

Prof. Dr. Rembert Reemtsen

# **Modul 61415**

## **Nichtlineare Optimierung**

**Lehrveranstaltung**  
**Einführung in die nichtlineare Optimierung**

**LESEPROBE**

Fakultät für  
**Mathematik und  
Informatik**

---

Das Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung sowie der Übersetzung und des Nachdrucks bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Kein Teil des Werkes darf in irgendeiner Form (Druck, Fotokopie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung der FernUniversität reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

### 3 Ein Modellalgorithmus

Wir wollen uns nun Verfahren zur Lösung der unrestringierten Optimierungsaufgabe

$$(P) : \quad \text{Minimiere } f(x) \text{ über alle } x \in \mathbb{R}^n$$

zuwenden. Im Allgemeinen fordern wir dabei, dass  $f$  mindestens einmal stetig differenzierbar ist und nehmen wir an bzw. werden wir durch geeignete Voraussetzungen garantieren, dass  $(P)$  einen globalen Minimalpunkt besitzt.

In Abschnitt 3.1 werden wir zwei unterschiedliche Ansätze für die Konstruktion von Verfahren zur Lösung von  $(P)$  beschreiben und zwar einen, der neben einer Regel für die Bestimmung geeigneter *Abstiegsrichtungen* eine Regel für die *Schrittweiten* erfordert, und einen zweiten, der beides miteinander kombiniert. Anschließend werden wir uns der Klasse von Verfahren zuwenden, welche eine Schrittweitenregel beinhaltet, und werden wir zunächst in Abschnitt 3.2 für diese Klasse einen *Modellalgorithmus* aufstellen.

In Abschnitt 3.3 werden wir als Nächstes *Standardvoraussetzungen* diskutieren, die für den Nachweis der Konvergenz von Verfahren des betreffenden Typs benötigt werden. Zu diesem Zweck werden wir in Abschnitt 3.4 weitere Hilfsmittel bereit stellen. Damit werden wir dann in Abschnitt 3.5 allgemeine Bedingungen herleiten, denen eine Schrittweitenregel im Hinblick auf die Konvergenz eines Verfahrens genügen sollte. Schließlich werden wir in Abschnitt 3.6 unter entsprechenden Forderungen an die Schrittweiten allgemeine Konvergenzaussagen für den Modellalgorithmus beweisen, welche wir später auch für den Nachweis der Konvergenz spezieller Verfahren heranziehen werden.

#### 3.1 Schrittweitenbestimmung und Trust-Region

Ein ideales Verfahren zur Lösung des Problems  $(P)$ , d. h. ein Verfahren mit Abbruchschranke  $\nabla f(x) = 0$ , soll zwei Kriterien genügen. Es soll erstens in jeder Iteration einen Abstieg oder zumindest keinen Aufstieg hinsichtlich des Funktionswertes erzeugen, d. h., es soll  $f(x^{k+1}) < f(x^k)$  oder zumindest  $f(x^{k+1}) \leq f(x^k)$  für die Iterierten  $x^k$ ,  $k = 0, 1, 2, \dots$ , gelten. Man spricht daher auch von einem *Abstiegsverfahren*. Und zweitens soll jeder Häufungspunkt der vom Verfahren erzeugten unendlichen Iteriertenfolge  $\{x^k\}$  ein stationärer Punkt von  $f$  sein, sofern das Verfahren nicht bereits nach endlich vielen Iterationen mit einem stationären Punkt von  $f$  abbricht. (Den Grenzwert einer konvergenten Teilfolge einer Folge bezeichnet man als einen *Häufungspunkt*.)

Wenn der Startpunkt  $x^0$  für ein solches Verfahren nicht gerade ein lokaler Maximalpunkt von  $f$  ist, so dass das Verfahren mit diesem stationären Punkt terminiert (man sollte es dann mit einem anderen  $x^0$  nochmals beginnen), und wenn eine sinnvolle Schrittweitensteuerung verwendet wird, dann ist der gefundene stationäre Punkt  $x^*$  entweder ein lokaler Minimalpunkt oder ein Sattelpunkt

von  $f$ . Dass  $x^*$  ein Sattelpunkt ist, ist z. B. ausgeschlossen, wenn  $f$  eine konvexe Funktion ist, oder es kann möglicherweise durch die Überprüfung von Optimalitätsbedingungen zweiter Ordnung ausgeschlossen werden. Zumindest für größere Probleme ist aber eine Überprüfung solcher Bedingungen zweiter Ordnung oft numerisch zu teuer, so dass man sich zumeist damit zufrieden gibt zu wissen, dass  $x^*$  entweder ein lokaler Minimalpunkt oder ein Sattelpunkt von  $f$  ist.

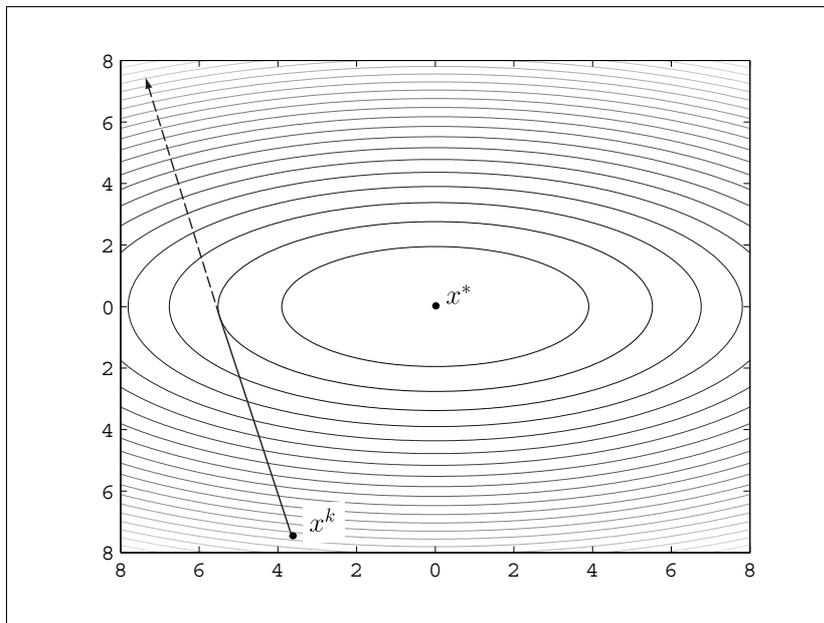


Abb. 3.1: Strahl  $z(t) := x^k + tp^k$ ,  $t \geq 0$

Leider ist es bis heute nur für relativ „einfache“ nichtlineare Funktionen möglich, globale Minimalpunkte, an denen man ja eigentlich interessiert ist, zu bestimmen. Es lässt sich daher nicht ausschließen, dass man bei der beschriebenen Vorgehensweise in einem lokalen Minimalpunkt hängenbleibt, der einen im Vergleich mit dem Minimalwert des Problems relativ großen Funktionswert besitzt. Aus diesem Grund ist es häufig sinnvoll, ein Verfahren für ein gegebenes Problem von verschiedenen, eventuell mit einem Zufallsgenerator erzeugten Startpunkten aus zu starten, um auf diesem Wege möglicherweise unterschiedliche stationäre Punkte der Zielfunktion zu erhalten.

