \quad \text{u.d.N.} \quad x \in X.$$

herleiten:

**Satz 2.13. (Existenz und Eindeutigkeit für konvexe Optimierungsaufgaben)**

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar und  $X \subseteq \mathbb{R}^n$  konvex. Dann gelten:

1. Ist  $f$  konvex auf  $X$ , so ist die Lösungsmenge von (1.2) konvex (gegebenenfalls leer).
2. Ist  $f$  strikt konvex auf  $X$ , so besitzt (1.2) höchstens eine Lösung.
3. Sind  $f$  gleichmäßig konvex auf  $X$ ,  $X \neq \emptyset$  und abgeschlossen, so besitzt (1.2) genau eine Lösung.

**Beweis.**

1. Seien  $x^1, x^2$  Lösungen von (1.2), also  $f(x^1) = f(x^2) = \min_{x \in X} f(x)$ . Für  $\lambda \in (0, 1)$  gelten dann

$$\lambda x^1 + (1 - \lambda)x^2 \in X \quad \& \quad f(\lambda x^1 + (1 - \lambda)x^2) \leq \lambda f(x^1) + (1 - \lambda)f(x^2) = \min_{x \in X} f(x).$$

Also nimmt  $f$  auch an  $\lambda x^1 + (1 - \lambda)x^2$  sein Minimum an.

2. Angenommen, (1.2) hat zwei verschiedene Lösungen  $x^1, x^2$ . Für  $\lambda \in (0, 1)$  gelten dann

$$\lambda x^1 + (1 - \lambda)x^2 \in X \quad \& \quad f(\lambda x^1 + (1 - \lambda)x^2) < \lambda f(x^1) + (1 - \lambda)f(x^2) = \min_{x \in X} f(x),$$

was einen Widerspruch ergibt.

3. Sei  $x^0 \in X$  beliebig gewählt. Dann ist wegen

$$\forall x, y \in \mathcal{L}(x^0) : f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) = f(x^0)$$

die Menge  $\mathcal{L}(x^0)$  konvex und nach Lemma 2.12 kompakt. Dann ist  $X \cap \mathcal{L}(x^0)$  kompakt und nichtleer. Daher nimmt die stetige Funktion  $f$  ihr globales Minimum in  $X \cap \mathcal{L}(x^0)$  an, welches natürlich auch ein Minimum von (1.2) ist.  $\square$

**Bemerkung 2.14.**

1. Das Problem (1.2) muss selbst für strikt konvexes  $f$  keine Lösung besitzen. Betrachte dazu zum Beispiel  $f(x) = \exp(x)$  auf  $X = \mathbb{R}$ .
2. Auf die Forderung nach Abgeschlossenheit von  $X$  in Satz 2.13 (3) kann nicht verzichtet werden, betrachte etwa  $f(x) = x^2$  für  $x \in (0, 1]$ .  $\diamond$

Das zentrale Resultat dieses Abschnitts wird in Satz 2.15 angegeben. Man kann daraus ersehen, dass  $\nabla f(x^*) = 0$  auch hinreichend ist dafür, dass  $x^*$  eine globale Minimalstelle ist.

**Korollar 2.15. (Charakterisierung von Minimalstellen konvexer Funktionen)**

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar und konvex und  $x^* \in \mathbb{R}^n$  ein stationärer Punkt von  $f$ . Dann ist  $x^*$  eine globale Minimalstelle von  $f$  auf  $\mathbb{R}^n$ .

**Beweis.**

Sei  $x \in \mathbb{R}^n$ . Nach dem Satz von Taylor existiert ein  $\vartheta \in (0, 1)$ , so dass für  $\xi = x^* + \vartheta(x - x^*)$  gilt

$$f(x) = f(x^*) + \nabla f(x^*)(x - x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 f(\xi)(x - x^*).$$

Nach Satz 2.13 (1) ist  $\nabla^2 f(\xi)$  positiv semidefinit. Ferner gilt  $\nabla f(x^*) = 0$ , also  $f(x) \geq f(x^*)$ .  $\square$

Wir schließen das Kapitel mit einer Aussage darüber, wie nahe ein Punkt mindestens bei einer Minimalstelle liegen muss, wenn sein Funktionswert nahe dem Minimalwert ist.

**Korollar 2.16.**

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar,  $x^0 \in \mathbb{R}^n$ , die Niveaumenge  $\mathcal{L}(x^0)$  konvex,  $f$  gleichmäßig konvex auf  $\mathcal{L}(x^0)$  und  $x^*$  die eindeutige globale Minimalstelle von  $f$ . Dann existiert ein  $C > 0$  mit

$$\forall x \in \mathcal{L}(x^0) : \quad C \|x - x^*\|^2 \leq |f(x) - f(x^*)|.$$

### 3. Abstiegsverfahren und Schrittweitenstrategien

Wir betrachten ein Abstiegsverfahren zur Lösung des unrestringierten Optimierungsproblems

$$\min_{x \in \mathbb{R}^n} f(x)$$

mit stetig differenzierbarer Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Die zentrale Idee der Verfahren in diesem Abschnitt ist wie folgt:

1. Ist man in einem Punkt  $x \in \mathbb{R}^n$ , so sucht man eine Richtung  $d \in \mathbb{R}^n$  aus, längs welcher der Funktionswert fällt (**Abstiegsverfahren**).
2. Entlang dieser Richtung  $d$  geht man so lange, bis man den Funktionswert von  $f$  hinreichend verkleinert hat (**Schrittweitenstrategie**).

Diese Schritte wollen wir formalisieren.

#### Definition 3.1.

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  und  $x \in \mathbb{R}^n$ . Ein Vektor  $d \in \mathbb{R}^n$  heißt **Abstiegsrichtung** von  $f$  im Punkt  $x$ , wenn es ein  $t_0 > 0$  gibt mit

$$f(x + td) < f(x) \quad \text{für alle } t \in (0, t_0].$$

#### Bemerkung 3.2. (Hinreichende Abstiegsbedingung)

Ist  $f$  stetig differenzierbar, dann ist die Bedingung

$$\nabla f(x)^T d < 0 \tag{3.1}$$

hinreichend dafür, dass  $d \in \mathbb{R}^n$  eine Abstiegsrichtung von  $f$  in  $x$  ist.

Um dies einzusehen, definieren wir  $\varphi(t) = f(x + td)$ . Aus  $f \in \mathcal{C}^1$  folgt dann

$$\varphi(t) = \varphi(0) + t\varphi'(0) + r(t) \quad \text{mit} \quad \frac{1}{t}r(t) \rightarrow 0 \quad \text{für } t \searrow 0,$$

d.h.  $r(t) = \mathcal{O}(t)$  in der Landau-Symbol-Formalisierung. Es gelten  $\varphi(0) = f(x)$  und  $\varphi'(0) = \nabla f(x)^T d < 0$ . Also existiert ein  $t_0 > 0$  mit

$$\frac{\varphi(t) - \varphi(0)}{t} = \nabla f(x)^T d + \frac{r(t)}{t} \quad \implies \quad \forall t \in (0, t_0] : \quad \frac{\varphi(t) - \varphi(0)}{t} < 0,$$

d.h.  $f(x + td) < f(x)$  und somit ist  $d$  eine Abstiegsrichtung von  $f$  in  $x$ .  $\diamond$

#### Bemerkung 3.3.

1. Die Bedingung (3.1) bedeutet, dass der Winkel zwischen  $d$  und dem negativen Gradienten  $-\nabla f(x)$  von  $f$  in  $x$  weniger als  $90^\circ$  beträgt.
2. Das Kriterium (3.1) ist nicht notwendig. Ist  $x$  beispielsweise eine strikte lokale Maximalstelle, so sind alle  $d \in \mathbb{R}^n$  Abstiegsrichtungen von  $f$  in  $x$ , aber (3.1) ist natürlich nicht für alle diese Richtungen erfüllt.  $\diamond$

#### Beispiel 3.4.

1.  $d = -\nabla f(x)$  ist stets eine Abstiegsrichtung, sofern  $x$  kein stationärer Punkt ist, denn dann gilt  $\nabla f(x)^T d = -\|\nabla f(x)\|^2 < 0$ .  $-\nabla f(x)$  ist sogar die Richtung des steilsten Abstiegs: Unter allen Richtungen  $d \in \mathbb{R}^n$  mit  $\|d\| = 1$  ist der normierte Gradient von  $f$  in  $x$  Minimalstelle von  $\min \nabla f(x)^T d$  u.d.N.  $\|d\| = 1$ . Wählt man  $d = -\nabla f(x)$ , so bezeichnet man das Abstiegsverfahren als **Gradientenverfahren**.
2. Ist  $M \in \mathcal{S}_n$  positiv definit, dann ist  $d = -M\nabla f(x)$  im Fall  $\nabla f(x) \neq 0$  eine Abstiegsrichtung, denn es gilt  $\nabla f(x)^T d = -\nabla f(x)^T M \nabla f(x) < 0$ . Man bezeichnet das Abstiegsverfahren dann als **gradientenähnliche Verfahren**.  $\diamond$



### 3.1. Abstiegsverfahren

Wir wollen in algorithmischer Form ein allgemeines Abstiegsverfahren angeben. Hier werden iterativ Punkte  $x^k \in \mathbb{R}^n$  erzeugt mit  $f(x^{k+1}) < f(x^k)$ , bis eine Abbruchbedingung beispielsweise von der Form  $\|x^k - x^*\| < \epsilon$  erfüllt ist.

---

#### Algorithmus 3.5 (Allgemeines Abstiegsverfahren)

---

**Eingabe:** Startpunkt  $x_0 \in \mathbb{R}^n$ .

- 1: Setze  $k = 0$ .
  - 2: **while** Konvergenzkriterium nicht erfüllt **do**
  - 3:   Bestimme Abstiegsrichtung  $d^k$  von  $f$  in  $x^k$ .
  - 4:   Bestimme eine Schrittweite  $t_k > 0$  mit  $f(x^k + t_k d^k) < f(x^k)$ .
  - 5:   Setze  $x^{k+1} = x^k + t_k d^k$  und  $k = k + 1$ .
  - 6: **end while**
- 

In theoretischen Konvergenzuntersuchungen betrachten wir kein Konvergenzkriterium, d.h. wir nehmen an, dass eine unendliche Folge  $(x^k)_{k \in \mathbb{N}}$  erzeugt wird, und untersuchen diese Folge auf Grenzwerte bzw. Häufungspunkte.

#### Satz 3.6. (Hinreichendes Konvergenzkriterium des Allgemeinen Abstiegsverfahrens)

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar und  $(x^k)_{k \in \mathbb{N}}$  eine durch Algorithmus 3.5 erzeugte Folge. Weiter mögen gelten:

1. Es gibt eine Konstante  $\theta_1 > 0$  existiert, so dass die folgende **Winkelbedingung** erfüllt ist:

$$-\nabla f(x^k)^T d^k \geq \theta_1 \|\nabla f(x^k)\| \|d^k\| \quad \text{für alle } k \in \mathbb{N}.$$

2. Es gibt eine von  $(x^k)_{k \in \mathbb{N}}$  und  $(d^k)_{k \in \mathbb{N}}$  unabhängige Konstante  $\theta_2 > 0$  existiert mit

$$f(x^k + t_k d^k) \leq f(x^k) - \theta_2 \left( \frac{\nabla f(x^k)^T d^k}{\|d^k\|} \right)^2 \quad \text{mit } t_k > 0 \text{ für alle } k \in \mathbb{N}.$$

Dann ist jeder Häufungspunkt der Folge  $(x^k)_{k \in \mathbb{N}}$  ein stationärer Punkt von  $f$ .

#### Beweis.

Sei  $x^*$  ein Häufungspunkt der durch Algorithmus 3.5 erzeugten Folge  $(x^k)_{k \in \mathbb{N}}$ . Dann gibt es eine Teilfolge  $(x^{k_\nu})_{\nu \in \mathbb{N}}$  mit  $x^{k_\nu} \rightarrow x^*$  für  $\nu \rightarrow \infty$ . Nach Konstruktion ist  $(f(x^k))_{k \in \mathbb{N}}$  monoton fallend, also folgt aus  $f(x^{k_\nu}) \rightarrow f(x^*)$  für  $\nu \rightarrow \infty$ , dass auch die gesamte Folge der Funktionswerte  $(f(x^k))_{k \in \mathbb{N}}$  gegen  $f(x^*)$  konvergiert.

Weiter liefern die beiden Bedingungen des Satzes die Abschätzung

$$0 \stackrel{k \rightarrow \infty}{\longleftarrow} f(x^{k+1}) - f(x^k) \leq -\theta_2 \left( \frac{\nabla f(x^k)^T d^k}{\|d^k\|} \right)^2 \leq -\theta_1^2 \theta_2 \|\nabla f(x^k)\|^2 \leq 0,$$

also ist  $\lim_{k \rightarrow \infty} \|\nabla f(x^k)\| = 0$ . Damit ist jeder Häufungspunkt von  $f$  auch ein stationärer Punkt von  $f$ .  $\square$

#### Bemerkung 3.7.

Bezeichne  $\eta_k$  den Winkel zwischen  $d^k$  und  $-\nabla f(x^k)$ , dann bedeutet die Winkelbedingung (1) aus Satz 3.6, dass

$$\cos(\eta_k) = \frac{-\nabla f(x^k)^T d^k}{\|\nabla f(x^k)\| \|d^k\|}$$

gleichmäßig größer als 0 ist. Ein wichtiges Beispiel einer Abstiegsrichtung, für welche diese Bedingung erfüllt ist, ist die Wahl  $d^k = -\nabla f(x^k)$ .  $\diamond$

### 3.2. Schrittweitenstrategien

Das Allgemeine Abstiegsverfahren (Algorithmus 3.5) besitzt in der Wahl der Abstiegsrichtung  $d^k$  und der Schrittweite  $t_k > 0$  große Freiheitsgrade. Die nahe liegende Minimierungsregel  $t_k = t_k^{\min}$  für die zu wählende Schrittweite mit

$$f(x^k + t_k^{\min} d^k) = \min_{t \geq 0} f(x^k + t d^k)$$

ist unter der Annahme, dass die Niveaumenge  $\mathcal{L}(x^0)$  kompakt ist und  $\nabla f$  Lipschitz-stetig ist auf  $\mathcal{L}(x^0)$ , wohldefiniert. Allerdings ist diese Regel im Allgemeinen nicht praktikabel: In jedem Schritt müsste dazu schließlich ein eindimensionales Optimierungsproblem exakt gelöst werden. Wir betrachten im Folgenden **Schrittweitenstrategien**, für welche die zugehörige Folge  $(x^k)_{k \in \mathbb{N}}$  gegen  $x^*$  konvergiert.

#### 3.2.1. Die Armijo-Regel

Wir behandeln **gradientenähnliche Richtungen** d.h.  $d = -M \nabla f(x)$  mit positiv definitem  $M \in \mathcal{S}_n$ . Sei dazu  $\alpha \in [0, 1]$  fest vorgegeben. Die **Armijo-Regel** ist eine Bedingung, die einen hinreichenden Abstieg sichert, und lautet

$$f(x + td) \leq f(x) + \alpha t \nabla f(x)^T d. \quad (3.2)$$

Wegen der stetigen Differenzierbarkeit von  $f$  und der Eigenschaft  $\nabla f(x)^T d < 0$  ist die Existenz eines  $t_0 > 0$  gesichert, so dass für alle  $t \in (0, t_0]$  die Bedingung (3.2) erfüllt ist.

Wir visualisieren die Armijo-Regel (3.2): Die durchgezogene Linie repräsentiert den Graphen der 1D-Abbildung  $f_1 : t \mapsto f(x + td)$  für  $t \geq 0$  während die gestrichelte Gerade die affin lineare Funktion  $f_2 : t \mapsto f(x) + \alpha t \nabla f(x)^T d$  darstellt.

Da  $d$  eine Abstiegsrichtung ist, ist die Existenz von  $a > 0$  gesichert, so dass  $f_2(t) \leq f_1(t)$  für alle  $t \in [0, a]$  erfüllt ist.

Zufällig ist die Armijo-Bedingung auch für alle  $t \in [b, c]$  erfüllt (dies wäre beispielsweise bei einer strikt konvexen Funktion nicht der Fall).

Zur tatsächlichen Berechnung einer geeigneten **Schrittweite**  $t$  überprüft man (3.2) sequenziell zum Beispiel für  $t = \beta^l$ ,  $l = 0, 1, 2, \dots$  mit einem fest vorgegebenen  $\beta \in (0, 1)$ , etwa  $\beta = \frac{1}{2}$ . Bei erstmaliger Gültigkeit von (3.2) bricht man ab.

Im Folgenden verallgemeinern wir die Wahl der Schrittweite: Ist (3.2) für ein  $t_c$  nicht erfüllt, dann wird ein  $t_+ < t_c$  so konstruiert, dass  $t_+ \in [\nu_0 t_c, \nu^0 t_c]$  gilt für gewisse Parameter  $0 < \nu_0 \leq \nu^0 < 1$ .

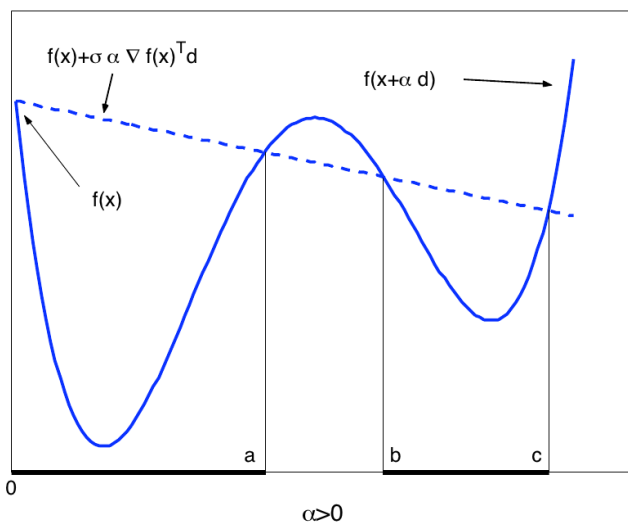


Fig. 4: Die Armijo-Regel für gradientenähnliche Abstiegsrichtungen.

---

#### Algorithmus 3.8 (Armijo-Schrittweitenalgorithmus)

---

**Eingabe:** Abstiegsrichtung  $d$ , Punkt  $x$ , Parameter  $\alpha, \nu_0, \nu^0$ .

- 1: Setze  $l = 0$  und  $t^{(0)} = 1$ .
  - 2: **while** Armijo-Bedingung (3.2) nicht erfüllt **do**
  - 3: Bestimme  $t^{(l+1)} \in [\nu_0 t^{(l)}, \nu^0 t^{(l)}]$  und setze  $l = l + 1$ .
  - 4: **end while**
- 

#### Lemma 3.9. (Endliche Terminierungsbedingung des Armijo-Schrittweitenalgorithmus)

Seien  $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ ,  $\nabla f$  Lipschitz-stetig zur Konstante  $L$ ,  $\alpha \in (0, 1)$ ,  $x \in \mathbb{R}^n$  mit  $\nabla f(x) \neq 0$  und  $M \in \mathcal{S}_n$  positiv definit. Bezeichne  $\lambda_s$  den kleinsten und  $\lambda_g$  den größten Eigenwert von  $M^{-1}$  und  $\varkappa = \lambda_g \lambda_s^{-1}$  die spektrale Konditionszahl von  $M^{-1}$ . Dann ist (3.2) für alle  $0 < t \leq 2\lambda_s(1 - \alpha)(L\varkappa)^{-1}$  erfüllt.

**Beweis.**

Es gilt

$$f(x + td) - f(x) = \int_0^1 \nabla f(x + \tau td)^T d \, d\tau.$$

Wir erhalten daraus

$$f(x + td) = f(x) + t \nabla f(x)^T d + t \int_0^1 (\nabla f(x + \tau td) - \nabla f(x))^T d \, d\tau.$$

Aus der Lipschitz-Stetigkeit des Gradienten von  $f$  folgt

$$f(x + td) \leq f(x) + t \nabla f(x)^T d + \frac{Lt^2}{2} \|d\|^2.$$

Wegen  $\lambda_g^{-1} \|z\|^2 \leq z^T M z \leq \lambda_s^{-1} \|z\|^2$  für alle  $z \in \mathbb{R}^n$  können wir  $\|d\|^2$  abschätzen via

$$\begin{aligned} \|d\|^2 &= \|M \nabla f(x)\|^2 = \nabla f(x)^T M^2 \nabla f(x) \leq \lambda_s^{-2} \|\nabla f(x)\|^2 \\ &\leq \lambda_g \lambda_s^{-2} \nabla f(x)^T M \nabla f(x) = -\varkappa \lambda_s^{-1} \nabla f(x)^T d. \end{aligned}$$

Einsetzen ergibt

$$f(x + td) \leq f(x) + t \left(1 - \frac{\varkappa L t}{\lambda_s}\right) \nabla f(x)^T d,$$

die Armijo-Bedingung (3.2) ist also erfüllt, falls  $\alpha < (1 - \frac{\varkappa L t}{\lambda_s})$  bzw.  $t \leq 2\lambda_s(1 - \alpha)(L\varkappa)^{-1}$ .  $\square$

**Lemma 3.10. (Endliche Terminierung & maximale Iterationszahl für gradientenähnliche Richtungen)**

Seien  $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ ,  $\nabla f$  Lipschitz-stetig zur Konstante  $L$ ,  $\alpha \in (0, 1)$ ,  $(x^k)_{k \in \mathbb{N}} \subseteq \mathbb{R}^n$  eine durch das Allgemeines Abstiegsverfahren 3.5 mit Armijo-Schrittweitenwahl 3.8 generierte Iterationsfolge. Die Folge  $(M^k)_{k \in \mathbb{N}} \subseteq \mathcal{S}_n$  der zu den Abstiegsrichtungen  $d^k = -M^k \nabla f(x^k)$  gehörigen, positiv definiten Matrizen erfülle  $\lambda_0 \leq \lambda_s^k \leq \lambda_g^k \leq \lambda^0$  für geeignete  $0 < \lambda_0 \leq \lambda^0$ . Dann gelten:

1. Jede Schrittweite  $t_k$  genügt der Abschätzung  $t_k \geq \tilde{t} = 2\nu_0 \lambda_0 (1 - \alpha) (L\tilde{\varkappa})^{-1}$  mit  $\tilde{\varkappa} = \lambda^0 \lambda_0^{-1}$ .
2. Jede Schrittweitensuche terminiert nach höchstens  $\log(2\lambda_0(1 - \alpha)(L\tilde{\varkappa})^{-1})(\log \nu^0)^{-1}$  Iterationen.

**Beweis.**

1. Mit  $\lambda_s^k \geq \lambda_0 > 0$  und  $\varkappa^k = \lambda_g^k \lambda_s^{-k} \leq \lambda^0 \lambda_0^{-1} = \tilde{\varkappa}$  folgt aus Lemma 3.10:

$$\forall k \in \mathbb{N} : \frac{2\lambda_s^k(1 - \alpha)}{L\varkappa} \geq \frac{2\lambda_0(1 - \alpha)}{L\tilde{\varkappa}} > 0 \quad \implies \quad \forall k \in \mathbb{N} : t_k \geq \tilde{t}.$$

2. Wegen  $t^{(0)} = 1$  und  $t^{(k+1)} \leq \nu^0 t^{(k)}$  wird eine zulässige Schrittweite  $t^{(k)}$  spätestens nach  $m$  Schritten gefunden mit

$$(\nu^0)^m < \frac{2\lambda_0(1 - \alpha)}{L\tilde{\varkappa}} \leq \frac{2\lambda_s^k(1 - \alpha)}{L\varkappa^k} \quad \implies \quad \frac{1}{\log \nu^0} \log \left( \frac{2\lambda_0(1 - \alpha)}{L\tilde{\varkappa}} \right). \quad \square$$

**Satz 3.11. (Hinreichendes Konvergenzkriterium des Abstiegsverfahrens mit Armijo-Schrittweiten)**

Seien  $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ ,  $\nabla f$  Lipschitz-stetig zur Konstante  $L$ ,  $\alpha \in (0, 1)$ ,  $(x^k)_{k \in \mathbb{N}} \subseteq \mathbb{R}^n$  eine durch das Allgemeines Abstiegsverfahren 3.5 mit Armijo-Schrittweitenwahl 3.8 generierte Iterationsfolge. Die Folge  $(M^k)_{k \in \mathbb{N}} \subseteq \mathcal{S}_n$  der zu den Abstiegsrichtungen  $d^k = -M^k \nabla f(x^k)$  gehörigen, positiv definiten Matrizen erfülle  $\lambda_0 \leq \lambda_s^k \leq \lambda_g^k \leq \lambda^0$  für geeignete  $0 < \lambda_0 \leq \lambda^0$ .

Ist  $(f(x^k))_{k \in \mathbb{N}}$  von unten beschränkt, so ist jeder Häufungspunkt von  $f$  ein stationärer Punkt von  $f$ .

**Beweis.**

Sei  $(f(x^k))_{k \in \mathbb{N}}$  von unten beschränkt. Da die Folge nach Konstruktion zugleich monoton fällt, folgt die Konvergenz und damit  $0 \leftarrow f(x^{k+1}) - f(x^k) < -\alpha \tilde{\kappa} \lambda_0^{-1} \|\nabla f(x^k)\|^2 \leq 0$ , also ist  $\lim \|\nabla f(x^k)\| = 0$ .  $\square$

**Bemerkung 3.12.**

1. Im Allgemeinen gibt es keine Garantie, dass ein (eindeutiger) Häufungspunkt existiert.
2. Die folgende Variation der Armijo-Regel führt ebenfalls zu einem im Sinne von Satz 3.11 konvergenten Verfahren: Seien  $r > 0$  ein *Skalierungsverfahren* und  $\beta \in (0, 1)$ . Bestimme  $t = \max\{r\beta^l\}$  für Exponenten  $l = 0, 1, 2, \dots$  so, dass die Armijo-Bedingung erfüllt ist. Diese Bestimmung der Schrittweite  $t$  wird in der englischsprachigen Literatur als *Backtracking* bezeichnet.
3. Weitere Strategien basieren auf Polynommodellen, die  $\varphi(t) = f(x + td)$  durch ein *quadratisches* oder *kubisches Modell* ersetzen und dann dieses Modell minimieren.  $\diamond$

**3.2.2. Die Wolfe-Powell-Regel**

Seien  $\alpha \in (0, \frac{1}{2})$  und  $\rho \in [\alpha, 1]$  gegeben. Die **Wolfe-Powell-Regel** lautet:

Zum Punkt  $x \in \mathbb{R}^n$  und zum Richtungsvektor  $d \in \mathbb{R}^n$  mit  $\nabla f(x)^T d < 0$  bestimme eine Schrittweite  $t > 0$  so, dass gelten

$$f(x + td) \leq f(x) + \alpha t \nabla f(x)^T d \quad (3.3a)$$

$$\nabla f(x + td)^T d \geq \rho \nabla f(x)^T d \quad (3.3b)$$

bzw. gleichbedeutend  $\varphi(t) \leq \varphi(0) + \alpha t \varphi'(0)$  und  $\varphi'(t) \geq \rho \varphi'(0)$ .

Zusätzlich zur Armijo-Regel (3.3a), welche sicherstellt, dass die gewählte Schrittweite  $t$  *hinreichend klein* ist, wird also gefordert, dass der Schritt nicht *zu klein* wird (3.3b).

Diese Bedingung wird dadurch motiviert, dass an einem lokalen Minimum  $t^*$  von  $\varphi$  gilt

$$\varphi'(t^*) = \nabla f(x + td)^T d = 0$$

und es daher sinnvoll ist, zu verhindern, dass  $\varphi$  bei  $t$  zu steil fällt.

In der Graphik ist die Armijo-Bedingung auf ganz  $[0, b]$  erfüllt, die Wolfe-Powell-Regel allerdings nur auf dem Teilintervall  $[a, b]$ . Weiterhin genügen auch alle Schrittweiten  $t \in [c, d]$  den Wolfe-Powell-Bedingungen.

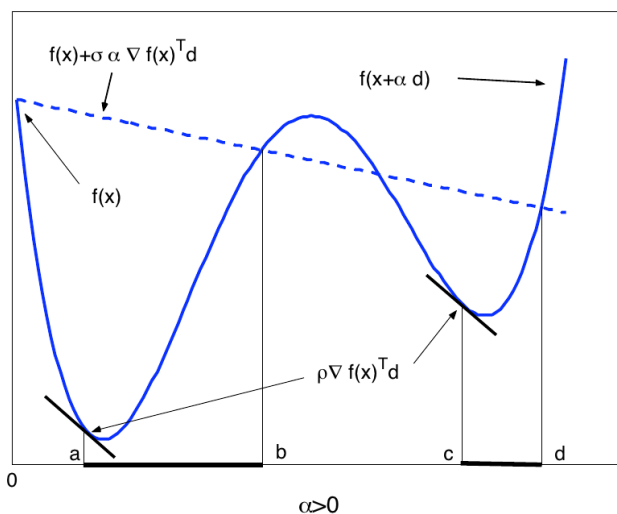


Fig. 5: Die Wolfe-Powell-Regel als modifizierte Armijo-Bedingung.

Ohne Beweis geben wir den folgenden Satz an:

**Satz 3.13. (Existenz zulässiger Wolfe-Powell-Schrittweiten)**

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar und  $\alpha \in (0, \frac{1}{2})$ ,  $\rho \in [\alpha, 1)$  und  $x^0 \in \mathbb{R}^n$  fest vorgegeben. Zu  $x \in \mathcal{L}(x^0)$  und einer Richtung  $d \in \mathbb{R}^n$  mit  $\nabla f(x)^T d < 0$  sei

$$\mathcal{T}_{\text{WP}}(x, d) = \{t > 0 \mid (3.3a) \text{ und } (3.3b) \text{ sind erfüllt}\}.$$

Dann gelten:

1. Ist  $f$  nach unten beschränkt, so ist  $\mathcal{T}_{\text{WP}}(x, d) \neq \emptyset$ , d.h. die Wolfe-Powell-Schrittweitenstrategie ist wohldefiniert.
2. Ist weiter  $\nabla f$  auf  $\mathcal{L}(x^0)$  Lipschitz-stetig, so gibt es eine Konstante  $\theta > 0$  (unabhängig von  $x$  und  $d$ ) mit

$$f(x + td) \leq f(x) - \theta \left( \frac{\nabla f(x)^T d}{\|d\|} \right)^2 \quad \text{für alle } x \in \mathcal{T}_{\text{WP}}(x, d).$$

**Algorithmus 3.14 (Wolfe-Powell-Schrittweitenalgorithmus)**

**Eingabe:** Abstiegsrichtung  $d \in \mathbb{R}^n$ , aktueller Punkt  $x \in \mathbb{R}^n$ , Parameter  $\alpha \in (0, \frac{1}{2})$ ,  $\rho \in [\alpha, 1]$ ,  $\gamma > 1$ , Kontraktionskonstanten  $\tau_1, \tau_2 \in (0, \frac{1}{2}]$ , Startschrittweite  $t^{(0)} > 0$ .