Wir gehen nun davon aus, dass  $x^k$  kein kritischer Punkt von  $f$  ist, also  $\nabla f(x^k) \neq 0$  gilt. Erstes Ziel, wie beschrieben, ist es dann, einen neuen Punkt  $x^{k+1}$  zu bestimmen, für den der Funktionswert von  $f$  kleiner oder zumindest nicht größer als der Funktionswert bei  $x^k$  ist. In diesem Zusammenhang definieren wir:

**Definition 3.1** Ein Vektor  $p \in \mathbb{R}^n$  heißt Abstiegsrichtung für  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  in  $x \in \mathbb{R}^n$ , falls ein  $t_1 > 0$  existiert, so dass gilt:

$$f(x + tp) < f(x), \quad t \in (0, t_1]. \quad (3.1)$$

Schreiben wir  $x^{k+1} := x^k + t_k p^k$ , so geht es also darum, für einen nichtkritischen Punkt  $x^k$  eine Abstiegsrichtung  $p^k$  und eine geeignete Schrittweite  $t_k > 0$  zu bestimmen bzw., wenn wir  $\tilde{p}^k := t_k p^k$  setzen, eine Abstiegsrichtung  $\tilde{p}^k$  mit geeigneter Länge  $\|\tilde{p}^k\|$  zu finden. (Mit  $p^k$  ist offenbar auch  $\tilde{p}^k$  Abstiegsrichtung.) Die Schrittweite  $t_k$  bzw. die Länge  $\|\tilde{p}^k\|$  von  $\tilde{p}^k$  darf dabei nicht zu groß sein, da anderenfalls auch für eine Abstiegsrichtung die Funktionswerte normalerweise wieder ansteigen. Letzteres verdeutlicht Abb. 3.1, in welcher Höhenlinien der quadratischen Funktion

$$f(x_1, x_2) := x_1^2 + 4x_2^2$$

sowie ein Strahl  $z(t) := x^k + t p^k$ ,  $t \geq 0$ , für eine Abstiegsrichtung  $p^k$  dargestellt werden. (Hinter dem Punkt, in dem der Strahl eine Höhenlinie tangential berührt, gelangt man wieder zu weiter außen befindlichen Höhenlinien und damit zu größeren Funktionswerten.)

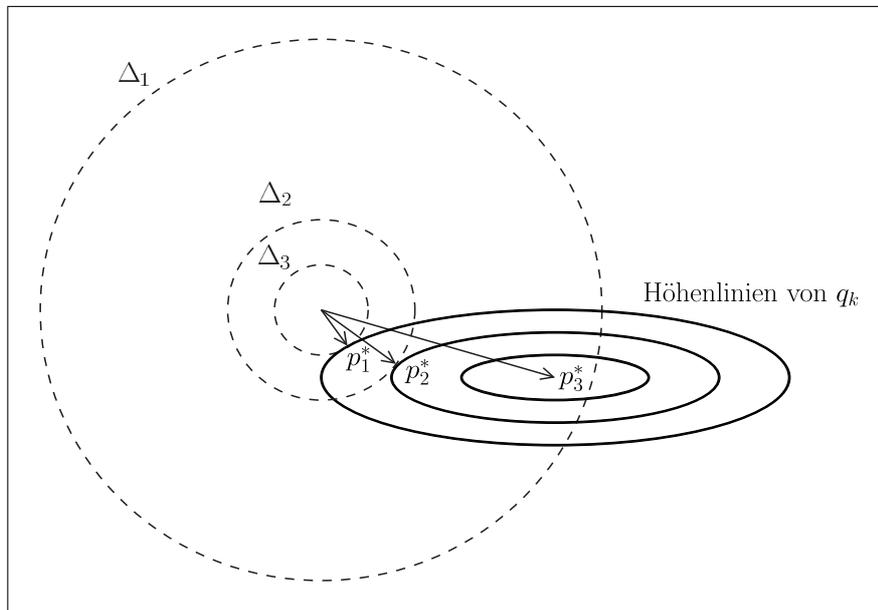


Abb. 3.2: Lösungen  $p_i^*$  des Trust-Region-Teilproblems für  $\Delta_i$

In diesem Zusammenhang gibt es nun zwei grundsätzlich verschiedene Vorgehensweisen. Bei den meisten bekannten Verfahren wird zunächst eine Abstiegsrichtung und anschließend eine geeignete Schrittweite bestimmt. Bei den *Trust-Region-Verfahren* dagegen kombiniert man die Richtungssuche und die Schrittweitenbestimmung, indem man  $f(x^k + p)$  z. B. durch eine quadratische Näherung

$q_k(p)$  ersetzt und für eine passend gewählte Konstante  $\Delta_k$  das folgende Teilproblem mit einer Nebenbedingung bezüglich  $p \in \mathbb{R}^n$  löst:

$$\begin{aligned} &\text{Minimiere } q_k(p) \\ &\text{u. d. N. } \quad \|p\| \leq \Delta_k. \end{aligned} \tag{3.2}$$

Die Konstante  $\Delta_k$  in diesem Problem definiert ein „trust region“, einen Vertrauensbereich, in dem die gesuchte Richtung liegen darf. Ob eine Lösung  $\tilde{p}^k$  dieses Problems akzeptabel ist, muss anschließend entschieden werden. (Nach dem Satz 2.39 von Weierstraß besitzt das Problem (3.2) für stetiges  $q_k$  eine Lösung.) Falls eine Lösung nicht brauchbar ist, muss  $\Delta_k$  verkleinert und muss das Problem (3.2) erneut gelöst werden. Die Funktion  $f(x^k + p)$  selbst kann man in einem solchen Teilproblem übrigens nicht als Zielfunktion verwenden, da dieses Problem dann von demselben Schwierigkeitsgrad oder wegen der zusätzlichen Restriktion sogar von einem höheren Schwierigkeitsgrad wie das eigentlich zu lösende Problem ( $P$ ) wäre.

Die meisten Verfahren, die wir vorstellen werden, sind vom ersten Typ und erfordern eine Schrittweitenbestimmung. Jedoch werden wir in Kapitel 9 auch Trust-Region-Verfahren diskutieren. Die Kritiker von Trust-Region-Verfahren bemängeln vor allem, dass die Richtung, in welche eine Lösung  $\tilde{p}^k$  des Teilproblems zeigt, zumeist stark von der Wahl von  $\Delta_k$  abhängt (siehe Abb. 3.2) und dass man die Richtung für einen Abstiegschritt nicht von einer für den Schritt vorgegebenen Länge abhängig machen sollte.

Letzteres ist so, als ob man sich im Gebirge, wenn man dort einen Weg ins Tal sucht, als Erstes eine Länge für den nächsten Schritt vorgibt und man erst anschließend eine geeignete Richtung wählt. Auf der anderen Seite kennt man bei den Verfahren, welche eine Schrittweitenbestimmung erfordern, zwar die Richtung, in die man einen Abstieg erzielen möchte (die Verfahren unterscheiden sich vor allem durch die Wahl der Richtungen), die Schrittweitenbestimmung kann aber sehr viele Funktionsauswertungen erfordern, und es besteht die Gefahr, dass Schrittweiten gewählt werden, die für den weiteren Verlauf des Verfahrens z. B. zu klein und damit ungünstig sind. In der Praxis haben sich aber beide Typen von Verfahren bewährt, zum Teil in unterschiedlichen Zusammenhängen.

### 3.2 Ein Modellalgorithmus

Wir wollen nun zunächst nur Algorithmen mit einer Schrittweisenstrategie betrachten und in diesem Abschnitt ein Grundmodell für derartige Algorithmen diskutieren. Zunächst stellen wir eine einfache Bedingung bereit, mit deren Hilfe wir leicht Abstiegsrichtungen in einem vorgegebenen Punkt angeben können.