- 1: Setze  $i = 0$ .
- 2: **if**  $\varphi(t^{(i)}) \geq \varphi(0) + \alpha t^{(i)} \varphi'(0)$  **then**
- 3:   Setze  $a = 0$  &  $b = t^{(i)}$ , **goto** Zeile 11
- 4: **else**
- 5:   **if**  $\varphi'(t^{(i)}) \geq \rho \varphi'(0)$  **then**
- 6:     Wähle  $t = t^{(i)}$ ; **return** I
- 7:   **else**
- 8:     Setze  $t^{(i+1)} = \gamma t^{(i)}$  und  $i = i + 1$ ; **goto** Zeile 2
- 9:   **end if**
- 10: **end if**
- 11: Wähle  $j = 0$ ;  $t_1^{(0)} = a$ ;  $t_2^{(0)} = b$ ;  $\Delta^{(0)} = t_2^{(0)} - t_1^{(0)}$ .
- 12: Wähle  $t^{(j)} \in [t_1^{(j)} + \tau_1 \Delta^{(j)}, t_2^{(j)} - \tau_2 \Delta^{(j)}]$
- 13: **if**  $\varphi(t^{(j)}) \geq \varphi(0) + \alpha t^{(j)} \varphi'(0)$  **then**
- 14:   Setze  $t_1^{(j+1)} = t_1^{(j)}$ ;  $t_2^{(j+1)} = t^{(j)}$  und  $j = j + 1$ ; **goto** Zeile 12
- 15: **else**
- 16:   **if**  $\varphi'(t^{(j)}) \geq \rho \varphi'(0)$  **then**
- 17:     Wähle  $t = t^{(j)}$ ; **return** II
- 18:   **else**
- 19:     Setze  $t_2^{(j+1)} = t_2^{(j)}$ ;  $t_1^{(j+1)} = t^{(j)}$  und  $j = j + 1$ ; **goto** Zeile 12
- 20:   **end if**
- 21: **end if**

Um die Wohldefiniertheit der Wolfe-Powell-Schrittweitenstrategie 3.14 zu zeigen, zitieren wir zunächst folgendes Lemma:

**Lemma 3.15.**

Seien  $\alpha < \rho$  und  $\varphi'(0) < 0$ . Ist  $[a, b]$  mit  $0 \leq a < b$  ein Intervall, so dass

$$\varphi(a) \leq \varphi(0) + \alpha a \varphi'(0); \quad \varphi(b) \geq \varphi(0) + ab \varphi'(0) \quad \varphi'(a) < \alpha \varphi'(0), \quad (3.4)$$

so enthält  $[a, b]$  einen Punkt  $t_0$  mit

$$\varphi(t_0) < \varphi(0) + \alpha t_0 \varphi'(0); \quad \varphi'(t_0) = \alpha \varphi'(0) > \rho \varphi'(0).$$

Der Punkt  $t_0$  ist ein innerer Punkt eines Intervalls  $I$ , so dass für alle  $t \in I$  gilt

$$\varphi(t) \leq \varphi(0) + \alpha t \varphi'(0); \quad \varphi'(t) \geq \rho \varphi'(0),$$

d.h. alle Punkte von  $I$  sind erfüllen die Wolfe-Powell-Bedingungen:  $I \subseteq \mathcal{T}_{\text{WP}}(x, d)$ .

Mit diesem Lemma lässt sich nun zeigen, dass der Wolfe-Powell-Schrittweitenalgorithmus 3.14 nach endlich vielen Schritten terminiert und die ausgegebene Schrittweite in der zulässigen Menge  $\mathcal{T}_{\text{WP}}(x, d)$  liegt:

**Satz 3.16. (Endliche Terminierung des Wolfe-Powell-Schrittweitenalgorithmus)**

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar und nach unten beschränkt und  $\alpha \in (0, \frac{1}{2})$ ,  $\rho \in (\frac{1}{2}, 1)$ . Dann bricht Algorithmus 3.14 nach endlich vielen Schritten entweder bei **return** I oder bei **return** II mit einer Schrittweite  $t \in \mathcal{T}_{\text{WP}}(x, d)$  ab.

**Beweis.**

1. Terminiert der Wolfe-Powell-Schrittweitenalgorithmus bei **return** I, so genügt die generierte Schrittweite offenbar den Wolfe-Powell-Bedingungen.

2. Die erste Schleife, Zeilen 2 bis 10, terminiert nach endlich vielen Schritten: In dem Fall würden gelten  $t^{(i)} = \gamma t^{(0)}$  und  $\varphi(t^{(i)}) < \varphi(0) + \alpha t^{(i)} \varphi'(0)$  für alle  $i \in \mathbb{N}$ . Daraus würde aber folgen  $f(x + t^{(i)}d) < f(x) + \alpha t^{(i)} \nabla f(x)^T d$  für alle  $i$ ; wegen  $\gamma > 1$  und  $\varphi'(0) < 0$  wäre somit  $f$  nach unten unbeschränkt im Widerspruch zur Voraussetzung.
3. Sei jetzt also Position 11 erreicht. Dann hat das Intervall  $[a, b]$  die Eigenschaft aus Lemma 3.15 und es gilt  $\varphi'(a) < \rho \varphi'(0)$ . Terminiert der Algorithmus bei **return II**, so ist die Wolfe-Powell-Bedingung erfüllt.
4. Es bleibt also zu zeigen, dass auch die zweite Schleife, Zeilen 12 bis 21, nur endlich oft durchlaufen werden können. Per Induktion sieht man, dass das Intervall  $[t_1^{(j)}, t_2^{(j)}]$  (mit  $a = t_1^{(j)}$  und  $b = t_2^{(j)}$ ) für jedes  $j \in \mathbb{N}$  die Eigenschaft (3.4) besitzt und  $\varphi'(t_1^{(j)}) < \rho \varphi'(0)$  erfüllt ist. Würde nun eine unendliche Folge von Intervallen  $[t_1^{(j)}, t_2^{(j)}]$  erzeugt, so würde für diese gelten  $[t_1^{(j+1)}, t_2^{(j+1)}] \subseteq [t_1^{(j+1)}, t_2^{(j+1)}]$ ,  $\text{diam}[t_1^{(j)}, t_2^{(j)}] \rightarrow 0$  und die nach Lemma 3.15 existierende Folge von Punkten  $t_0^{(j)} \in (t_1^{(j)}, t_2^{(j)})$  mit  $\varphi(t_0^{(j)}) < \varphi(0) + \alpha t_0^{(j)} \varphi'(0)$  und  $\varphi'(t_0^{(j)}) > \rho \varphi'(0)$  würde somit gegen ein  $t_0$  konvergieren mit  $\varphi'(t_0) = \alpha \varphi'(0) > \rho \varphi'(0)$ . Andererseits würde  $\varphi'(t_1^{(j)}) < \rho \varphi'(0)$  für alle  $j$  wegen der Stetigkeit von  $\varphi'$  implizieren, dass auch  $\varphi'(t_0) \leq \rho \varphi'(0)$  erfüllt ist, ein Widerspruch.  $\square$

### 3.2.3. Die strenge Wolfe-Powell-Regel

Seien  $\alpha \in (0, \frac{1}{2})$  und  $\rho \in [\frac{1}{2}, 1)$  fest vorgegeben. Die **strenge Wolfe-Powell-Regel** lautet:

Zu  $x, d \in \mathbb{R}^n$  mit  $\nabla f(x)^T d < 0$  bestimme eine Schrittweite  $t > 0$  mit

$$f(x + td) \leq f(x) + \alpha t \nabla f(x)^T d, \quad (3.5a)$$

$$|\nabla f(x + td)^T d| \leq -\rho \nabla f(x)^T d. \quad (3.5b)$$

Im Gegensatz zur klassischen Wolfe-Powell-Regel wird also nicht nur gefordert, dass der Graph von  $\varphi$  nicht zu steil abfällt, sondern auch, dass er nicht zu steil ansteigt.

Für sehr kleines  $\rho$  (und damit auch kleines  $\alpha$ ) ist eine Schrittweite, welche Bedingung (3.5b) erfüllt, nahe an einem stationären Punkt von  $\varphi$ .

### 3.3. Praktische Aspekte

Die vorgestellten Schrittweitalgorithmen sind idealisiert. In der Praxis sind  $f$  und  $\nabla f$  nur maschinen- und/oder problemabhängig genau. Werden diese Ungenauigkeiten nicht berücksichtigt, führt dies schnell zu Endlosschleifen.

Ideal wäre es, wenn mit den Funktionswerten  $\varphi(t), \varphi(0)$  und den Ableitungen  $\varphi'(t), \varphi'(0)$  Fehlerschranken  $\varepsilon(t), \varepsilon(0)$  und  $\hat{\varepsilon}(t), \hat{\varepsilon}(0)$  mitgeliefert würden. Dann werden die beiden Bedingungen

$$\varphi(t) \leq \varphi(0) + \alpha t \varphi'(0) \quad \text{bzw.} \quad \varphi'(t) \geq \rho \varphi'(0)$$

ersetzt durch

$$\varphi(t) \leq \varphi(0) + \alpha t (\varphi'(0) + \hat{\varepsilon}(0)) + \varepsilon(0) + \varepsilon(t) \quad \text{bzw.} \quad \varphi'(t) \geq \rho (\varphi'(0) - \hat{\varepsilon}(0)) - \hat{\varepsilon}(t).$$

Weiter ist abzubrechen, wenn das Intervall  $[t_1^{(j)}, t_2^{(j)}]$  „zu klein“ wird, d.h. wenn  $t_2^{(j)} - t_1^{(j)}$  klein wird. Ferner sollte man stets eine untere Schranke für  $f$  mitführen.

### 3.4. Gradientenverfahren

Das allgemeine Abstiegsverfahren lässt noch einige Freiheiten in der Wahl der Abstiegsrichtung  $d^k$ . Eine nahe liegende Wahl für  $d$  – auch in Hinblick auf die Winkelbedingung aus Satz 3.6 – ergibt sich als Lösung von

$$\min \nabla f(x)^T d \quad \text{u.d.N.} \quad \|d\| = 1.$$

Das Ziel ist also,  $d$  als jene Richtung zu bestimmen, in welche  $f$  in  $x$  am steilsten fällt. Gemäß der Cauchy-Schwarzschen Ungleichung wird dieses Problem gelöst von  $d = -\frac{\nabla f(x)}{\|\nabla f(x)\|}$ .

Verwenden wir die Wolfe-Powell-Schrittweitenstrategie, dann folgt sofort aus den Sätzen 3.6 und 3.13, dass jeder Häufungspunkt der Folge  $(x^k)_{k \in \mathbb{N}}$  ein stationärer Punkt von  $f$  ist. Eine analoge Aussage gilt auch für die strenge Wolfe-Powell-Regel. Die Armijo-Bedingung hingegen erfüllt die Winkelbedingung nicht notwendigerweise, daher geben wir folgenden Satz an:

**Satz 3.17. (Konvergenz des Gradientenverfahrens mit Armijo-Schrittweitenstrategie)**

Ist  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar, so ist jeder Häufungspunkt einer durch das Allgemeine Abstiegsverfahren 3.5 konstruierten Folge  $(x^k)_{k \in \mathbb{N}}$  mit Abstiegsrichtungen  $d^k = \frac{-\nabla f(x^k)}{\|\nabla f(x^k)\|}$  und der Armijo-Schrittweitenstrategie 3.8 ein stationärer Punkt von  $f$ .

**Bemerkung 3.18.**

Das Konvergenzverhalten des steilsten Abstiegs kann sehr schlecht sein. Für eine *quadratische Form*  $f(x) = \frac{1}{2}x^T Q x + c^T x + \gamma$  mit positiv definiten Matrix  $Q \in \mathcal{S}_n$ ,  $c \in \mathbb{R}^n$  und  $\gamma \in \mathbb{R}$  lässt sich zeigen, dass

$$\|x^k - x^*\| \leq \sqrt{\varkappa} \left( \frac{\varkappa - 1}{\varkappa + 1} \right)^k \|x^0 - x^*\|,$$

erfüllt ist, wobei  $\varkappa = \frac{\lambda_g(Q)}{\lambda_s(Q)}$  die spektrale Konditionszahl bezeichnet.

Eine mögliche Abhilfe für die langsame Konvergenz des Verfahrens des steilsten Abstiegs besteht darin,  $d^k := -H^{-1} \nabla f(x^k)$  mit geeigneter positiv definiten Matrix  $H \in \mathcal{S}_n$  zu setzen.  $H$  soll überdies so gewählt sein, dass gilt

$$0 < \frac{\lambda_g(H^{-1}Q)}{\lambda_s(H^{-1}Q)} = \varkappa(H^{-1}Q) < \frac{\lambda_g(Q)}{\lambda_s(Q)} = \varkappa(Q). \quad \diamond$$

**Definition 3.19.**

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar und  $(x^k)_{k \in \mathbb{N}} \subseteq \mathbb{R}^n$ . Dann heißt  $(d^k)_{k \in \mathbb{N}} \subseteq \mathbb{R}^n$  **gradientenähnlich** bzgl.  $f$  und  $(x^k)_{k \in \mathbb{N}}$ , wenn für jede gegen einen nichtstationären Punkt von  $f$  konvergierende Teilfolge Konstanten  $c > 0$ ,  $\varepsilon > 0$ ,  $L > 0$  existieren, so dass gelten

$$\forall l \in \mathbb{N} : \|d^{k(l)}\| \leq c \quad \& \quad \forall l \geq L : \nabla f(x^{k(l)})^T d^{k(l)} \leq -\varepsilon.$$

**Bemerkung 3.20.**

1. Sei  $(H^k)_{k \in \mathbb{N}} \subseteq \mathcal{S}_n$  eine Folge positiv definiten Matrizen, so dass Konstanten  $c_1, c_2 > 0$  existieren mit

$$\forall x \in \mathbb{R}^n, \forall k \in \mathbb{N} : c_1 \|x\|^2 \leq x^T H^k x \leq c_2 \|x\|^2.$$

Dann ist  $(d^k)_{k \in \mathbb{N}}$ , gegeben durch  $H^k d^k = -\nabla f(x^k)$  für alle  $k \in \mathbb{N}$ , gradientenähnlich: Sei  $(x^{k(l)})_{l \in \mathbb{N}}$  eine konvergente Teilfolge von  $(x^k)_{k \in \mathbb{N}}$ , dann gilt

$$\|d^{k(l)}\| = \|(H^{k(l)})^{-1} \nabla f(x^{k(l)})\| \leq \|(H^{k(l)})^{-1}\| \|\nabla f(x^{k(l)})\| \leq \frac{1}{c_1} \|\nabla f(x^{k(l)})\| \leq C;$$

außerdem ist  $\frac{1}{c_2} \|x\|^2 \leq x^T (H^{k(l)})^{-1} x \leq \frac{1}{c_1} \|x\|^2$  für alle  $x$  erfüllt, also ist

$$-\nabla f(x^{k(l)})^T (H^{k(l)})^{-1} \nabla f(x^{k(l)}) \leq -\frac{1}{c_2} \|\nabla f(x^{k(l)})\|^2 \leq -\varepsilon,$$

2. Für Algorithmus 3.5 mit gradientenähnlichen Suchrichtungen und der Armijo-Schrittweitenstrategie gilt eine analoge Aussage zu Satz 3.17.