**Lemma 3.2** *Es sei  $f \in C^1(\mathbb{R}^n)$ . Gilt für  $p$*

$$\nabla f(x)^T p < 0,$$

*so ist  $p$  Abstiegsrichtung für  $f$  in  $x$ .*

**Beweis.** Die Definition der Richtungsableitung von  $f$  bei  $x$  in Richtung  $p$  liefert

$$\lim_{t \rightarrow 0+} \frac{f(x + tp) - f(x)}{t} = \nabla f(x)^T p < 0. \quad (3.3)$$

Folglich gilt für ein  $t_1 > 0$

$$\frac{f(x + tp) - f(x)}{t} < 0, \quad t \in (0, t_1].$$

Somit ist (3.1) richtig. ■

**Beispiel 3.3** *Es sei  $f \in C^1(\mathbb{R}^n)$ , und es sei  $H \in \mathbb{R}^{n \times n}$  eine positiv definite Matrix. Für ein  $x$  mit  $\nabla f(x) \neq 0$  ist dann der Vektor  $p$  mit*

$$p := -H \nabla f(x)$$

*Abstiegsrichtung für  $f$  in  $x$ , da gilt:*

$$\nabla f(x)^T p = -\nabla f(x)^T H \nabla f(x) < 0.$$

*Insbesondere erhält man für  $H := I$  die Abstiegsrichtung*

$$p := -\nabla f(x). \quad (3.4)$$

Zu der Abstiegsrichtung in (3.4) bemerken wir:

**Bemerkung 3.4** *Nach (3.3) hat man für alle genügend kleinen  $t \geq 0$*

$$f(x + tp) - f(x) = t \nabla f(x)^T p + \varepsilon(t)$$

*mit  $\varepsilon(t)/t \rightarrow 0$  für  $t \rightarrow 0+$ . Im Hinblick auf einen möglichst großen Abstieg für  $f$  in  $x$ , also einen möglichst kleinen Wert  $f(x + tp) - f(x)$ , macht es Sinn,*

nach einer auf 1 normierten Richtung  $p$  zu fragen, welche im Fall  $\nabla f(x) \neq 0$  das Problem

$$\begin{aligned} & \text{Minimiere } \nabla f(x)^T p \\ & \text{u. d. N. } \quad \|p\| = 1 \end{aligned} \tag{3.5}$$

löst. Die eindeutige Lösung dieses Problems ist

$$p^* := -\nabla f(x) / \|\nabla f(x)\|.$$

Denn mit der Cauchy-Schwarz-Ungleichung kann man den Zielfunktionswert von Problem (3.5) für alle zulässigen  $p$  durch

$$\nabla f(x)^T p \geq -\|\nabla f(x)\| \|p\| = -\|\nabla f(x)\| \tag{3.6}$$

nach unten abschätzen, wobei die untere Schranke offenbar gerade für  $p^*$  angenommen wird und damit den Minimalwert des Problems definiert. Die Eindeutigkeit von  $p^*$  folgt aus der Tatsache, dass man  $\nabla f(x) \neq 0$  und  $p \neq 0$  hat und daher Gleichheit in der Cauchy-Schwarz-Ungleichung in (3.6) genau dann vorliegt, wenn  $p = \lambda \nabla f(x)$  für ein  $\lambda \neq 0$  ist. Mit (3.6) ergibt sich für  $\lambda$  als einzige mögliche Wahl  $\lambda := -1/\|\nabla f(x)\|$ .

Die Richtung  $p := -\nabla f(x)$  nennt man daher auch die Richtung des steilsten Abstiegs für  $f$  in  $x$ . Man kann sie lokal als „beste“ Abstiegsrichtung ansehen. Global gesehen muss sie es jedoch nicht sein, wie auch im Gebirge die Richtung, die vom Standpunkt aus - vielleicht nur für ein kleines Stück - den steilsten Abstieg liefert, global gesehen nicht notwendig die beste Richtung für einen Abstieg ins Tal sein muss.

**Aufgabe 3.5** Sei  $A \in \mathbb{R}^{n \times n}$  eine symmetrische positiv definite Matrix und

$$\langle x, y \rangle_A := x^T A y, \quad x, y \in \mathbb{R}^n. \tag{3.7}$$

Wie sich leicht überprüfen lässt, wird durch  $\langle \cdot, \cdot \rangle_A$  ein Skalarprodukt und, wie aus der Analysis 2 bekannt ist, durch  $\|x\|_A := \langle x, x \rangle_A^{1/2}$  eine Norm, eine sog. elliptische Norm, auf dem  $\mathbb{R}^n$  definiert. Es sei weiter  $f \in C^1(\mathbb{R}^n)$  und  $x$  ein Punkt mit  $\nabla f(x) \neq 0$ . Man zeige nun, dass

$$p^* := -\frac{A^{-1} \nabla f(x)}{\|A^{-1} \nabla f(x)\|_A}$$

die eindeutige Lösung der Optimierungsaufgabe

$$\begin{aligned} & \text{Minimiere } \nabla f(x)^T p \\ & \text{u. d. N. } \quad \|p\|_A = 1, \end{aligned} \tag{3.8}$$

d. h. dass  $-A^{-1} \nabla f(x)$  die Richtung des steilsten Abstiegs in  $x$  bezüglich  $\|\cdot\|_A$  ist und dass für den optimalen Zielfunktionswert von (3.8)  $\nabla f(x)^T p^* < 0$  gilt. (Dieses Ergebnis ist für Kapitel 8 von Interesse.)

Wir wollen nun den folgenden Algorithmus betrachten, der ein Modell für Algorithmen mit Schrittweitenbestimmung zur Lösung von (P) darstellt.

**Modellalgorithmus 3.6**

- (0) (*Initialisierung*)  
Wähle  $x^0 \in \mathbb{R}^n$ . Setze  $k := 0$ .
- (1) (*Abbruchkriterium*)  
Falls  $\nabla f(x^k) = 0$  ist, stop!
- (2) (*Bestimmung einer Abstiegsrichtung*)  
Bestimme ein  $p^k \in \mathbb{R}^n$  mit  $\nabla f(x^k)^T p^k < 0$ .
- (3) (*Bestimmung einer Schrittweite*)  
Bestimme ein  $t_k > 0$  mit  $f(x^k + t_k p^k) < f(x^k)$ .
- (4) (*Bestimmung der nächsten Iterierten*)  
Setze  $x^{k+1} := x^k + t_k p^k$ ,  $k := k + 1$  und gehe nach (1).

Im Hinblick auf Konvergenzuntersuchungen wird im Modellalgorithmus das ideale Abbruchkriterium „ $\nabla f(x^k) = 0$ “ verwendet. Für die Praxis ist dieses durch ein realistisches *Abbruchkriterium* zu ersetzen. In diesem Zusammenhang wird zumeist das Kriterium

$$\|\nabla f(x^k)\| \leq \varepsilon$$

für ein vorgegebenes  $\varepsilon > 0$  genutzt. Man bedenke jedoch, dass es Funktionen wie die Funktion  $f(x) := 10^{-6}x^2$  gibt, für welche  $\nabla f(x^k) \approx 0$  auch für Punkte  $x^k$  gilt, die noch weit von einer kritischen Lösung  $x^*$  des Problems entfernt sind.

Deshalb kann es beispielsweise sinnvoll sein, für den Abbruch eines Verfahrens (zusätzlich) zu fordern, dass die Bedingung

$$|f(x^k) - f(x^{k-1})| / |f(x^{k-1})| \leq \delta_1$$

über einige Iterationen hinweg für ein genügend kleines  $\delta_1 > 0$  erfüllt ist, da dann keine signifikante Verminderung des Funktionswertes von  $f$  mehr zu erwarten ist.<sup>1</sup> Letzteres Kriterium zielt darauf ab, dass die Größe des beim Abbruch eines Verfahrens erreichten Funktionswertes  $f(x^k)$  in der Praxis häufig viel interessanter ist als die Größe der Abweichung  $\|x^k - x^*\|$ . Ist man an der Genauigkeit von  $x^k$  selbst interessiert, so kann man (zusätzlich) auch das Kriterium

$$\|x^k - x^{k-1}\| / \|x^{k-1}\| \leq \delta_2$$

für ein  $\delta_2 > 0$  heranziehen.

Die zentralen Schritte (2) und (3) des Algorithmus sind prinzipiell ausführbar (was nicht heißt, dass ein entsprechend spezifizierter Algorithmus auch konvergieren muss). Denn in Schritt (2) kann man offenbar z. B. die Richtung steilsten Abstiegs  $p^k := -\nabla f(x^k)$  als Abstiegsrichtung wählen. (Es ist dort ja  $\nabla f(x^k) \neq 0$ .)

<sup>1</sup>Mit einer *Iteration* des Modellalgorithmus 3.6 - und analog jedes anderen Verfahrens - ist das einmalige Durchlaufen der Schritte (1) bis (4) und mit der  $\ell$ -ten Iteration insbesondere das Durchlaufen dieser Schritte für  $k := \ell$  gemeint.

Bei dieser Richtungswahl spricht man von dem *Gradientenverfahren*, das wir in Abschnitt 5 genauer untersuchen wollen. Ferner existiert gemäß der Definition 3.1 einer Abstiegsrichtung und gemäß Lemma 3.2 immer eine Schrittweite  $t_k > 0$ , wie sie in Schritt (3) des Modellalgorithmus zu bestimmen ist.

Da der Wert  $\varphi_k(t) := f(x^k + tp^k)$  in der  $k$ -ten Iteration eines Verfahrens vom Typ des Modellalgorithmus möglichst klein werden sollte, liegt es zumindest aus theoretischer Sicht zunächst nahe, als Schrittweite ein  $t_k$  zu wählen, für welches

$$f(x^k + t_k p^k) = \min_{t \in [0, \infty)} f(x^k + tp^k) \quad (3.9)$$

gilt, d. h., für welches die eindimensionale Funktion  $\varphi_k$  auf  $[0, \infty)$  ein globales Minimum annimmt. Dass eine solche *Minimumschrittweite* existiert, werden wir unter relativ schwachen Voraussetzungen in Abschnitt 4.1 zeigen. Aus numerischer Sicht ist aber die Bestimmung des globalen Minimums einer nichtkonvexen Funktion eine Aufgabe, die im Allgemeinen nicht oder bestenfalls nur näherungsweise gelöst werden kann. Es wird deshalb erforderlich sein, auch noch andere Regeln zur Bestimmung von Schrittweiten zu diskutieren.

Der Modellalgorithmus bricht also entweder nach endlich vielen Schritten mit einer kritischen Lösung von  $(P)$  ab, oder er erzeugt eine unendliche Folge  $\{x^k\}$  mit

$$f(x^{k+1}) < f(x^k) < f(x^0), \quad k \in \mathbb{N}. \quad (3.10)$$

Für jedes spezielle Verfahren vom Typ des Modellalgorithmus ist demnach zu zeigen, dass eine solche durch das Verfahren erzeugte Iteriertenfolge für jeden geeigneten Startpunkt  $x^0$  mindestens einen Häufungspunkt besitzt und dass zumindest ein Häufungspunkt oder, besser noch, dass jeder Häufungspunkt dieser Folge eine kritische Lösung von  $(P)$  ist. Bevor wir auf spezielle Verfahren, d. h. spezielle Richtungs- und Schrittweitenstrategien eingehen werden, wollen wir eine Reihe von allgemeinen Aussagen zur Konvergenz des Modellalgorithmus selbst herleiten. Diese Aussagen werden zum einen Richtungs- und Schrittweitenwahlen motivieren und zum anderen für den Nachweis der Konvergenz im Spezialfall nützlich sein.

### 3.3 Standardvoraussetzungen

Zunächst einmal stellen wir Standardvoraussetzungen bereit, auf die wir uns durchgängig beziehen werden:

(V1)  $f \in C^1(\mathbb{R}^n)$ .

(V2) Für ein  $x^0 \in \mathbb{R}^n$  ist die Niveaumenge

$$N_0 := N(x^0) = \{x \in \mathbb{R}^n \mid f(x) \leq f(x^0)\} \quad (3.11)$$

kompakt, wobei  $x^0$  im Zusammenhang mit einem Verfahren der Startpunkt des Verfahrens ist.

(V3) Der Gradient  $\nabla f$  ist auf  $N_0$  Lipschitz-stetig, d. h., es existiert eine Konstante  $\gamma > 0$  derart, dass gilt:

$$\|\nabla f(x) - \nabla f(y)\| \leq \gamma \|x - y\|, \quad x, y \in N_0. \quad (3.12)$$

Gelegentlich, insbesondere für globale Konvergenzaussagen, werden wir zusätzlich noch folgende Bedingung voraussetzen (vgl. dazu Satz 2.29):

(V4) Die Niveaumenge  $N_0$  ist konvex, und  $f$  ist gleichmäßig konvex auf  $N_0$ , d. h., es existiert eine Konstante  $\beta > 0$  mit

$$\frac{\beta}{2} t(1-t) \|x - y\|^2 + f(tx + (1-t)y) \leq tf(x) + (1-t)f(y), \quad x, y \in N_0, \quad t \in [0, 1].$$

Diese Voraussetzungen haben einige Implikationen, welche in der nachstehenden Bemerkung erläutert sind.

**Bemerkung 3.7** (i) Die Voraussetzung (V2) sichert nach Satz 2.42, dass das Problem (P) eine Lösung besitzt. Zum anderen garantiert sie nach dem aus der Analysis 2 bekannten Satz von Bolzano-Weierstraß, dass jede Folge  $\{x^k\}$  mit  $x^k \in N_0$  bzw. mit  $f(x^k) \leq f(x^0)$  einen Häufungspunkt in  $N_0$  hat. Wegen (3.10) erzeugen Abstiegsverfahren solche Folgen.

(ii) Die Voraussetzung (V3) ist erfüllt, wenn (V2) erfüllt und  $f \in C^2(K)$  für eine kompakte, konvexe Menge  $K \supseteq N_0$  ist. (Letzteres ist wegen (V2) offenbar gegeben, wenn  $f \in C^2(\mathbb{R}^n)$  ist.) Denn dann existiert

$$\gamma := \max_{z \in K} \|\nabla^2 f(z)\| \quad (3.13)$$

und ist  $x + t(y - x) \in K$  für alle  $x, y \in N_0$  und  $t \in [0, 1]$ . Für

$$u(t) := \nabla f(x + t(y - x)), \quad t \in [0, 1],$$

folgt dann für alle  $x, y \in N_0$  mit einem  $\vartheta \in (0, 1)$  nach dem Mittelwertsatz  $u(1) - u(0) = u'(\vartheta)$  bzw.

$$\nabla f(y) - \nabla f(x) = \nabla^2 f(x + \vartheta(y - x))(y - x).$$

Letzteres impliziert die Ungleichung (3.12) mit  $\gamma$  aus (3.13).

(iii) Die Voraussetzungen (V1) und (V4) zusammen implizieren, dass die gemäß (V4) konvexe Niveaumenge  $N_0$  kompakt ist, also (V2) gilt, und dass das konvexe Optimierungsproblem „Minimiere  $f(x)$  über alle  $x \in N_0$ “ genau eine Lösung  $x^* \in N_0$  besitzt (Satz 2.43 für  $Z := N_0$ ). Demnach hat das Problem (P) unter (V1) und (V4) genau eine globale Lösung  $x^*$ , und diese ist die einzige kritische Lösung von (P), die in  $N_0$  liegt (wende Satz 2.29 (i) auf  $K := N_0$  und einen kritischen Punkt  $y \in N_0$  an). Wenn  $f$  überdies auf dem ganzen Raum  $\mathbb{R}^n$  gleichmäßig konvex ist, ist  $x^*$  die einzige kritische Lösung von (P) (Korollar 2.50).

Da die Voraussetzung (V4) in der Praxis für ein  $f$  normalerweise nur dadurch verifiziert werden kann, dass man zeigt, dass  $f$  auf  $\mathbb{R}^n$  gleichmäßig konvex ist, sprechen wir der Einfachheit halber auch im Fall, dass nur (V4) vorausgesetzt wird, von der eindeutigen Lösung  $x^*$  von (P), obwohl dann (P) auch noch lokale Lösungen außerhalb von (P) besitzen kann.

Als Beispiel betrachten wir quadratische Funktionen.