3. Manchmal bringt die Wahl  $H^k = \text{diag}(h_{ii}^k)$  mit  $h_{ii}^k = \frac{\partial^2 f}{\partial x_i^2}(x^k)$  für alle  $1 \leq i \leq n$  eine deutliche Verbesserung (wobei diese Matrizen nicht notwendig positiv definit sind).

4. Von besonderem Interesse ist die Wahl  $H^k = \nabla^2 f(x^k)$ . Mit diesen Abstiegsrichtungen beschäftigt sich das nächste Kapitel.  $\diamond$

## 4. Newton-Verfahren

### 4.1. Das lokale Newton-Verfahren

Wir setzen ab nun voraus, dass  $f$  und die lokale Minimalstelle  $x^*$  von  $f$  folgende Voraussetzungen erfüllen:

$$f \text{ ist zweimal stetig differenzierbar.} \quad (4.1a)$$

$$\exists \gamma > 0 : \forall x, y \in \mathbb{R}^n : \|\nabla^2 f(x) - \nabla^2 f(y)\| \leq \gamma \|x - y\|. \quad (4.1b)$$

$$x^* \text{ ist ein stationärer Punkt von } f: \nabla f(x^*) = 0. \quad (4.1c)$$

$$\nabla^2 f(x^*) \text{ ist positiv definit.} \quad (4.1d)$$

Zur Vereinfachung der Schreibweise bezeichne  $x_a$  den aktuellen Iterationspunkt und  $x_+$  die neue Iterierte. Wir betrachten ein **quadratisches Modell** von  $f$  um  $x_a$ :

$$m_a(x) = f(x_a) + \nabla f(x_a)^T(x - x_0) + \frac{1}{2}(x - x_0)^T \nabla^2 f(x_0)(x - x_0).$$

Wenn  $\nabla^2 f(x_a)$  positiv definit ist, dann definieren wir  $x_+$  als die (eindeutige) Minimalstelle dieses Modells. Es gilt  $0 = \nabla m_a(x_+) = \nabla f(x_a) + \nabla^2 f(x_a)(x_+ - x_a)$ . Umformen liefert die **Iterationsvorschrift des Newton-Verfahrens**, d.h.

$$x_+ = x_a - \nabla^2 f(x_a)^{-1} \nabla f(x_a). \quad (4.2)$$

Natürlich wird bei der numerischen Implementierung nicht  $\nabla^2 f(x_a)^{-1}$  berechnet, sondern es wird in jedem Iterationsschritt  $\nabla^2 f(x_a)d = -\nabla f(x_a)$  gelöst und  $x_+ = x_a + d$  gesetzt.

Falls  $x_a$  weit von einer lokalen Minimalstelle  $x^*$ , die 4.1 erfüllt, entfernt ist, dann kann  $\nabla^2 f(x_a)$  negative Eigenwerte haben. Also kann  $x_+$  eine lokale Minimalstelle oder ein Sattelpunkt sein. Um dies zu vermeiden, müssen geeignete Modifikationen eingeführt werden. Zunächst sei aber vorausgesetzt, dass  $x_a$  hinreichend nahe an  $x^*$  ist.

#### Satz 4.1. (Quadratische Konvergenz des lokalen Newton-Verfahrens)

1. Sei (4.1) erfüllt. Dann existieren Konstanten  $C > 0$  und  $\delta > 0$  derart, dass für alle  $x_a$  aus der Menge  $B(x^*, \delta) = \{x \in \mathbb{R}^n \mid \|x - x^*\| < \delta\}$  der **Newton-Schritt** (4.2) folgende Abschätzung erfüllt:

$$\|x_+ - x^*\| \leq C \|x_a - x^*\|^2.$$

2. Es sei (4.1) erfüllt. Dann existieren ein Radius  $\delta > 0$  und eine Konstante  $C > 0$ , so dass das Newton-Verfahren  $x^{k+1} = x^k - \nabla^2 f(x^k)^{-1} \nabla f(x^k)$  für  $x^0 \in B(x^*, \delta)$  gegen  $x^*$  konvergiert mit

$$\|x^{k+1} - x^*\| \leq C \|x^k - x^*\|^2.$$

#### Beweis.

1. Wähle  $\delta > 0$  derart, dass für alle  $x \in B(x^*, \delta)$  gelten

$$\|\nabla^2 f(x)\| \leq 2\|\nabla^2 f(x^*)\|, \quad \|(\nabla^2 f(x))^{-1}\| \leq 2\|(\nabla^2 f(x^*))^{-1}\|$$

sowie

$$\frac{\|x - x^*\|}{2\|(\nabla^2 f(x^*))^{-1}\|} \leq \|\nabla f(x)\| \leq 2\|\nabla^2 f(x^*)\| \|x - x^*\|.$$

Da sich die Differenz zwischen  $x_+$  und  $x^*$  darstellen lässt als

$$\begin{aligned} x_+ - x^* &= x_a - x^* - (\nabla^2 f(x_a))^{-1} \nabla f(x_a) = (\nabla^2 f(x_a))^{-1} (\nabla^2 f(x_a)(x_a - x^*) - \nabla f(x_a)) \\ &= (\nabla^2 f(x_a))^{-1} \int_0^1 (\nabla^2 f(x_a) - \nabla^2 f(x^* + t(x_a - x^*))) (x_a - x^*) dt, \end{aligned}$$



erhalten wir daraus

$$\begin{aligned} \|x_+ - x^*\| &\leq \|(\nabla^2 f(x_a))^{-1}\| \int_0^1 \|(\nabla^2 f(x_a) - \nabla^2 f(x^* + t(x_a - x^*)))\| dt \cdot \|x_a - x^*\| \\ &\leq 2\|(\nabla^2 f(x^*))^{-1}\| \gamma \int_0^1 \|x_a - x^* - t(x_a - x^*)\| dt \cdot \|x_a - x^*\| \\ &= 2\gamma\|(\nabla^2 f(x^*))^{-1}\| \cdot \|x_a - x^*\|^2 \int_0^1 1 - t dt = \underbrace{\gamma\|(\nabla^2 f(x^*))^{-1}\|}_{=C} \cdot \|x_a - x^*\|^2. \end{aligned}$$

2. Wähle  $\tilde{\delta} = \min(\delta, \frac{1}{2K})$ . Dann gilt für jedes  $x^k \in B(x^*, \tilde{\delta})$ , dass auch  $x^{k+1}$  in  $B(x^*, \tilde{\delta})$  liegt, denn

$$\|x^{k+1} - x^*\| \leq C\|x^k - x^*\|^2 \leq C\tilde{\delta}\|x^k - x^*\| \leq \frac{1}{2}\|x^k - x^*\| < \delta. \quad (4.3)$$

Aus  $x^0 \in B(x^*, \tilde{\delta})$  folgt also die Wohldefiniertheit des Newton-Verfahrens und  $(x^k)_{k \in \mathbb{N}} \subseteq B(x^*, \tilde{\delta})$ .  $\square$

**Bemerkung 4.2. (Terminierungsbedingung des lokalen Newton-Verfahrens)**

Ein natürliches Abbruchkriterium für das Newton-Verfahren (wie auch für die Gradientenverfahren des letzten Abschnitts) setzt sich aus einer **relativen** und einer **absoluten Fehlerschranke** zusammen: Sei  $\tau_{\text{rel}} \in (0, 1)$  eine erwünschte Reduktion in der Gradientennorm und  $\tau_{\text{abs}}$  mit  $1 \gg \tau_{\text{abs}} > 0$  eine absolute Fehlerschranke, dann stoppt man das Verfahren, wenn

$$\|\nabla f(x^k)\| \leq \tau_{\text{rel}}\|\nabla f(x^0)\| + \tau_{\text{abs}}$$

erfüllt ist. Ist  $\|\nabla f(x^0)\|$  groß, so ist  $\tau_{\text{rel}}\|\nabla f(x^0)\|$  der dominante Term. Ist hingegen  $\|\nabla f(x^0)\|$  klein, dann dominiert  $\tau_{\text{abs}}$ .  $\diamond$

**Bemerkung 4.3. (Approximationsfehler diskreter Ableitungen)**

Wir nehmen nun an, dass  $f$  nur approximativ ausgewertet werden kann. Betrage der Einfachheit halber die Raumdimension  $n = 1$ , dann ist

$$\tilde{f}(x) = f(x) + \tilde{\varepsilon}_f(x) \quad \text{mit } \tilde{\varepsilon}_f(x) \geq 0 \text{ und } |\tilde{\varepsilon}_f(x)| \leq \varepsilon_f \text{ für ein } \varepsilon_f > 0.$$

Bestimmen wir nun die Ableitungen numerisch, z.B. durch Vorwärtsdifferenzen, so ergibt sich

$$D_h^+ f(x) = \frac{\tilde{f}(x+h) - \tilde{f}(x)}{h}.$$

Wir schätzen die Ordnung der Differenz zwischen  $D_h^+ f$  und  $f'$  ab: Nach dem Satz von Taylor existiert eine Zwischenstelle  $\xi \in (x, x+h)$  mit

$$\begin{aligned} \|D_h^+ f(x) - f'(x)\| &= \left\| \frac{\tilde{f}(x+h) - \tilde{f}(x)}{h} - f'(x) \right\| \\ &= \left\| \frac{f(x+h) + \tilde{\varepsilon}_f(x+h) - f(x) - \tilde{\varepsilon}_f(x)}{h} - f'(x) \right\| \\ &\leq \left\| \frac{f(x+h) - f(x)}{h} - f'(x) \right\| + \frac{2\varepsilon_f}{h} \\ &\leq \left\| \frac{f(x) + f'(x)h + \frac{1}{2}f''(\xi)h^2 - f(x)}{h} - f'(x) \right\| + \frac{2\varepsilon_f}{h} \\ &= \frac{h}{2}\|f''(\xi)\| + \frac{2\varepsilon_f}{h} = \mathcal{O}\left(h + \frac{\varepsilon_f}{h}\right). \end{aligned}$$

Die Minimalstelle  $h^*$  der Fehlerfunktion  $\text{err}_+(h) = h + \frac{\varepsilon_f}{h}$  erfüllt  $\text{err}'_+(h^*) = 1 - \frac{\varepsilon_f}{(h^*)^2} = 0$ . In der Tat ist  $h^* = \sqrt{\varepsilon_f}$  Minimalstelle der Fehlerfunktion. Der Fehler  $\varepsilon_g$  in der Ableitung ist also von der Ordnung  $\mathcal{O}(h^*) = \mathcal{O}(\sqrt{\varepsilon_f})$ .

Verwenden wir nun abermals Vorwärtsdifferenzen zur Approximation der Hessematrix, so ergibt sich für den Fehler  $\varepsilon_H$  die Größenordnung  $\mathcal{O}(\sqrt{\varepsilon_g}) = \mathcal{O}(\varepsilon_f^{1/4})$ . Dies impliziert, dass Hessematrizen, basierend auf zweifacher numerischer Differenziation, relativ ungenau sind – selbst wenn  $\varepsilon_f = 10^{-16}$  die Größenordnung des **Maschinen-Epsilons** erreicht, folgt nur  $\varepsilon_H \approx 10^{-4}$ .

Im Falle **zentraler Differenzenapproximationen** ergibt sich ein besseres Ergebnis:  $\varepsilon_H = \mathcal{O}(\varepsilon_f^{4/9})$ . Betragen die Fehler in der ersten Ableitung  $\varepsilon_f = 10^{-16}$ , so erhalten wir so immerhin  $\varepsilon_H \approx 10^{-7.1}$ .

Konvergenz des Newton-Verfahrens ist nur zu erwarten, wenn  $\varepsilon_g \rightarrow 0$  im Laufe der Iteration gilt.  $\diamond$

**Satz 4.4. (Konvergenz des lokalen Newton-Verfahrens bei Diskretisierungsfehlern)**

Es seien die Bedingungen (4.1) erfüllt. Dann existieren Konstanten  $\bar{K} > 0$ ,  $\delta > 0$  und ein  $\varepsilon > 0$ , so dass für  $x_a \in B(x^*, \delta)$  und  $\|\varepsilon_H(x_a)\| < \varepsilon$  gilt:

$$x_+ = x_a - (\nabla^2 f(x_a) + \varepsilon_H(x_a))^{-1} (\nabla f(x_a) + \varepsilon_g(x_a))$$

ist wohldefiniert, d.h.  $(\nabla^2 f(x_a) + \varepsilon_H(x_a))$  ist regulär, und erfüllt

$$\|x_+ - x^*\| \leq \bar{K} (\underbrace{\|x_a - x^*\|^2}_{\text{beeinflusst Konv.geschw.}} + \underbrace{\|\varepsilon_H(x_a)\| \|x_a - x^*\|}_{\text{Genauigkeit}} + \underbrace{\|\varepsilon_g(x_a)\|}_{\text{Genauigkeit}}).$$

**Bemerkung 4.5.**

Wir betrachten das Newton-ähnliche Verfahren

$$x^{k+1} = x^k - \nabla^2 f(x^0)^{-1} \nabla f(x^k), \quad x^0 \text{ Startwert, } k = 0, 1, \dots \quad (4.4)$$

Im Gegensatz zum klassischen Newton-Verfahren ist es hier nicht notwendig, in jedem Schritt die Hessematrix  $\nabla^2 f(x^k)$  aufzustellen. Es gelten

$$\varepsilon_H(x^k) = \nabla^2 f(x^0) - \nabla^2 f(x^k), \quad \|\varepsilon_H(x^k)\| \leq \varepsilon_H, \quad (4.5a)$$

$$\|\varepsilon_H(x^k)\| = \|\nabla^2 f(x^0) - \nabla^2 f(x^k)\| \leq \gamma \|x^0 - x^k\| \leq \gamma (\|x^0 - x^*\| + \|x^* - x^k\|). \quad (4.5b)$$

Die Konvergenz des Verfahrens (4.4) folgt aus Satz 4.4 mit  $\varepsilon_g = 0$  und  $\varepsilon_H = \mathcal{O}(\|x^0 - x^*\|)$ .  $\diamond$

**Satz 4.6. (Lineare Konvergenz des Newton-ähnlichen Verfahrens)**

Es sei (4.1) erfüllt. Dann existieren Konstanten  $K \in (0, 1)$  und  $\delta > 0$ , so dass für  $x^0 \in B(x^*, \delta)$  gilt: Die Folge der Iterierten  $(x^k)_{k \in \mathbb{N}}$ , erzeugt durch das Verfahren (4.4), konvergiert linear gegen  $x^*$  und man hat

$$\|x^{k+1} - x^*\| \leq K \|x^* - x^k\|.$$

**Beweis.**

Sei  $\delta > 0$  so gewählt, dass Satz 4.4 angewendet werden kann. Mit den Bedingungen (4.5) folgt:

$$\begin{aligned} \|x^{k+1} - x^*\| &\leq \bar{K} (\|x^k - x^*\|^2 + \gamma (\|x^0 - x^*\| + \|x^* - x^k\|) \|x^k - x^*\|) \\ &= \bar{K} (\underbrace{\|x^k - x^*\|}_{\leq \delta} (1 + \gamma) + \underbrace{\gamma \|x^0 - x^*\|}_{\leq \delta}) \|x^k - x^*\| \leq \bar{K} (1 + 2\gamma) \delta \|x^k - x^*\|. \end{aligned}$$

Um Konvergenz zu garantieren, verkleinere  $\delta$ , sodass  $\bar{K}(1 + 2\gamma)\delta < 1$  erfüllt ist.  $\square$

**4.2. Das inexakte Newton-Verfahren**

Während beim klassischen Newtonverfahren das Update  $x_+ = x_a + d_a$  in jedem einzelnen Schritt das (exakte) Lösen des Gleichungssystems  $\nabla^2 f(x_a) d_a = -\nabla f(x_a)$  erfordert, verwendet man beim **inexakten Newton-Verfahren** eine approximative Richtung  $\tilde{d}_a$ , die mit geringerem Aufwand berechnet werden kann und für ein  $\eta_a > 0$  die folgende Bedingung erfüllt:

$$\|\nabla^2 f(x_a) \tilde{d}_a + \nabla f(x_a)\| \leq \eta_a \|\nabla f(x_a)\|. \quad (4.6)$$

Wir wissen, dass  $\nabla^2 f(x_a)$  positiv definit ist für  $x_a$  nahe  $x^*$ . Daher eignet sich zum Beispiel das **CG-Verfahren** (Verfahren der konjugierten Gradienten) zur iterativen Lösung von  $\nabla^2 f(x_a)\tilde{d}_a = -\nabla f(x_a)$ . Man spricht dann vom **Newton-CG-Verfahren**.

**Satz 4.7. (Konvergenzverhalten des inexacten Newton-Verfahrens)**

1. Es sei (4.1) erfüllt. Dann existieren eine Konstante  $K \geq 0$  und ein Radius  $\delta > 0$ , so dass für alle Punkte  $x_a \in B(x^*, \delta)$  mit zugehöriger Richtung  $\tilde{d}$ , welche (4.6) erfüllt, und Update  $x_+ = x_a + \tilde{d}$  gilt:

$$\|x_+ - x^*\| \leq K(\|x_a - x^*\| + \eta_a)\|x_a - x^*\|.$$

2. Sei Voraussetzung (4.1) erfüllt. Dann existieren  $\delta > 0$  und  $\bar{\eta} \geq 0$ , sodass die inexacte Newton-Iteration  $x^{k+1} = x^k + \tilde{d}^k$  mit

$$\|\nabla^2 f(x^k)\tilde{d}^k + \nabla f(x^k)\| \leq \eta_k \|\nabla f(x^k)\| \quad \& \quad (\eta_k)_{k \in \mathbb{N}} \subseteq [0, \bar{\eta}]$$

linear gegen  $x^*$  konvergiert.

Weiterhin gelten:

- a) Falls  $\eta_k \xrightarrow{k \rightarrow \infty} 0$ , dann ist die Konvergenz **superlinear**, d.h.  $\lim_{k \rightarrow \infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 0$ .

- b) Falls  $\eta_k \leq K_\eta \|\nabla f(x^k)\|$  für ein  $K_\eta > 0$  erfüllt ist, so ist die Konvergenz sogar quadratisch.

#### 4.2.1. Das Newton-CG-Verfahren

Bezeichne  $D_h^2 f(x; d)$  eine hinreichend genaue Approximation von  $\nabla^2 f(x)d$ , etwa eine Diskretisierung der zweiten Ableitung von  $f$  im Punkt  $x$  in Richtung  $d$ . Wir betrachten im Folgenden eine vorkonditionierte CG-Methode, welche mit einer Fehlermeldung abbricht, falls  $\nabla^2 f(x)$  in Richtung  $d$  degeneriert, d.h.  $d^T \nabla^2 f(x)d = 0$  gilt, oder  $d$  eine Richtung mit negativer Krümmung ist, d.h.  $d^T \nabla^2 f(x)d < 0$  erfüllt ist.

**Algorithmus 4.8 (Newton-CG-Verfahren)**

**Eingabe:** Startpunkt  $x_0 \in \mathbb{R}^n$ , Toleranzen  $\tau_{\text{rel}}, \tau_{\text{abs}} > 0$ .

- 1: Setze  $r_0 = \|\nabla f(x_0)\|$  und  $k = 0$ .
- 2: **while**  $\|\nabla f(x^k)\| > \tau_{\text{rel}} r_0 + \tau_{\text{abs}}$  **do**
- 3:   Wähle Parameter  $\eta_k > 0$  und positiv definite Matrix  $W^k \in \mathcal{S}_n$ .
- 4:   Setze  $d^0 = 0$ ,  $r^0 = \nabla f(x^k)$ ,  $p^0 = -(W^k)^{-1} r^0$  und  $l = 0$ .
- 5:   **while**  $\|r^l\| > \eta_k \|\nabla f(x^k)\|$  **do**
- 6:     Bestimme  $w^l = D_h^2 f(x, p^l)$ .
- 7:     **if**  $(p^l)^T w^l = 0$  **then**
- 8:       **return** mit Fehlermeldung Indefinitheit.
- 9:     **else if**  $(p^l)^T w^l < 0$  **then**
- 10:       **return** mit Fehlermeldung Negative Krümmung.
- 11:     **else**
- 12:       Setze  $\alpha_l = \frac{(r^l)^T (W^k)^{-1} r^l}{(p^l)^T w^l}$ ,  $d^{l+1} = d^l + \alpha_l p^l$  und  $r^{l+1} = r^l + \alpha_l w^l$ .
- 13:       Setze  $\beta_{l+1} = \frac{(r^{l+1})^T (W^k)^{-1} r^{l+1}}{(r^l)^T (W^k)^{-1} r^l}$ ,  $p^{l+1} = -(W^k)^{-1} r^{l+1} + \beta_{l+1} p^l$  und  $l = l + 1$ .
- 14:     **end if**
- 15:   **end while**
- 16:   Bestimme Update  $x^{k+1} = x^k + d^k$ .
- 17:   **if**  $f(x^{k+1}) > f(x^k)$  **then**
- 18:     **return** mit Fehlermeldung Kein Abstieg.
- 19:   **end if**
- 20: **end while**

Wir werden im Kontext von Trust-Region-Verfahren sehen, wie im Falle negativer Krümmungen verfahren werden kann.

### 4.3. Das globalisierte Newton-Verfahren

Bisher haben wir im Zusammenhang mit dem Newton-Verfahren *lokale Konvergenzaussagen* hergeleitet; wir mussten dabei stets voraussetzen, dass der Startpunkt  $x^0$  hinreichend nahe beim Minimum  $x^*$  gewählt war. Um auf diese manchmal problematische Forderung verzichten zu können, führen wir *Globalisierungen* ein, welche die Startpunktwahl relaxieren.

Kann man sicherstellen, dass die **Newton-Iterationsmatrix**  $\nabla^2 f(x^k)$  – oder eine entsprechende Approximation für diese – die Bedingung

$$\forall k \in \mathbb{N}, \forall d \in \mathbb{R}^n : c_1 \|d\|^2 \leq d^T \nabla^2 f(x^k) d \leq c_2 \|d\|^2$$

für gewisse Konstanten  $c_1, c_2 > 0$  erfüllt, so ist  $d^k$  als Lösung von  $\nabla^2 f(x^k) d^k = -\nabla f(x^k)$  eine gradientenähnliche Abstiegsrichtung. Eingesetzt in das Allgemeine Abstiegsverfahren 3.5, folgt dann die globale Konvergenz des **globalisierten Newton-Verfahrens**, d.h. der Startpunkt  $x^0$  kann dann beliebig gewählt werden.

#### 4.3.1. Die Trust-Region-Methode

Das klassische Newton-Verfahren könnte als Abstiegsverfahren mit gradientenähnlichen Richtungen aufgefasst werden (wobei das Problem der Schrittweitsuche entfällt), sofern sichergestellt wäre, dass alle Komponenten der Matrixfolge  $(\nabla^2 f(x^k))_{k \in \mathbb{N}}$  positiv definit wären. Das Ziel der Trust-Region-Methode besteht darin, das Verfahren des steilsten Abstiegs, welches keine Anforderungen an die zweiten Ableitungen stellt, in geeigneter Weise mit dem Newton-Verfahren, das höhere Konvergenzraten zur Verfügung stellt, zu verbinden. Wir verwenden dabei eine Umgebung, in der wir einem Modell von  $f$  *vertrauen*.

Sei  $m_a$  ein quadratisches Modell von  $f$  um  $x_0$ , gegeben durch

$$m_a(x) = f(x_a) + \nabla f(x_a)^T (x - x_a) + \frac{1}{2} (x - x_a)^T \nabla^2 f(x_a) (x - x_a).$$

Weiter sei  $\Delta$  der Radius einer Kugel um  $x_a$ , in welcher wir dem Modell von  $f$  trauen.  $\Delta$  nennt man den **Trust-Region-Radius** und die Kugel  $\mathcal{T}(\Delta) = \{x \in \mathbb{R}^n \mid \|x - x_a\| \leq \Delta\}$  **Trust-Region**. Die nächste Iterierte wird dann als approximative Minimalstelle von  $m_a$  in  $\mathcal{T}(\Delta)$  gewählt. Das zugehörige **Trust-Region-Hilfsproblem** lautet daher

$$\min m_a(x + d) \quad \text{u.d.N. } \|d\| \leq \Delta. \quad (4.7)$$

Wir bezeichnen die Lösung von (4.7), die (**Versuchslösung**, mit  $d_V$  und setzen  $x_V = x_a + d_V$ . Im Wesentlichen überprüft man, ob das quadratische Modell eine „gute“ Approximation von  $f$  in  $\mathcal{T}(\Delta)$  ist. Dazu definiert man

$$\begin{aligned} \text{ared}_a &= f(x_a) - f(x_V) && \text{(tatsächliche Reduktion)} \\ \text{pred}_a &= m_a(x_a) - m_a(x_V). && \text{(erwartete Reduktion)} \end{aligned}$$

Es gilt (mit  $H_a = \nabla^2 f(x_a)$ ):

$$\begin{aligned} \text{pred}_a &= m_a(x_a) - m_a(x_V) \\ &= f(x_a) - f(x_a) - \nabla f(x_a)^T (x_V - x_a) - \frac{1}{2} (x_V - x_a)^T H_a (x_V - x_a) \\ &= -\nabla f(x_a)^T (x_V - x_a) - \frac{1}{2} (x_V - x_a)^T H_a (x_V - x_a). \end{aligned}$$

Im folgenden Algorithmus benötigen wir Parameter  $0 < \mu_0 \leq \underline{\mu} < \bar{\mu}$ , um zu entscheiden, ob

1. der Versuchsschritt verworfen wird (im Fall  $\frac{\text{ared}_a}{\text{pred}_a} < \mu_0$ )
2. und/oder ob der Trust-Region verkleinert werden soll (falls  $\frac{\text{ared}_a}{\text{pred}_a} < \underline{\mu}$ ),
3. ob er vergrößert werden soll (falls  $\frac{\text{ared}_a}{\text{pred}_a} > \bar{\mu}$ )
4. oder ob man ihn unverändert belässt.

Die Änderung von  $\Delta$  wird mit Hilfe von  $0 < \underline{\omega} < 1 < \bar{\omega}$  durchgeführt. Weiter sei eine Konstante  $C > 1$  gegeben.

**Algorithmus 4.9 (Trust-Region-Routine)****Eingabe:**  $x_a \in \mathbb{R}^n$ ,  $x_V \in \mathbb{R}^n$ ,  $\Delta \in \mathbb{R}^+$ .

- 1: Setze  $z^0 = x_a$ ,  $z_V^0 := x_V$ ,  $\hat{\Delta}^{(0)} = \Delta$  und  $l = 0$ .
- 2: **while**  $z^l = x_a$  **do**
- 3:   Setze  $\text{ared}^{(l)} = f(x_a) - f(z_V^l)$  und  $d_V^l = z_V^l - x_a$ .
- 4:   Setze  $\text{pred}^{(l)} = -\nabla f(x_a)^T d_V^l - \frac{1}{2}(d_V^l)^T H_a d_V^l$ .
- 5:   **if**  $\frac{\text{ared}^{(l)}}{\text{pred}^{(l)}} < \mu_0$  **then**
- 6:     Wähle  $z^{l+1} = x_a$  und setze  $\hat{\Delta}^{(l+1)} = \underline{\omega} \hat{\Delta}^{(l)}$ .
- 7:     **if**  $l > 1$  &  $\hat{\Delta}^{(l)} > \hat{\Delta}^{(l-1)}$  **then**
- 8:       Wähle  $z^{l+1} = z_V^l$  und setze  $\hat{\Delta}^{(l+1)} = \hat{\Delta}^{(l-1)}$ .
- 9:     **else**
- 10:       Berechne die Lösung  $d_V^{l+1}$  des Trust-Region-Hilfsproblems (4.7) mit Radius  $\hat{\Delta}^{(l+1)}$ .
- 11:       Wähle  $z_V^{l+1} = x_a + d_V^{l+1}$ .
- 12:     **end if**
- 13:   **else if**  $\mu_0 \leq \frac{\text{ared}^{(l)}}{\text{pred}^{(l)}} \leq \underline{\mu}$  **then**
- 14:     Wähle  $z^{l+1} = z_V^l$  und setze  $\hat{\Delta}^{(l+1)} = \underline{\omega} \hat{\Delta}^{(l)}$ .
- 15:   **else if**  $\underline{\mu} \leq \frac{\text{ared}^{(l)}}{\text{pred}^{(l)}} \leq \bar{\mu}$  **then**
- 16:     Wähle  $z^{l+1} = z_V^l$  und setze  $\hat{\Delta}^{(l+1)} = \hat{\Delta}^{(l)}$ .
- 17:   **else if**  $\bar{\mu} \leq \frac{\text{ared}^{(l)}}{\text{pred}^{(l)}}$  **then**
- 18:     **if**  $\|d_V^l\| = \hat{\Delta}^{(l)} \leq C \|\nabla f(x_a)\|$  **then**
- 19:       Wähle  $z^{l+1} = x_a$  und setze  $\hat{\Delta}^{(l+1)} = \bar{\omega} \hat{\Delta}^{(l)}$ .
- 20:       Berechne die Lösung  $d_V^{l+1}$  des Trust-Region-Hilfsproblems (4.7) mit Radius  $\hat{\Delta}^{(l+1)}$ .
- 21:       Wähle  $z^{l+1} = x_a + d_V^{l+1}$ .
- 22:     **else**
- 23:       Wähle  $z^{l+1} = z_V^l$  und setze  $\hat{\Delta}^{(l+1)} = \hat{\Delta}^{(l)}$ .
- 24:     **end if**
- 25:   **end if**
- 26:   Setze  $l = l + 1$ .
- 27: **end while**
- 28: Bestimme Updates des Punktes  $x_+ = z^l$  und des Trust-Region-Radius  $\Delta_+ = \hat{\Delta}^{(l)}$ .