**Beispiel 3.8** Es sei  $Q \in \mathbb{R}^{n \times n}$  symmetrisch und positiv semidefinit und  $f$  die somit konvexe quadratische Funktion

$$f(x) := \frac{1}{2}x^T Q x + c^T x + \alpha, \quad x \in \mathbb{R}^n.$$

Unter Verwendung der oberen Schranke aus (2.22) und der Symmetrie von  $Q$  erhält man dann, wenn  $\lambda_i(A)$  die Eigenwerte einer Matrix  $A \in \mathbb{R}^{n \times n}$  bezeichnet:

$$\begin{aligned} \|\nabla f(x) - \nabla f(y)\|^2 &= \|Q(x - y)\|^2 = [Q(x - y)]^T Q(x - y) \\ &= (x - y)^T Q^T Q(x - y) \leq \lambda_{\max}(Q^T Q) \|x - y\|^2 \\ &= \lambda_{\max}(Q^2) \|x - y\|^2 = [\lambda_{\max}(Q)]^2 \|x - y\|^2. \end{aligned}$$

Wenn  $Q$  positiv definit ist, genügt  $f$  also neben (V1) auch der Voraussetzung (V3) mit

$$\gamma := \lambda_{\min}(Q) > 0, \quad (3.14)$$

wobei dieses  $\gamma$  offenbar die kleinst mögliche Konstante für  $f$  ist, für welche (V3) gilt. Ist  $Q$  positiv definit, so ist zusätzlich (V4) erfüllt (Lemma 2.34) und damit auch (V2) (Bemerkung 3.7 (iii)). Es kann dann insbesondere

$$\beta := \lambda_{\min}(Q) > 0 \quad (3.15)$$

gewählt werden. Für positiv definites  $Q$  gilt somit gemäß (2.21) hinsichtlich der Spektralnorm

$$\text{cond}(Q) = \frac{\lambda_{\max}(Q)}{\lambda_{\min}(Q)} = \frac{\gamma}{\beta}. \quad (3.16)$$

Dabei ist  $\beta$  die gleichmäßige Konvexitätskonstante von  $f$  und  $\gamma$  die Lipschitz-Konstante von  $\nabla f$ .

### 3.4 Hilfsmittel

In diesem Abschnitt stellen wir einige Hilfsmittel bereit, die wir zur Untersuchung des Modellalgorithmus benötigen werden.

In jeder Iteration des Modellalgorithmus 3.6 sind, ausgehend von einem Punkt  $x$  mit  $\nabla f(x) \neq 0$ , eine Abstiegsrichtung  $p$ , eine Schrittweite  $t > 0$  und damit ein neuer Punkt  $x+tp$  zu bestimmen, so dass ein Abstieg bezüglich des Funktionswertes von  $f$  erzielt wird. (Der Einfachheit halber lassen wir hier den Iterationsindex weg.) Das folgende Lemma gibt für solche Vektoren  $x$  und  $p$  ein Intervall von Schrittweiten an, in dem die Funktion

$$\psi(t) := f(x) - f(x + tp) \quad (3.17)$$

positive Werte annimmt und demnach eine Reduktion des Zielfunktionswertes von Problem  $(P)$  möglich ist.

**Lemma 3.9** *Es seien (V1) und (V2) erfüllt und*

$$x \in N_0, \quad p \in \mathbb{R}^n \quad \text{mit} \quad \nabla f(x)^T p < 0$$

*gegeben. Dann besitzt die in (3.17) definierte Funktion  $\psi \in C^1(\mathbb{R})$  eine kleinste positive Nullstelle  $\hat{t} := \hat{t}(x, p) > 0$ , und es gilt mit einem  $\kappa \geq \hat{t}$*

$$\psi(t) > 0, \quad t \in (0, \hat{t}); \quad x + tp \in N_0, \quad t \in [0, \hat{t}]; \quad x + tp \notin N_0, \quad \psi(t) < 0, \quad t \geq \kappa.$$

**Beweis.** Es ist  $\psi(0) = 0$ , und es folgt

$$\psi'(t) = -\nabla f(x + tp)^T p, \quad \psi'(0) = -\nabla f(x)^T p > 0. \quad (3.18)$$

Somit existiert ein  $\delta > 0$ , so dass  $\psi(t) > 0$  für alle  $t \in (0, \delta)$  gilt. Da  $N_0$  beschränkt und  $p \neq 0$  ist, hat man ferner  $x + tp \notin N_0$  für alle genügend großen  $t > 0$  und hat man somit für diese  $t$

$$f(x + tp) > f(x^0) \geq f(x).$$

Also existiert ein  $\kappa > 0$ , so dass  $x + tp \notin N_0$  sowie  $\psi(t) < 0$  für alle  $t \geq \kappa$  gilt. Damit ist die Menge

$$N := \{t \in [\delta, \kappa] \mid \psi(t) = 0\}$$

der Nullstellen von  $\psi$  nichtleer. Sie ist ferner kompakt, so dass nach dem Satz 2.39 von Weierstraß ein  $\hat{t} := \min_{t \in N} t$ , d. h. eine kleinste positive Nullstelle existiert. Mit den anfangs genannten Eigenschaften von  $\psi$  schließt man daher, dass  $\psi(t) > 0$  für  $t \in (0, \hat{t})$  sowie  $\psi(t) \geq 0$  für  $t \in [0, \hat{t}]$  gilt und damit  $\kappa \geq \hat{t}$  ist. Weil  $x$  in  $N_0$  liegt, impliziert Letzteres

$$f(x + tp) \leq f(x) \leq f(x^0)$$

und somit  $x + tp \in N_0$  für alle  $t \in [0, \hat{t}]$ . ■

Unter den Voraussetzungen (V1) - (V4) können wir einige sehr nützliche Abschätzungen für gleichmäßig konvexe Funktionen beweisen. Es sei daran erinnert, dass das Problem (P) unter diesen Voraussetzungen eine eindeutige Lösung  $x^*$  besitzt (Bemerkung 3.7 (iii)). Wegen  $f(x^*) \leq f(x^0)$  liegt diese in  $N_0$ .

**Lemma 3.10** *Es seien (V1)–(V4) erfüllt, und es sei  $x^* \in N_0$  die Lösung von (P). Ist weiter  $\beta$  die gleichmäßige Konvexitätskonstante aus (V4) und  $\gamma$  die Lipschitz-Konstante aus (V3), so gilt für alle  $x, y \in N_0$ :*

$$(i) \quad \frac{\beta}{2} \|y - x\|^2 + \nabla f(x)^T(y - x) \leq f(y) - f(x) \leq \frac{\gamma}{2} \|y - x\|^2 + \nabla f(x)^T(y - x),$$

$$(ii) \quad \beta \|y - x\|^2 \leq (\nabla f(y) - \nabla f(x))^T(y - x),$$

$$(iii) \quad \frac{\beta}{2} \|x - x^*\|^2 \leq f(x) - f(x^*),$$

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

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

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

**Beweis.** Seien  $x, y$  und  $x^*$  wie vorgegeben. Man beachte, dass  $N_0$  gemäß (V4) eine konvexe Menge ist und somit  $x + s(y - x) \in N_0$  für alle  $s \in [0, 1]$  gilt.

(i) Die linke Ungleichung von (i) wird durch (V4) impliziert (vgl. Satz 2.29 (iii)). Sei nun  $\phi \in C^1[0, 1]$  durch  $\phi(t) := f(x + t(y - x))$  definiert. Dann ist

$$\phi'(s) = \nabla f(x + s(y - x))^T(y - x)$$

und

$$f(y) - f(x) = \phi(1) - \phi(0) = \phi'(0) + \int_0^1 [\phi'(s) - \phi'(0)] ds. \quad (3.19)$$

Unter Verwendung der Cauchy-Schwarz-Ungleichung und von (V3) gilt dabei

$$\begin{aligned} \int_0^1 [\phi'(s) - \phi'(0)] ds &= \int_0^1 [\nabla f(x + s(y-x)) - \nabla f(x)]^T (y-x) ds \\ &\leq \int_0^1 \|\nabla f(x + s(y-x)) - \nabla f(x)\| \|y-x\| ds \\ &\leq \int_0^1 \gamma s \|y-x\|^2 ds = \frac{\gamma}{2} \|y-x\|^2, \end{aligned} \quad (3.20)$$

Damit ergibt sich die rechte Ungleichung von (i) aus (3.19).

(ii) Die linke Ungleichung in (i) liefert

$$\frac{\beta}{2} \|y-x\|^2 + \nabla f(x)^T (y-x) \leq f(y) - f(x)$$

und, mit vertauschten Rollen von  $x$  und  $y$ ,

$$\frac{\beta}{2} \|y-x\|^2 + \nabla f(y)^T (x-y) \leq f(x) - f(y).$$

Addition von beiden Ungleichungen ergibt die gewünschte Beziehung.

(iii), (iv) Für  $y := x$  und  $x := x^*$  folgt wegen  $\nabla f(x^*) = 0$  aus der linken Ungleichung von (i) die Abschätzung in (iii) und aus der Ungleichung in (ii) die Abschätzung in (iv) mit

$$\beta \|x-x^*\|^2 \leq \nabla f(x)^T (x-x^*) \leq \|\nabla f(x)\| \|x-x^*\|.$$

(v) Die Funktion

$$g_x(h) := \frac{\beta}{2} h^T h + \nabla f(x)^T h, \quad h \in \mathbb{R}^n,$$

ist nach Lemma 2.34 gleichmäßig konvex und nimmt folglich ihr Minimum in genau einem Punkt  $h^*$  an, für den gilt:

$$\nabla g_x(h^*) = 0 \quad \Leftrightarrow \quad \beta h^* + \nabla f(x) = 0 \quad \Leftrightarrow \quad h^* = -\frac{1}{\beta} \nabla f(x).$$

Wegen der Kompaktheit von  $N_0$  und unter Verwendung von (i) folgt damit

$$\begin{aligned} g_x(h^*) &= -\frac{1}{2\beta} \|\nabla f(x)\|^2 = \min_{h \in \mathbb{R}^n} g_x(h) \\ &\leq \min_{y \in N_0} g_x(y-x) = \min_{y \in N_0} \left\{ \frac{\beta}{2} \|y-x\|^2 + \nabla f(x)^T (y-x) \right\} \\ &\leq \min_{y \in N_0} \{f(y) - f(x)\} = f(x^*) - f(x). \end{aligned}$$

(vi) Es sei  $\hat{y} := x - \frac{1}{\gamma} \nabla f(x)$ . Wir zeigen zunächst, dass  $\hat{y} \in N_0$  ist. Da dies für  $\nabla f(x) = 0$  trivial ist, können wir dabei  $\nabla f(x) \neq 0$  annehmen.

Nach Lemma 3.9, angewandt auf  $p := -\nabla f(x)$ , besitzt die Funktion

$$\chi(t) := f(x) - f(x - t\nabla f(x)), \quad t \geq 0,$$

eine erste positive Nullstelle  $\hat{t}$  und ist  $[x - t\nabla f(x)] \in N_0$  für alle  $t \in [0, \hat{t}]$ . Ferner folgt mit der rechten Ungleichung von (i)

$$0 = f(x - \hat{t}\nabla f(x)) - f(x) \leq \frac{\gamma}{2} \hat{t}^2 \|\nabla f(x)\|^2 - \hat{t} \|\nabla f(x)\|^2$$

und daraus  $\hat{t} \geq 2/\gamma \geq 1/\gamma$ . Demnach ist  $\hat{y} \in N_0$ . Unter Ausnutzung der Optimalität von  $x^*$  und der von Aussage (i) erhalten wir schließlich

$$\begin{aligned} f(x^*) - f(x) &\leq f(\hat{y}) - f(x) \\ &\leq \frac{1}{2\gamma} \|\nabla f(x)\|^2 - \frac{1}{\gamma} \|\nabla f(x)\|^2 = -\frac{1}{2\gamma} \|\nabla f(x)\|^2. \quad \blacksquare \end{aligned}$$

**Bemerkung 3.11** Für  $f(x) := \frac{1}{2}x^T x$  können wir gemäß Beispiel 3.8

$$\beta = \lambda_{\min}(I) = 1, \quad \gamma = \lambda_{\max}(I) = 1$$

wählen. Wie man leicht nachprüft, gelten für dieses  $f$  alle Ungleichungen in Lemma 3.10 mit Gleichheit. Sie können also nicht mehr verschärft werden.

Man beachte, dass die gleichmäßige Konvexität von  $f$  unter den Voraussetzungen von Lemma 3.10 also insbesondere eine obere Abschätzung des Fehlers  $\|x - x^*\|$  durch  $f(x) - f(x^*)$  liefert, was für den Nachweis der Konvergenz von Verfahren genutzt werden wird. Eine obere Abschätzung umgekehrt von  $f(x) - f(x^*)$  durch  $\|x - x^*\|$  hat man ja für jede Lipschitz-stetige Funktion, also für jede Funktion  $f \in C^1(\mathbb{R}^n)$ .

### 3.5 Bedingungen an die Schrittweiten

Als die ersten Verfahren zur Lösung unrestringierter Optimierungsprobleme entwickelt wurden, wurde immer wieder Konvergenz von Verfahren bewiesen, die sich „nur“ durch die Schrittweitenregel unterschieden. Es war daher irgendwann sinnvoll zu fragen, welche Eigenschaften eine Schrittweitenregel besitzen sollte, damit Konvergenz für ein Verfahren nachgewiesen werden kann. Die Herleitung zweier Bedingungen in diesem Zusammenhang, welche zu den Definitionen einer *effizienten* und einer *semieffizienten Schrittweitenregel* führen werden, sind Thema dieses Abschnitts.

Lemma 3.9 gibt für ein  $x$  und eine Abstiegsrichtung  $p$  ein offenes Intervall  $(0, \hat{t})$  an, in dem die Funktion

$$\psi(t) := f(x) - f(x + tp) \quad (3.21)$$

positiv ist, also der Zielfunktionswert von Problem  $(P)$  verkleinert werden kann. Es ist jedoch zu vermuten, dass die Iteriertenfolge eines Abstiegsverfahrens bei einer, je nach Entfernung von dem angestrebten kritischen Punkt, zu geringen Verminderung des Zielfunktionswertes von  $(P)$  pro Iteration nicht konvergieren könnte (s. [Alt02, S. 76] für ein Beispiel). Wegen  $\psi(0) = \psi(\hat{t}) = 0$  sollte daher die Schrittweite nicht zu nahe bei 0 oder  $\hat{t}$  liegen bzw. sollte  $\psi(t)$  genügend groß sein. Das nächste Lemma dient dazu, eine geeignete Forderung an die Schrittweiten zu formulieren.

**Lemma 3.12** *Es seien (V1)–(V3) erfüllt und*

$$x \in N_0, \quad p \in \mathbb{R}^n \quad \text{mit} \quad \nabla f(x)^T p < 0$$

gegeben. Weiter sei  $\hat{t} := \hat{t}(x, p) > 0$  die nach Lemma 3.9 existierende erste positive Nullstelle von  $\psi$  in (3.21). Dann gilt:

$$(i) \quad \hat{t} \geq -\frac{2 \nabla f(x)^T p}{\gamma \|p\|^2} := \tilde{t} > 0.$$

$$(ii) \quad \psi(t) \geq -t \nabla f(x)^T p - t^2 \frac{\gamma}{2} \|p\|^2 := \Psi(t), \quad t \in [0, \hat{t}].$$

**Beweis.** Nach Lemma 3.9 hat man  $x + tp \in N_0$  für alle  $t \in [0, \hat{t}]$ . Weiter ist  $\psi(0) = 0$  und somit

$$\psi(t) = \psi'(0)t + \int_0^t [\psi'(s) - \psi'(0)] ds.$$

Für alle  $t \in [0, \hat{t}]$  folgt daher unter Anwendung von (V3) mit einer ähnlichen Abschätzung für das Integral wie in (3.20)

$$\begin{aligned} f(x) - f(x + tp) &= -t \nabla f(x)^T p - \int_0^t [\nabla f(x + sp) - \nabla f(x)]^T p ds \\ &\geq -t \nabla f(x)^T p - t^2 \frac{\gamma}{2} \|p\|^2. \end{aligned} \quad (3.22)$$

Damit ist (ii) bewiesen. Setzt man  $t := \hat{t}$  in (3.22) ein, so erhält man wegen  $\psi(\hat{t}) = 0$  die Abschätzung in (i). ■