In Algorithmus 4.9 ist der Trust-Region-Radius nach oben durch  $C \|\nabla f(x_a)\|$  beschränkt. Die While-Schleife sollte nach endlich vielen Schritten terminieren. Die Trust-Region-Routine wird nun in folgendem Programm iterativ aufgerufen:

**Algorithmus 4.10 (Trust-Region-Framework)****Eingabe:** Startpunkt  $x^0 \in \mathbb{R}^n$ , Startradius  $\Delta_0 > 0$ .

- 1: Setze  $k = 0$  und  $\tau_0 = \|\nabla f(x^0)\|$ .
- 2: **while**  $\|\nabla f(x^k)\| > \tau_{\text{rel}} \tau_0 + \tau_{\text{abs}}$  **do**
- 3:   Berechne eine Approximation  $H^k$  der Hessematrix  $\nabla^2 f(x^k)$ .
- 4:   Berechne eine Versuchslösung  $d_V^k$  via  $\min f(x^k) + \nabla f(x^k)d + \frac{1}{2}d^T H^k d$  u.d.N.  $\|d\| \leq \Delta_k$ .
- 5:   Berechne  $(x^{k+1}, \Delta_{k+1})$  durch Algorithmus 4.9 mit Inputs  $x^k$ ,  $x_V^k = x^k + d_V^k$  und  $\Delta_k$ , setze  $k = k + 1$ .
- 6: **end while**

**4.3.2. Globale Konvergenz des Trust-Region-Verfahrens****Bemerkung 4.11.**

Wir studieren das Trust-Region-Konvergenzverhalten. Dazu treffen wir die folgenden Annahmen:

1. Es existiert eine Konstante  $\sigma > 0$ , so dass gilt

$$\text{pred}_a = m_a(x_a) - m_a(x_V) = f(x_a) - m_a(x_V) \geq \sigma \|\nabla f(x_a)\| \min\{\|d_V\|, \|\nabla f(x_a)\|\}.$$

2. Es existiert ein  $M > 0$ , so dass gilt  $\|d_V\| \geq \frac{\|\nabla f(x_a)\|}{M}$  oder  $\|d_V\| = \Delta_a$ . ◇

**Satz 4.12. (Globale Konvergenz des Trust-Region-Verfahrens)**

Sei  $\nabla f$  Lipschitz-stetig mit Konstante  $L > 0$ . Sei die Folge  $(x_n)_{n \in \mathbb{N}}$  erzeugt von Algorithmus 4.10 und sei angenommen, dass die Lösungen des Trust-Region-Hilfsproblems (4.7) die beiden Voraussetzungen aus Bemerkung 4.11 erfüllen. Ferner seien die Matrizen  $(H^k)_{k \in \mathbb{N}}$  beschränkt. Dann ist entweder  $f$  nach unten unbeschränkt,  $\nabla f(x^k) = 0$  für ein  $k$  oder es gilt  $\lim_{k \rightarrow \infty} \|\nabla f(x^k)\| = 0$ .

**Bemerkung 4.13.**

Eine einfache Idee zur Lösung des Hilfsproblems beruht auf dem Fixieren der Richtung gemäß des Verfahrens des steilsten Abstiegs unter Berücksichtigung des Trust-Regions. Seien  $x_a$  die aktuelle Iterierte und  $\Delta_a$  der aktuelle Trust-Region-Radius. Der Versuchspunkt  $x_V = x_V(t) = x_a - t_a \nabla f(x_a)$  ist dann definiert über die Minimalstelle  $t_a$  von

$$\min_{t \geq 0} \psi_a(t) = m_a(x_a - t \nabla f(x_a)) \quad \text{u.d.N. } x_V(t) - x_a \in \mathcal{T}(\Delta_a).$$

Es gelten

$$\begin{aligned} \psi_a(t) &= f(x_a) - t \|\nabla f(x_a)\|^2 + \frac{t^2}{2} \nabla f(x_a)^T H_a \nabla f(x_a), \\ \psi'_a(t) &= -\|\nabla f(x_a)\|^2 + t \nabla f(x_a)^T H_a \nabla f(x_a). \end{aligned}$$

Zur Bestimmung von  $t_a$  unterscheiden wir zwei Fälle (vgl. das Problem mit negativen Krümmungen beim Newton-CG-Verfahren 4.8):

1.  $\nabla f(x_a)^T H_a \nabla f(x_a) \leq 0$ : Offensichtlich wird die Trust-Region-Restriktion aktiv, d.h.

$$\|x_V(t_a) - x_a\| = t_a \|\nabla f(x_a)\| = \Delta_a \quad \implies \quad t_a = \frac{\Delta_a}{\|\nabla f(x_a)\|}.$$

2.  $\nabla f(x_a)^T H_a \nabla f(x_a) > 0$ : Dann impliziert die Optimalitätsbedingung  $\psi'(t_a) = 0$ , dass  $t_a$  gegeben ist als

$$t_a = \frac{\|\nabla f(x_a)\|^2}{\nabla f(x_a)^T H_a \nabla f(x_a)}.$$

Zusammenfassend setzen wir

$$t_a = \begin{cases} \frac{\Delta_a}{\|\nabla f(x_a)\|} & \text{falls } \nabla f(x_a)^T H_a \nabla f(x_a) \leq 0 \\ \min\left\{ \frac{\Delta_a}{\|\nabla f(x_a)\|}, \frac{\|\nabla f(x_a)\|^2}{\nabla f(x_a)^T H_a \nabla f(x_a)} \right\} & \text{falls } \nabla f(x_a)^T H_a \nabla f(x_a) > 0 \end{cases}.$$

Die Minimalstelle  $x_V(t_a)$  ( $:= x_a - t_a \nabla f(x_a)$ ) des quadratischen Modells  $m_a$  in Richtung des negativen Gradienten heißt **Cauchy-Punkt** und wird mit  $x_a^{\text{CP}}$  bezeichnet. Man kann zeigen, dass der Cauchy-Punkt Annahme 4.11 erfüllt.  $\diamond$

**Bemerkung 4.14.**

1. Die Verwendung des Cauchy-Punktes führt zwar zu globaler Konvergenz, aber unter Umständen ist die Konvergenzgeschwindigkeit lokal nur linear.
2. Die sog. **dogleg-Technik** leistet einen „glatten“ Übergang von der Richtung des steilsten Abstiegs zur Newton-Richtung. Lokal liegt dann quadratische Konvergenz vor, wenn  $H^k = \nabla^2 f(x^k)$ ,  $k \in \mathbb{N}$ , gilt.  $\diamond$

**4.4. Quasi-Newton-Verfahren**

Wir haben gesehen, dass das (klassische) Newton-Verfahren eine Reihe von Nachteilen hat:

- Zweite Ableitungen werden benötigt (Rechenaufwand).
- Positive Definitheit muss gesichert werden.
- $\mathcal{O}(n^3)$  Multiplikationen (Lösung des linearen Systems mit direktem Verfahren).

**Quasi-Newton-Verfahren** gleichen diese Nachteile teilweise aus:

- Zweite Ableitungen werden durch erste Ableitungen approximiert.
- Positive Definitheit bleibt bei einigen QN-Verfahren erhalten.
- $\mathcal{O}(n^2)$  Multiplikationen bei QN-Varianten, welche direkt  $\nabla f(x^k)^{-1}$  approximieren.

$H^k$  wird dabei von Iteration zu Iteration aufdatiert. Die allgemeine Grundstruktur ist wie folgt:

1. Wähle eine Richtung  $d^k = -(H^k)^{-1}\nabla f(x^k)$ .
2. Bestimme das Update  $x^{k+1} = x^k + t_k d^k$  mit Schrittweitenstrategie.
3. Verwende  $x^k$ ,  $x^{k+1}$  und  $H^k$ , um  $H^k$  zu  $H^{k+1}$  aufzudatieren.

**Bemerkung 4.15. (Aufdatierungsformeln)**

Seien  $s_a = t_a d_a = -t_a H_a^{-1} \nabla f(x_a)$ ,  $y_a = \nabla f(x_+) - \nabla f(x_a)$  und  $x_+ = x_a + s_a$ . Dann gilt

$$\begin{aligned} y_a &= \nabla f(x_+) - \nabla f(x_a) \\ &= \nabla f(x_a) + \nabla^2 f(x_a)(x_+ - x_a) + \mathcal{O}(\|x_+ - x_a\|) - \nabla f(x_a) \\ &= \nabla^2 f(x_a)s_a + \mathcal{O}(\|x_+ - x_a\|) + \mathcal{O}(\|s_a\|). \end{aligned}$$

1. Eine nahe liegende Forderung für das Update  $H_+$  von  $H_a$  ist die **QN-Bedingung** oder auch **Sekantenbedingung**  $H_+ s_a = y_a$ . Ein einfacher Ansatz für  $H_+$  ist dann die **symmetrische Rang 1-Korrektur**  $H_+ = H_a + \alpha u u^T$  für zu bestimmendes  $\alpha \in \mathbb{R}$  und  $u \in \mathbb{R}^n$ . Einsetzen in die QN-Bedingung liefert  $H_+ s_a = H_a s_a + \alpha u (u^T s_a) = y_a$ , d.h.  $u$  ist proportional zu  $y_a - H_a s_a$ . Wähle daher  $u = y_a - H_a s_a$  und berücksichtige die Länge im Parameter  $\alpha$ : Aus  $\alpha (u^T s_a) = 1$  folgt  $\alpha = (y_a^T s_a - s_a^T H_a s_a)^{-1}$ . Also lautet die Update-Formel für  $H$ :

$$H_+ = H_a + \frac{(y_a - H_a s_a)(y_a - H_a s_a)^T}{(y_a - H_a s_a)^T s_a}.$$

Die Rang-1-Korrektur hat aber zwei Nachteile: Die positive Definitheit von  $H$  geht meist verloren und  $y_a - H_a s_a$  liegt nahe bei 0.

2. Flexibler sind **Rang 2-Korrekturen**  $H_+ = H_a + \alpha u u^T + \beta v v^T$ . Eingesetzt in die QN-Bedingung ergibt sich  $H_+ s_a = H_a s_a + \alpha u (u^T s_a) + \beta v (v^T s_a) = y_a$ . Die Vektoren  $u$  und  $v$  sind dann allerdings nicht mehr eindeutig bestimmt. Es bietet sich an,  $u = y_a$  und  $v = H_a s_a$  zu wählen (**Broyden/Fletcher/Goldfarb/Shanno**). Dann gilt

$$H_a s_a + \alpha y_a y_a^T s_a + \beta (H_a s_a)(H_a s_a)^T s_a = y_a \iff \alpha y_a (y_a^T s_a) + \beta H_a s_a (s_a^T H_a s_a) = y_a - H_a s_a$$

ergeben sich  $\alpha (y_a^T s_a) = 1$  und  $\beta (s_a^T H_a s_a) = -1$ . Die **BFGS-Formel** zum Update der Hessematrix lautet dann

$$H_+ = H_a + \frac{y_a y_a^T}{y_a^T s_a} - \frac{H_a s_a (H_a s_a)^T}{s_a^T H_a s_a}.$$

3. Man kann auch die inverse Hessematrix  $\nabla^2 f(x^k)^{-1}$  durch ein geeignetes  $B^k$  approximieren. Die QN-Bedingung lautet in dem Fall  $B_+ y_a = s_a$ . Verwenden wir die symmetrische Rang 2-Korrektur mit  $u = s_a$  und  $v = B_a y_a$  (**Davidon/Fletcher/Powell**), dann erhalten wir die (**DFP-Formel**)

$$B_+ = B_a + \frac{s_a s_a^T}{s_a^T y_a} - \frac{(B_a y_a)(B_a y_a)^T}{y_a^T B_a y_a}. \quad \diamond$$

**Lemma 4.16. (Positive Definitheit des BFGS-Updates)**

Seien  $H_a \in \mathcal{S}_n$  positiv definit,  $y_a^T s_a > 0$  und  $H_+$  gemäß der BFGS-Formel bestimmt.

Dann ist  $H_+$  symmetrisch und positiv definit.

**Beweis.**

Aus  $H_a$  positiv definit und  $y_a^T s_a \neq 0$  folgt für alle Vektoren  $z \neq 0$ , dass

$$z^T H_a z = z^T H_a z + \frac{z^T y_a y_a^T z}{y_a^T s_a} - \frac{z^T (H_a s_a) (H_a s_a)^T z}{s_a^T H_a s_a} = \frac{(z^T y_a)^2}{y_a^T s_a} + z^T H_a z - \frac{(z^T H_a s_a)^2}{s_a^T H_a s_a}$$

erfüllt ist. Da  $H_a$  symmetrisch und positiv definit ist, existiert  $H_a^{1/2}$  mit  $H_a = H_a^{1/2} H_a^{1/2}$ . Damit gilt

$$\begin{aligned} z^T H_a s_a &= z^T H_a^{1/2} H_a^{1/2} s_a \leq \|H_a^{1/2} z\| \|H_a^{1/2} s_a\| \\ \implies (z^T H_a s_a)^2 &\leq \underbrace{\|H_a^{1/2} z\|^2}_{z^T H_a z} \underbrace{\|H_a^{1/2} s_a\|^2}_{s_a^T H_a s_a} \\ \implies z^T H_a z - \frac{(z^T H_a s_a)^2}{s_a^T H_a s_a} &\geq z^T H_a z - z^T H_a z = 0. \end{aligned}$$

Ferner gilt  $z^T H_a z - \frac{(z^T H_a s_a)^2}{s_a^T H_a s_a} = 0$ , wenn  $H_a^{1/2} z$  und  $H_a^{1/2} s_a$  linear abhängig sind. Da  $H_a^{1/2}$  regulär ist, gilt in diesem Fall  $z = \lambda s_a \neq 0$ , d.h.  $\lambda \neq 0$ . Also ist  $z^T y_a = \lambda s_a^T y_a \neq 0$  und wir erhalten

$$z^T H_a z \geq \frac{(z^T y_a)^2}{y_a^T s_a} > 0 \text{ für } H_a^{1/2} z = \lambda H_a^{1/2} s_a \quad \text{und} \quad z^T H_a z > \frac{(z^T y_a)^2}{y_a^T s_a} \geq 0 \text{ sonst.} \quad \square$$

**Bemerkung 4.17.**

Die Bedingung  $y_a^T s_a > 0$  ist realistisch. Für quadratische Probleme von der Form  $f(x) = \frac{1}{2} x^T G x + g^T x + b$  mit symmetrischer, positiv definiter Hessematrix  $G$  gilt die Beziehung

$$y_a^T s_a = (\nabla f(x_+) - \nabla f(x_a))^T (x_+ - x_a) = (x_+ - x_a)^T G (x_+ - x_a) > 0 \quad \text{für } x_+ \neq x_a.$$

Für allgemeine Probleme wird  $y_a^T s_a > 0$  durch Verwendung von Schrittweitenstrategien (Wolfe-Powell) sichergestellt.  $\diamond$

**Satz 4.18. (Lokale Konvergenz des BFGS-Verfahrens)**

Es sei Annahme 4.1 erfüllt. Dann existiert ein  $\delta > 0$ , so dass für

$$\|x^0 - x^*\| \leq \delta \quad \text{und} \quad \|H^0 - \nabla^2 f(x^*)\| \leq \delta$$

die BFGS-Methode wohldefiniert ist und superlinear gegen ein  $x^*$  konvergiert, d.h.

$$\lim_{k \rightarrow \infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 0.$$

**Übungsaufgaben**

**Exercise 1. (Local extreme points)**

Find all local extreme points of the **Rosenbrock function**

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x_1, x_2) = 100 \cdot (x_2 - x_1^2)^2 + (1 - x_1)^2.$$

**Exercise 2. (Distance optimization)**

Let  $a, b, c, d$  be vectors in  $\mathbb{R}^n$  with  $(b, d)$  linear independent. We parametrize two lines  $X, Y \subseteq \mathbb{R}^n$  by  $x(s) = a + sb$  and  $y(t) = c + td$  ( $s, t \in \mathbb{R}$ ). Find all global extreme points of the distance function  $\text{dist} : (s, t) \mapsto \|x(s) - y(t)\|$ .