In Lemma 3.12 werden ein  $\tilde{t} \in [0, \hat{t}]$  und die Parabel

$$\Psi(t) := -t \nabla f(x)^T p - t^2 \frac{\gamma}{2} \|p\|^2, \quad t \in \mathbb{R},$$

definiert. Wie man nachrechnet, gilt insbesondere  $\Psi(0) = \Psi(\tilde{t}) = 0$  und nimmt die Funktion  $\Psi$  ihr Maximum bei

$$\frac{\tilde{t}}{2} = -\frac{1}{\gamma} \frac{\nabla f(x)^T p}{\|p\|^2} \in [0, \hat{t}] \quad (3.23)$$

an. Das Maximum von  $\Psi$  hat den Wert

$$\max_{t \in [0, \hat{t}]} \Psi(t) = \Psi\left(\frac{\tilde{t}}{2}\right) = \frac{1}{2\gamma} \left\{ \frac{\nabla f(x)^T p}{\|p\|} \right\}^2. \quad (3.24)$$

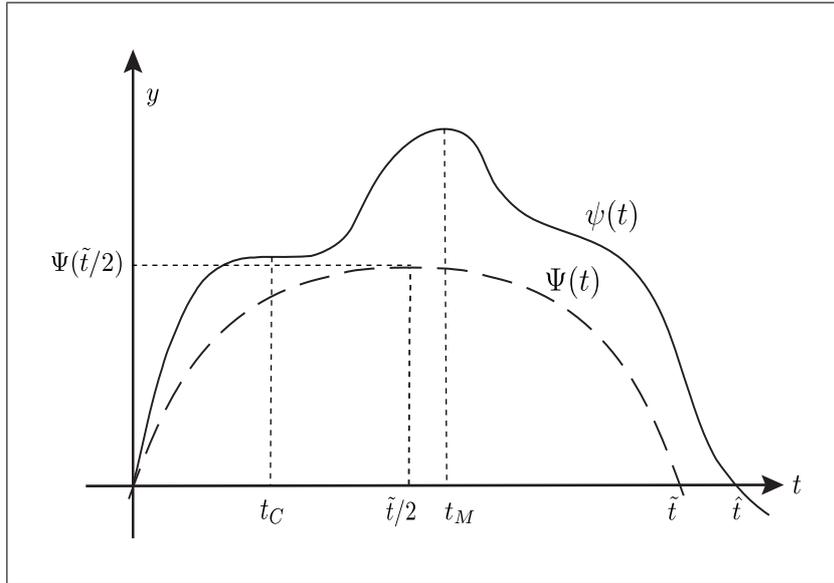


Abb. 3.3: Schrittweitenbedingung

Nach Lemma 3.12 gilt ferner  $\psi(t) \geq \Psi(t)$  für alle  $t \in [0, \hat{t}]$ . Wünschenswert ist es nun, dass  $\psi(t)$ , d. h. die Verminderung des Zielfunktionswertes von  $f$  beim Übergang von  $x$  zu  $x + tp$ , die Größenordnung von  $\{\nabla f(x)^T p / \|p\|\}^2$  besitzt und idealerweise größer oder gleich dem maximalen Wert von  $\Psi$  aus (3.24) ist (vgl. Abb. 3.3). Letzteres ist insbesondere für eine Minimumschrittweite (3.9) der Fall, wie mit Satz 4.3 gezeigt werden wird. Aus diesem Grund definiert man:

**Definition 3.13** Eine Schrittweitenregel heißt effizient (mit Konstante  $\vartheta$ ), wenn sie jedem Paar

$$x \in N_0, \quad p \in \mathbb{R}^n \quad \text{mit} \quad \nabla f(x)^T p < 0$$

ein wohldefiniertes (nicht notwendig eindeutig bestimmtes)  $t := t(x, p) > 0$  zugeordnet und wenn ein von  $x$ ,  $p$  und  $t$  unabhängiges  $\vartheta > 0$  existiert, so dass gilt:

$$f(x) - f(x + tp) \geq \vartheta \left\{ \frac{\nabla f(x)^T p}{\|p\|} \right\}^2. \quad (3.25)$$

Gilt entsprechend nur

$$f(x) - f(x + tp) \geq \vartheta \min \left( -\nabla f(x)^T p, \left\{ \frac{\nabla f(x)^T p}{\|p\|} \right\}^2 \right), \quad (3.26)$$

so heißt die *Schrittweitenregel semieffizient*. Bei Verwendung einer effizienten bzw. semieffizienten Schrittweitenregel bezeichnen wir auch kurz die Schrittweiten selbst als *effizient* bzw. *semieffizient*.

Spellucci [Spe93] führt für eine Schrittweitenregel das *Prinzip des hinreichenden Abstiegs* ein, welches die Bedingung (3.25) der Effizienz unmittelbar impliziert. (Eine Motivation dafür wird in [Alt02] gegeben.) Der Begriff einer effizienten Schrittweitenregel geht auf die in diesem Zusammenhang grundlegende Arbeit [WaWe77] von Warth und Werner zurück. Er wurde von Kosmol [Kos89] aufgegriffen, der zusätzlich den Begriff der semieffizienten Schrittweitenregel definierte. Da jede effiziente Schrittweitenregel offenbar auch semieffizient ist und man sich diese Implikation anhand der gewählten Benennungen gut merken kann, haben wir diese Bezeichnungen übernommen. (Alle unten in Kapitel 4 eingeführten Schrittweitenregeln sind somit zumindest semieffizient.)

Für manche Verfahren kann man Konvergenz beweisen, wenn sie mit einer effizienten Schrittweitenregel verbunden werden, während dies für eine semieffiziente Regel nicht gelingt oder bisher nicht gelungen ist. Die Bezeichnungen in Definition 3.13 sind aber insofern etwas irreführend, als sie nichts über die numerische Effizienz eines Verfahrens bei Verwendung einer entsprechenden Regel aussagen und eine „semieffiziente“ Schrittweitenregel im Hinblick auf die Konvergenzgeschwindigkeit oder den numerischen Aufwand eines Verfahrens nicht notwendig weniger effizient als eine „effiziente“ ist.

### 3.6 Konvergenzaussagen

Wir kehren nun zu dem Modellalgorithmus 3.6 zurück. Wie wir dort festgestellt haben, ist dieser generell durchführbar. In diesem Abschnitt wollen wir Konvergenzaussagen für dieses allgemeine Modell eines Abstiegsverfahrens beweisen, wobei wir das Modell nur insoweit spezifizieren, als dass wir voraussetzen, dass in Schritt (3) eine effiziente Schrittweitenregel verwendet wird. Anschließend werden wir für eine wichtige Klasse von Abstiegsrichtungen auch Konvergenz des Verfahrens für semieffiziente Schrittweiten zeigen.

Wir beginnen damit, dass wir in dem folgenden Lemma unter den relativ schwachen Voraussetzungen (V1) und (V2) einige, zum Teil offenkundige Aussagen für den Modellalgorithmus zusammenfassen. Es sei daran erinnert, dass die Niveaumenge  $N_0$  in (V2) durch den Startpunkt  $x^0$  des Verfahrens definiert wird und dass die Voraussetzungen (V1) und (V2) die Existenz einer globalen Lösung des unrestringierten Optimierungsproblems (P) garantieren (Bemerkung 3.7). Folglich dürfen wir „ $\min_{x \in \mathbb{R}^n} f(x)$ “ schreiben.

**Lemma 3.14** *Es seien (V1) und (V2) erfüllt, und der Modellalgorithmus 3.6 sei mit einer beliebigen Schrittweitenregel versehen. Dann gilt für alle  $k$ :*<sup>2</sup>

(i)  $f(x^{k+1}) < f(x^k) < f(x^0)$ .

(ii)  $x^k \in N_0$ .

*Bricht der Algorithmus nicht nach endlich vielen Schritten mit einer kritischen Lösung von Problem (P) ab, so erzeugt er eine unendliche Folge  $\{x^k\}$  mit folgenden Eigenschaften:*

(iii)  $\{x^k\}$  besitzt einen Häufungspunkt.