**Exercise 3. (Convexity criteria)**

Consider the quadratic function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(x) = \frac{1}{2}\langle x, Qx \rangle + \langle c, x \rangle + \gamma$  where  $Q \in \mathcal{S}_n$ ,  $c \in \mathbb{R}^n$  and  $\gamma \in \mathbb{R}$  ( $\mathcal{S}_n$  denotes the vector space of symmetric  $n \times n$  matrices).

Show that Remark 2.9 holds, true, i.e.

1.  $f$  is convex  $\Leftrightarrow Q$  is positive semidefinite,
  2.  $f$  is strictly convex  $\Leftrightarrow f$  is uniformly convex  $\Leftrightarrow Q$  is positive definite.
- 

**Exercise 4. (Exact step sizes)**

Consider again the quadratic function in equation  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(x) = \frac{1}{2}\langle x, Qx \rangle + \langle c, x \rangle + \gamma$  where in addition  $Q$  is assumed to be positive definite. Let  $x^k \in \mathbb{R}^n$  be arbitrary and  $d^k \in \mathbb{R}^n$  be a descent direction of  $f$  in  $x^k$ . Find the exact step-length  $t^k$  in direction  $d^k$  such that the decreasing of  $f$  is maximal.

---

**Exercise 5. (Convergence of the general descent method)**

Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a continuous function,  $(x^k)_{k \in \mathbb{N}} \subseteq \mathbb{R}^n$  a sequence generated by the general descent method, Algorithm 3.5. Show that if  $x^*$  and  $x^{**}$  are two accumulation points of the sequence  $(x^k)_{k \in \mathbb{N}}$ , then  $f(x^*) = f(x^{**})$  holds true.

---

**Exercise 6. (Ill-posed step size strategies)**

Consider the general descent method, Algorithm 3.5 for the parabel function  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x^2$ , with starting point  $x^0 = 1$  and direction  $d^k$  and step size  $t^k$  as follows:

1.  $d^k = -1$  and  $t^k = (\frac{1}{2})^{k+2}$  with  $k \in \mathbb{N}_0$ ,
2.  $d^k = (-1)^{k+1}$  and  $t^k = 1 + \frac{3}{2^{k+2}}$  with  $k \in \mathbb{N}_0$ .

Verify that these choices of the parameters for  $k \in \mathbb{N}_0$  lead to a decreasing of the function  $f$ . In order to do that, present the sequence  $x^k$  generated by the Algorithm 3.5 using induction with respect to  $k$ . Determine in both cases  $\lim f(x^k)$  and compare them to the minimum of  $f(x)$ .

---

**Exercise 7. (Program: Armijo step size algorithm)**

1. Implement the Armijo step size Algorithm 3.8 using MATLAB. Generate a file `armijo.m` for the function

`t = armijo(fhandle,x0,d,t0,alpha,beta)`

where `fhandle` represents the handle to a function, `x0` the initial point, `d` the descent direction, `t0` the initial step-size and `alpha`, `beta` the parameters. The program should return a step-size `t` which complies with the Armijo condition 3.2.

2. Test your program by using the Rosenbrock function  $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$  with points  $x = (x_1, x_2)^T \in \mathbb{R}^2$  and the following parameter sets:

- a) `x0=[1.7;1.5]`, `d=[-1;0]`, `t0=4`, `alpha=0.1` and `beta=0.5`
- b) `x0=[0;0]`, `d=[1;0]`, `t0=1`, `alpha=0.1` and `beta=0.5`.

Implement the rosenbrock function in a separate function file `rosenbrock.m` and call the algorithm from a main file `main.m` which includes the parameter samples.

**Hint:** For later on, it would be useful if the function `rosenbrock` would return the value of the first derivative in addition to the function value, i.e. `[valf,gradf]=rosenbrock(x)`.

---

**Exercise 8. (Program: Steepest descent method)**

1. Implement now the general descent method 3.5 with descent directions  $d^k = -\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|}$  using the Armijo step size strategy 3.8. Generate a file `gradmethod.m` for the function

`X = gradmethod(fhandle,x0,epsilon,t0,alpha, beta)`

with initial point  $x_0$ , tolerance `epsilon` for the termination condition  $\|\nabla f(x_k)\| < \epsilon$  and parameters `t0`, `alpha` and `beta` for the Armijo rule. The program should return a matrix  $X = [x_0; x_1; x_2; \dots]$  containing the whole history of points generated within the iterations.

2. Test your program by using the following functions:

- a) The function  $f(x) = \cos(x)$  with `x0=1.1656`, `epsilon=1.0e-3`, `t0=1`, `alpha=1.0e-2`, `beta=0.5`.
- b) The Rosenbrock function with `x0=[1;-0,5]`, `epsilon=1.0e-2`, `t0=1`, `alpha=1.0e-2`, `beta=0.5`.

**Hint:** Comment your code in detail: What is this variable for? How is the function defined – what are the input arguments and output value? What are their dimensions? What is the function doing? What is computed in the loops?

**Exercise 9. (Descent directions)**

Consider the function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  with  $(x_1, x_2) \mapsto (x_1 + x_2^2)^2$  in the point  $x = (1, 0)$ . Show that  $d = (-1, 1)$  is a direction of descent and find all minimizers of the problem  $\min f(x + \alpha d)$  u.d.N.  $\alpha > 0$ .

**Exercise 10. (Local minimizers and saddle points)**

Consider the function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  with  $(x_1, x_2) \mapsto 3x_1^4 - 4x_1^2x_2 + x_2^2$ . Prove that  $x_0 = (0, 0)$  is a stationary point of  $f$ . Show that  $f$ , restricted on any line through  $x_0$ , has a strict local minimum in  $x_0$ . Is  $x_0$  a local minimizer of  $f$ ?

**Exercise 11. (Linear Regression)**

Let  $\mathcal{H}$  denote a  $\mathbb{K}$ -Hilbert space with scalar product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$  where  $\mathbb{K}$  is one of the fields  $\mathbb{R}, \mathbb{C}$ .

1. Let  $b \in \mathcal{H}$  and  $A \in \mathcal{L}_b(\mathcal{H}, \mathcal{H})$ , the space of all linear, continuous operators on  $\mathcal{H}$ . Show that  $x_0 \in \mathcal{H}$  is a minimal point of  $\varphi : \mathcal{H} \rightarrow \mathbb{R}$  with  $x \mapsto \|Ax - b\|_{\mathcal{H}}$  if and only if the **Gaussian normal equation**  $A^*Ax_0 = A^*b$  holds.

**Hint:**  $A^*$  denotes the *adjoint operator* to  $A$ , i.e. the well-defined, implicitly given operator  $A \in \mathcal{L}_b(\mathcal{H}, \mathcal{H})$  such that  $\langle Ax, y \rangle_{\mathcal{H}} = \langle x, A^*y \rangle_{\mathcal{H}}$  holds for all  $x, y \in \mathcal{H}$ .

2. Use this characterization to solve the following **linear regression problem**: Find parameters  $x_1, x_2 \in \mathbb{R}$  such that the corresponding regression line  $\gamma_x : \mathbb{R} \rightarrow \mathbb{R}, t \mapsto x_1 + x_2t$  approximates the measuring points

$t_i$	1975	1980	1985	1990	1995
$\gamma_i$	30	35	38	42	44

optimally, i.e.  $(x_1, x_2)$  is the minimizer of  $(y_1, y_2) \mapsto \sum_{i=1}^5 (\gamma_i - \gamma_y(t_i))^2$ .

**Exercise 12. (Best approximation)**

Again, let  $\mathcal{H}$  denote a  $\mathbb{K}$ -Hilbert space with scalar product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$  where  $\mathbb{K}$  is one of the fields  $\mathbb{R}, \mathbb{C}$ . Let  $x \in \mathcal{H}$  and  $F$  be a convex, nonempty, closed subset of  $\mathcal{H}$ .

1. Show that there is a unique Point  $y \in F$  such that  $\|x - y\|_{\mathcal{H}} = \text{dist}(x, F)$  holds.

**Hint:**  $\text{dist}$  denotes the distance function  $\text{dist}(y_0, Y) = \inf_{y \in Y} \|y_0 - y\|_{\mathcal{H}}$  on  $\mathcal{H}$ .

2. Show that for all  $x \in F$  the following equivalence holds true:

$$\|x_0 - x\|_{\mathcal{H}} = \text{dist}(x_0, F) \iff \forall y \in F : \text{Re}\langle x_0 - x, y - x \rangle_{\mathcal{H}} \leq 0.$$

**Exercise 13. (Wolfe-Powell Lemma)**

Let  $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$  and  $d_k \in \mathbb{R}^n$  be a direction of descent in the point  $x_k \in \mathbb{R}^n$ . Further, assume that  $f$  is limited from below on the ray  $\{x_k + td_k \mid t > 0\}$ . Show that for any given parameters  $0 < \alpha < \rho < 1$  there is a step size  $t > 0$  such that the Wolfe-Powell conditions

$$\begin{aligned} f(x_k + td_k) &\leq f(x_k) + \alpha t \langle \nabla f(x_k), d_k \rangle \\ \langle \nabla f(x_k + td_k), d_k \rangle &\geq \rho \langle \nabla f(x_k), d_k \rangle \end{aligned}$$

or the strict Wolfe-Powell conditions

$$\begin{aligned} f(x_k + td_k) &\leq f(x_k) + \alpha t \langle \nabla f(x_k), d_k \rangle \\ |\langle \nabla f(x_k + td_k), d_k \rangle| &\leq \rho |\langle \nabla f(x_k), d_k \rangle|, \end{aligned}$$

respectively, hold in an open neighbourhood of  $t$ .

---

**Exercise 14. (Finite termination of the Wolfe-Powell algorithm)**

Show that the step size intervals  $[t_1^{(j)}, t_2^{(j)}]$  which are iteratively generated in line 14 of the Wolfe-Powell step size algorithm 3.14 satisfies the conditions

$$\varphi(a) \leq \varphi(0) + \alpha a \varphi'(0); \quad \varphi(b) \geq \varphi(0) + \alpha b \varphi'(0) \quad \varphi'(a) < \alpha \varphi'(0),$$

with  $\varphi(t) = f(x + td)$  and that  $\varphi'(t_1^{(j)}) < \rho \varphi'(0)$  holds for all  $j \in \mathbb{N}$ .

---

**Exercise 15. (Convergence speed of the steepest descent method)**

Consider the quadratic form  $f(x) = \frac{1}{2}x^T Qx + c^T x + \gamma$  with symmetric, positive definite matrix  $Q \in \mathbb{R}^{n \times n}$ , vector  $c \in \mathbb{R}^n$  and scalar  $\gamma \in \mathbb{R}$ . Let  $(x^k)_{k \in \mathbb{N}}$  be generated by the steepest descent method 3.5. Show that

$$\|x^k - x^*\| \leq \sqrt{\varkappa} \left( \frac{\varkappa - 1}{\varkappa + 1} \right)^k \|x^0 - x^*\|$$

holds where  $\varkappa = \frac{\lambda_g(Q)}{\lambda_s(Q)}$  denotes the spectral condition number of  $Q$ .

---

**Exercise 16. (Necessary first-order optimality condition for constrained optimization problems)**

Let  $\Omega \subseteq \mathbb{R}^n$  be a compact interval,  $\Omega = [a_1, b_1] \times \dots \times [a_n, b_n]$ , and  $f \in \mathcal{C}^1(\Omega, \mathbb{R})$ , i.e.  $f \in \mathcal{C}^1(\Omega^\circ, \mathbb{R})$  and there is a continuous continuation of  $\nabla f$  on  $\partial\Omega$ .

- Let  $x^* \in \Omega$  be a local minimizer of  $f$ , i.e.  $\exists \epsilon > 0 : \forall x \in B_\epsilon(x^*) \cap \Omega : f(x^*) \leq f(x)$ . Prove that the following first-order condition holds:

$$\forall x \in \Omega : \langle \nabla f(x^*), x - x^* \rangle \geq 0.$$

**Remark:** Any  $x^*$  that fulfills this condition is called a **stationary point** of  $f$ .

- Let  $\mathcal{P} : \mathbb{R}^n \rightarrow \Omega$  denote the projection  $(\mathcal{P}x)_i = \max(\min x_i, b_i), a_i$  and  $x(\lambda) = \mathcal{P}(x - \lambda \nabla f(x))$ . Prove that

$$\forall x, y \in \Omega : \langle y - x(\lambda), x(\lambda) - x + \lambda \nabla f(x) \rangle \geq 0.$$


---

**Exercise 17. (Armijo condition for constrained optimization problems)**

Assume that  $\nabla f$  is Lipschitz continuous with Lipschitz constant  $L$ . Prove that for all  $\lambda \in (0, \frac{2(1-\alpha)}{L}]$ , the **projected Armijo condition** holds:

$$f(x(\lambda)) - f(x) \leq -\frac{\alpha}{\lambda} \|x - x(\lambda)\|^2.$$


---

**Exercise 18. (Program: Projected gradient algorithm)**

The projected gradient method is a modification of the steepest descent method 3.5 which can deal with optimal points on the boundary.

A suitable termination criterion in case of possible minimizers with active constraints is

$$\|x - x(1)\| < \varepsilon \quad \text{or} \quad \|\nabla f(x)\| < \varepsilon \quad \text{or} \quad k \geq k_{\max}$$

where  $\varepsilon$  denotes a small tolerance parameter and  $k_{\max}$  is the maximal number of iterations to be executed.

**Algorithm (Projected Gradient Method)**

**Input:** Point  $x^0 \in \mathbb{R}^n$ , tolerance  $\varepsilon > 0$ , maximal number of iterations  $k_{\max} \in \mathbb{N}$ , initial step size  $t_0 > 0$ , step size parameters  $\alpha, \beta > 0$ .

- 1: Set  $k = 0$ .
- 2: **while** convergence criterion not fulfilled **do**
- 3:   Choose steepest descent direction  $d^k$ .
- 4:   Determine suitable projected Armijo step size  $t_k$ .
- 5:   Update  $x^{k+1} = \mathcal{P}(x^k + t_k d^k)$  and  $k = k + 1$ .
- 6: **end while**

1. Implement the modified Armijo step size algorithm `projarmijo` and the projected gradient method `projgradmethod` with the termination criterion given above.

**Hint:** Additionally, the vectors of boundary points  $a, b \in \mathbb{R}^n$  have to be handed over both to the gradient method and to the Armijo algorithm. As before, the gradient method shall return the whole history matrix of iteration points.

2. Test your program by using the Rosenbrock function with the following parameters and comment on the results: `epsilon=1.0e-2, N=1.0e+04, t0=1, alpha=1.0e-02, beta=0.5` and
  - a) `x0=[1;-0,5]`, `a=[-1;-1]` and `b=[2;2]`
  - b) `x0=[-1;-0,5]`, `a=[-2;-2]` and `b=[2;0]`

**Exercise 19. (Convergence of the projected gradient method)**

Our objective is to prove that the generated sequence has a convergent subsequence which converges towards a stationary point of  $f$  in analogy to the unconstrained case, cp. Theorem 3.6.

Let  $(x_n)_{n \in \mathbb{N}}$  be an iteration sequence generated by `gradproj`.

1. Show that  $(f(x_n))_{n \in \mathbb{N}}$  converges.
2. Show that  $(x_n)_{n \in \mathbb{N}}$  has at least one convergent subsequence and that all accumulation points of  $(x_n)_{n \in \mathbb{N}}$  are stationary points of  $f$ .
3. Show that  $x^*$  is a stationary point of  $f$  if and only if  $x^* = P(x^* - \lambda \nabla f(x^*))$  holds.

**Exercise 20. (Discrete derivatives and approximation errors)**

Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a four times differentiable, approximative function – which means that the function values  $f(x)$  are not known exactly, but with some error tolerance  $\epsilon_{\text{tol}}$ :

$$f_{\text{known}}(x) = f_{\text{exact}}(x) + \epsilon_{\text{tol}}(x) \quad \text{where } |\epsilon_{\text{tol}}(x)| \leq \epsilon \text{ for some known } \epsilon > 0.$$

Determine the derivative of  $f$  numerically by central differences and prove that the error  $\epsilon_H$  arising in the numerical approximation of the second derivative of  $f$  is of order  $\epsilon_H = O(\epsilon^{4/9})$ .

**Hint:** The task is to estimate the difference between  $f''_{\text{exact}}$  and  $D_h^d D_h^c f_{\text{known}}$  where  $D_h^c f(x) = \frac{f(x+h) - f(x-h)}{2h}$ .

**Exercise 21. (Discrete approximation of the Hessian matrix)**

Let  $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ . Verify the formula

$$\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f(x) = \frac{f(x + \epsilon e_i + \epsilon e_j) - f(x + \epsilon e_i) - f(x + \epsilon e_j) + f(x)}{\epsilon^2} + O(\epsilon)$$

for the approximation of the Hessian matrix by evaluations of  $f$ .

---

**Exercise 22. (Angle condition for Newton-like methods)**

The angle  $\eta_k$  between the search direction  $d_k$  and the steepest descent direction  $-\nabla f(x_k)$  is

$$\cos \eta_k = \frac{-\langle \nabla f(x_k), d_k \rangle}{\|\nabla f(x_k)\|_2 \|d_k\|_2}.$$

Consider the Newton-like method  $x_{k+1} = x_k + \alpha_k d_k$  with direction  $d_k = -B_k^{-1} \nabla f(x_k)$  where all matrices  $B_k \in \mathbb{R}^{n \times n}$  are positive definite with uniformly bounded condition numbers, i.e. there is some  $M > 0$  such that

$$\text{cond}_2(B_k) = \|B_k\|_2 \|B_k^{-1}\|_2 \leq M \quad \text{for all } k \geq 0.$$

Show that  $\cos \eta_k \geq \frac{1}{M}$  holds true.

**Hint:** First prove that  $\|Bx\|_2 \geq \frac{\|x\|_2}{\|B^{-1}\|_2}$  for any non-singular matrix  $B$  holds.

---

**Exercise 23. (Quadratic models and curvature condition)**

Let  $\phi$  be given by  $\phi(\alpha) = f(x_k + \alpha d_k)$  where  $d_k$  is a descent direction, i.e.  $\langle \nabla f(x_k), d_k \rangle < 0$  holds true.

1. Show that the quadratic function interpolating  $\phi(0)$ ,  $\phi'(0)$  and  $\phi(\alpha_0)$  is given by

$$\phi_q(\alpha) = \left( \frac{\phi(\alpha_0) - \phi(0) - \phi'(0)\alpha_0}{\alpha_0^2} \right) \alpha^2 + \phi'(0)\alpha + \phi(0)$$

by making the ansatz  $\phi_q(\alpha) = a_0 + a_1\alpha + a_2\alpha^2$  and using the interpolators to calculate the coefficients  $a_0, a_1, a_2$  of  $\phi_q$ .

2. Assume now that the sufficient decrease condition  $f(x_k + \alpha d_k) \leq f(x_k) + c_1 \alpha \langle \nabla f(x_k), d_k \rangle$  is not satisfied at  $\alpha_0$ . Show that  $\phi_q$  has positive curvature and that the minimizer  $\alpha^*$  of  $\phi_q$  satisfies  $\alpha^* < \frac{\alpha_0}{2(1-c_1)}$ .

**Hint:** Since  $c_1$  is chosen to be quite small in practise, this indicates that  $\alpha^*$  cannot be much greater than  $\frac{\alpha_0}{2}$  (and may be smaller) which gives us an idea of the new step length.

---

**Exercise 24. (Trust region auxiliary problem and Cauchy point)**

A simple strategy to solve the trust region auxiliary problem (4.7) approximatively bases on the steepest descent method, respecting the radius  $\Delta_k$  for which we trust the model: Consider

$$(\star) \quad \min_{p \in \mathbb{R}^n} f(x_k) + \langle \nabla f(x_k), p \rangle \text{ s.t. } \|p\| \leq \Delta_k, \quad (\star\star) \quad \min_{0 \leq t \leq 1} m_k(x_k + tp_k) \text{ s.t. } \|tp_k\| \leq \Delta_k$$

where the vector  $p_k$  in  $(\star\star)$  is the solution to  $(\star)$ . With the solution  $t_k$  of  $(\star\star)$  we define the **Cauchy point**  $x_k^{\text{CP}}$  by  $x_k^{\text{CP}} = x_k + t_k p_k$ .

1. Assume that  $x_k$  is no stationary point of  $f$ . Find the solution  $p_k^*$  to the minimization problem  $(\star)$ .

2. Show that the solution  $t_k^*$  of  $(\star\star)$  is given by

$$t_k^* = \begin{cases} 1 & \text{if } \langle \nabla f(x_k), H_k \nabla f(x_k) \rangle \leq 0 \\ \min \left( 1, \frac{\|\nabla f(x_k)\|^3}{\Delta_k \langle \nabla f(x_k), \nabla^2 f(x_k) \nabla f(x_k) \rangle} \right) & \text{if } \langle \nabla f(x_k), H_k \nabla f(x_k) \rangle > 0 \end{cases}.$$


---

**Exercise 25. (Trust region auxiliary problem and dogleg strategy)**

One method to solve the trust-region auxiliary problem (4.7) is the **dogleg strategy**. Hereby, in each iteration step, the following optimization problem is solved instead of (4.7):

$$(\star\star\star) \quad \min_{0 \leq t \leq 2} m_k(x_k(t)) \text{ s.t. } \|x_k - x_k(t)\| \leq \Delta_k$$

with the piecewise linear path

$$x_k(t) = \begin{cases} x_k + t(x_k^{\text{CP}} - x_k) & \text{for } 0 \leq t \leq 1 \\ x_k^{\text{CP}} + (t-1)(x_k^{\text{N}} - x_k^{\text{CP}}) & \text{for } 1 \leq t \leq 2 \end{cases}$$

where  $x_k^{\text{N}} = x_k - H_k^{-1} \nabla f(x_k)$  is the Newton step and  $x_k^{\text{CP}}$  denotes the normalized Cauchy point

$$x_k^{\text{CP}} = x_k - \frac{\|\nabla f(x_k)\|^2}{\langle \nabla f(x_k), H \nabla f(x_k) \rangle} \nabla f(x_k).$$

Hereby, we assume that the approximation of the Hessian matrix  $H_k$  is positive definite which implies that  $\langle x_k^{\text{N}} - x_k^{\text{CP}}, x_k^{\text{CP}} - x_k \rangle > 0$  holds; this fact can be used in the following without proof.

1. Show that the distance function  $\|x_k - x_k(t)\|$  increases strictly monotonically in  $t$  and that the function of model values  $m_k(x_k(t))$  decreases strictly monotonically in  $t$ .
2. Why are these two abilities helpful by solving the problem  $(\star\star\star)$ ?
3. Design a pseudo-code algorithm to solve  $(\star\star\star)$ .

**Exercise 26. (Program: Trust region method)**

Implement a simplified version of the trust region method according to the framework 4.10 and the routine 4.9 in the lecture notes.

1. For the framework, write a function `X = trustregion(fhandle,x0,r0,epsilon,N)` where `fhandle` is a function handle, `x0` ist the initial point, `r0` is the initial trust region radius, `epsilon` denotes the termination tolerance and `N` the maximal number of iteration steps. The function should return the matrix `X=[x0;x1;x2;...]` containing the whole iteration history.

To ease the implementation, modify the routine 4.9 in the following way:

- Change the termination condition to  $\|\nabla f(x)\| > \epsilon$  instead of  $\|\nabla f(x)\| > \tau_r \tau_0 + \tau_a$ .
- Don't approximate the Hessian matrix – instead, hand over the analytical second derivative.
- Instead of solving the minimization problem to obtain  $d_V^k$ , you should compute the Cauchy point  $x_a^{\text{CP}}$  to get a candidate for the next iteration  $x_V^k = x_a^{\text{CP}}$ .

To compute the Cauchy point  $x_a^{\text{CP}}$ , solve the subproblem described in Remark 4.13. Write a function `t = trustsubproblem(fhandle,x,r)` with the function handle `fhandle`, current point `x` and current trust region radius `r` that returns the correct step size `t=t_a` and compute  $x_a^{\text{CP}} = x_a - t_a \nabla f(x_a) = x_V^k$ .

To compute the next point  $x^{k+1}$  and the next radius  $\Delta_{k+1}$ , implement the update routine 4.9 by writing a function `[x,delta]= trustupdate(fhandle,x,xv,r)` with the function handle `fhandle`, the current point `x`, the current candidate `xv` and the current trust region radius `r`. Here again, you should solve the Cauchy point subproblem at the respective positions in the algorithm instead of the minimization problem.

For the parameters within the algorithm, use the following configuration:  $\mu_0 = 0.01$ ,  $\underline{\mu} = 0.1$ ,  $\bar{\mu} = 0.9$ ,  $\omega = 0.5$ ,  $\bar{\omega} = 2.0$  and  $C = 1.0$ .

2. Test your program with the Rosenbrock function and the following parameters: `epsilon=1.0e-3`, `N=1.0e4` and
  - a) `x0=[-5;-5]`, `r0=1.0`
  - b) `x0=[-5;-5]`, `r0=0.1`
  - c) `x0=[-5;-5]`, `r0=0.01`.

Describe your observations in the written report and generate a contour plot of the Rosenbrock function in which you display the computed iteration points.

## Index

Abstiegsrichtung .....	8	striktes lokales .....	3
Abstiegsverfahren .....	8	Modell .....	12
Algorithmus		Modell, polynomiales .....	12
Abstiegs- .....	9	Modul .....	6
Armijo- .....	10	Nebenbedingung .....	3
Newton-CG- .....	19	Newton-CG-Verfahren .....	19
Trust-Region- .....	21	Newton-Iterationsmatrix .....	20
Wolfe-Powell .....	13	Newton-Schritt .....	16
Armijo-Regel .....	10	Newton-Verfahren .....	16
Backtracking .....	12	inexaktes .....	18
Bedingung erster Ordnung .....	4	Niveaumenge .....	6
Bedingung zweiter Ordnung .....	4f	Optimalitätskriterien .....	4
BFGS-Formel .....	23	Optimierungsproblem	
Cauchy-Punkt .....	22	diskretes .....	3
CG-Verfahren .....	19	endlichdimensionales .....	3
Definitheit der Hessematrix .....	4	kombinatorisches .....	3
DFP-Formel .....	23	nicht-differenzierbares .....	3
dogleg-Technik .....	22	restringiertes .....	3
Fehlerfunktion .....	17	stetiges .....	3
Fehlerschranke		unrestringiertes .....	3
absolute .....	17	QN-Bedingung .....	23
relative .....	17	QN-Verfahren .....	23
globalisiertes Newton-Verfahren .....	20	Rang 1-Korrektur, symmetrische .....	23
Gradient .....	3	Rang 2-Korrekturen .....	23
gradientenähnliche Richtung .....	10, 15	Reduktion	
gradientenähnliches Verfahren .....	8	erwartete .....	20
Gradientenverfahren .....	8, 14	tatsächliche .....	20
Hessematrix .....	4	Restriktion	
Konvergenzordnung, superlineare .....	19	Ganzzahligkeits- .....	3
konvex		Gleichungs- .....	3
gleichmäßig .....	6	Ungleichungs- .....	3
strikt .....	6	Sattelpunkt .....	4
konvexe Funktion .....	6	Schrittweite .....	10
konvexe Menge .....	6	Schrittweitenstrategie .....	8
Maschinen-Epsilon .....	18	Schrittweitenstrategien .....	10
Maximalstelle		Sekantenbedingung .....	23
globale .....	3	Spektralnorm .....	4
lokale .....	3	stationärer Punkt .....	3
strikte globale .....	3	steilster Abstieg .....	8
strikte lokale .....	3	Trust-Region .....	20
Minimalstelle		Trust-Region-Hilfsproblem .....	20
globale .....	3	Trust-Region-Radius .....	20
lokale .....	3	Versuchslösung .....	20
strikte globale .....	3	Vorwärtsdifferenz .....	17
strikte lokale .....	3	Winkelbedingung .....	9
Minimum		Wolfe-Powell-Regel .....	12
globales .....	3	Wolfe-Powell-Regel, strenge .....	14
lokales .....	3	zentrale Differenz .....	18
striktes globales .....	3	Zielfunktion .....	3
		Zulässigkeitsbereich .....	3

## Literaturverzeichnis

- [1] Boyd, S. & Vandenberghe, L.: *Convex Optimization*. Cambridge University Press, 2004.
- [2] Geletu, A.: *Solving Optimization Problems using the Matlab Optimization Toolbox*. Vorlesungsskript Technische Universität Ilmenau, 2007.
- [3] Harbrecht, H.: *Nichtlineare Optimierung*. Vorlesungsskript Universität Stuttgart, 2011.
- [4] Hintermüller, M.: *Nonlinear Optimization*. Vorlesungsskript Universität Berlin, 2010.
- [5] Kelley, C. T.: *Iterative Methods for Optimization*. SIAM Frontiers in Applied Mathematics, 1999.
- [6] Nocedal, J. & Wright, S. J.: *Numerical Optimization*. Springer Series in Operations Research, 1999.
- [7] Ulbrich, M. & Ulbrich, S.: *Nichtlineare Optimierung*. Birkhäuser Verlag, 2012.
- [8] Volkwein, S.: *Numerische Verfahren der restringierten Optimierung*. Vorlesungsskript Universität Graz, 2009.
- [9] Werner, J.: *Optimierung*. Vorlesungsskript Universität Hamburg, 2008.