(iv)  $\lim_{k \rightarrow \infty} f(x^k) = \hat{f}$  für ein  $\hat{f} \geq \min_{x \in \mathbb{R}^n} f(x)$ .

(v) Gilt zusätzlich  $\lim_{k \rightarrow \infty} \nabla f(x^k) = 0$ , so folgt:

( $\alpha$ ) Jeder Häufungspunkt von  $\{x^k\}$  ist kritische Lösung von (P).

( $\beta$ ) Hat (P) genau eine kritische Lösung  $x^*$  in  $N_0$ , so ist  $\lim_{k \rightarrow \infty} x^k = x^*$ .

**Beweis.** Aussage (i) ist nach Konstruktion des Algorithmus trivialerweise erfüllt und impliziert Aussage (ii). Ist  $\{x^k\}$  eine unendliche Folge, so folgt aus (ii) wegen (V2) die Aussage (iii) (vgl. Bemerkung 3.7). Schließlich garantiert Aussage (i) wegen

$$\min_{x \in \mathbb{R}^n} f(x) \leq f(x^{k+1}) < f(x^k), \quad k \in \mathbb{N}_0,$$

dass die Folge der Funktionswerte  $\{f(x^k)\}$  monoton fallend und nach unten beschränkt ist. Dies hat die in (iv) angegebene Konvergenz zur Folge.

Es gelte nun  $\lim_{k \rightarrow \infty} \nabla f(x^k) = 0$ . Ist  $x^*$  ein Häufungspunkt von  $\{x^k\}$ , so existiert eine Teilfolge  $\{x^{k_i}\}$  von  $\{x^k\}$  mit  $\lim_{i \rightarrow \infty} x^{k_i} = x^*$  und folgt wegen der vorausgesetzten Stetigkeit von  $\nabla f$

$$0 = \lim_{i \rightarrow \infty} \nabla f(x^{k_i}) = \nabla f(x^*).$$

Also ist die Aussage ( $\alpha$ ) richtig.

---

<sup>2</sup>Wir schreiben „für alle  $k$ “, wenn nicht klar ist, ob die Iterierten  $x^k$  nur für endlich viele  $k$  definiert sind, da der Algorithmus nach endlich vielen Schritten mit einer kritischen Lösung abbricht. Erzeugt er eine unendliche Folge  $\{x^k\}$ , so machen wir dies durch einen Zusatz wie „ $k \in \mathbb{N}_0$ “ deutlich.

Jetzt sei angenommen, dass  $(P)$  genau eine kritische Lösung  $x^*$  in  $N_0$  besitzt. Würde  $\{x^k\}$  nicht gegen  $x^*$  konvergieren, dann existierte ein  $\varepsilon > 0$ , so dass für unendlich viele  $x^k$ , d. h. für eine Teilfolge  $\{x^{k_i}\}$  von  $\{x^k\}$  gelten würde:

$$\|x^{k_i} - x^*\| \geq \varepsilon, \quad i \in \mathbb{N}_0. \quad (3.27)$$

Weil  $x^{k_i}$  gemäß Aussage (ii) in  $N_0$  liegt und  $N_0$  gemäß (V2) kompakt ist, könnte dann jedoch aus  $\{x^{k_i}\}$  eine konvergente Teilfolge  $\{x^{k_{i_j}}\}$  ausgewählt werden, die nach Aussage ( $\alpha$ ) notwendig gegen den einzigen kritischen Punkt  $x^*$  von  $(P)$  in  $N_0$  konvergieren würde. Letzteres steht aber im Widerspruch zu (3.27), wie man sieht, wenn man dort  $i := i_j$  mit  $j \in \mathbb{N}$  setzt. Damit ist alles bewiesen. ■

**Aufgabe 3.15** *Unter den Voraussetzungen von Lemma 3.14 (iii) und (iv) beweise man: Sind  $x^*$  und  $x^{**}$  zwei Häufungspunkte der Folge  $\{x^k\}$ , so ist  $f(x^*) = f(x^{**})$ .*

Wenn man von *Konvergenz eines Verfahrens* zur Lösung des unrestringierten Optimierungsproblems  $(P)$  spricht, meint man damit im Allgemeinen, dass jeder Häufungspunkt der durch das Verfahren erzeugten Iteriertenfolge eine kritische Lösung von  $(P)$  ist (siehe Aussage ( $\alpha$ ) von Lemma 3.14). Im Fall, dass  $(P)$  eine eindeutige Lösung besitzt, folgt dann unter den Voraussetzungen des Lemmas auch die Konvergenz der ganzen Folge (Aussage ( $\beta$ ) von Lemma 3.14).

Wir sind nun in der Lage, den zentralen Konvergenzsatz für den Modellalgorithmus zu beweisen. Mit Hilfe dieses Satzes werden wir später auch die Konvergenz von speziellen Verfahren verifizieren.

**Satz 3.16** *Es seien (V1) und (V2) erfüllt. Der Modellalgorithmus 3.6 sei mit einer effizienten Schrittweitenregel versehen und breche nicht nach endlich vielen Iterationen mit einer kritischen Lösung von Problem  $(P)$  ab. Weiter gelte für die durch ihn erzeugten unendlichen Folgen  $\{x^k\}$  und  $\{p^k\}$  die Zoutendijk-Bedingung*

$$\sum_{k=0}^{\infty} \alpha_k^2 = \infty \quad \text{für} \quad \alpha_k := -\frac{\nabla f(x^k)^T p^k}{\|\nabla f(x^k)\| \|p^k\|}. \quad (3.28)$$

Dann folgt:

(i) *Die Folge  $\{x^k\}$  hat mindestens einen Häufungspunkt, der kritische Lösung von  $(P)$  ist.*

(ii) *Sind zusätzlich (V3) und (V4) erfüllt und ist  $x^*$  die dann existierende eindeutige Lösung von  $(P)$ , so gilt  $\lim_{k \rightarrow \infty} x^k = x^*$ .*

**Beweis.** Wir nehmen an, dass der Algorithmus nicht nach endlich vielen Schritten abbricht. Da die Schrittweitenregel im Algorithmus effizient ist, gilt dann mit einem  $\vartheta > 0$  für alle  $k$

$$f(x^k) - f(x^{k+1}) \geq \vartheta \left\{ \frac{\nabla f(x^k)^T p^k}{\|p^k\|} \right\}^2 = \vartheta \alpha_k^2 \|\nabla f(x^k)\|^2 > 0. \quad (3.29)$$

(i) Durch Summation von (3.29) für  $k = 0, \dots, \ell$  erhalten wir

$$\sum_{k=0}^{\ell} \alpha_k^2 \|\nabla f(x^k)\|^2 \leq \frac{1}{\vartheta} \sum_{k=0}^{\ell} [f(x^k) - f(x^{k+1})] = \frac{1}{\vartheta} [f(x^0) - f(x^{\ell+1})].$$

Folglich bekommen wir unter Verwendung von Lemma 3.14 (iv) durch Grenzübergang für  $\ell \rightarrow \infty$

$$\sum_{k=0}^{\infty} \alpha_k^2 \|\nabla f(x^k)\|^2 \leq \frac{1}{\vartheta} [f(x^0) - \hat{f}] < \infty. \quad (3.30)$$

Wäre nun kein Häufungspunkt von  $\{x^k\}$  kritischer Punkt von  $(P)$ , so existierte ein  $\varepsilon > 0$  mit

$$\varepsilon \leq \|\nabla f(x^k)\|, \quad k \in \mathbb{N}_0. \quad (3.31)$$

(Denn sonst gäbe es eine Teilfolge  $\{x^{k_j}\}$  von  $\{x^k\}$  mit  $\lim_{j \rightarrow \infty} \nabla f(x^{k_j}) = 0$ ; da Lemma 3.14 aber genauso für die Folge  $\{x^{k_j}\}$  gültig ist, folgte daraus die Existenz eines Häufungspunktes von  $\{x^{k_j}\}$  und damit eines Häufungspunktes von  $\{x^k\}$ , der kritische Lösung von  $(P)$  wäre.) Mit (3.31) würde aber aus (3.30) die Abschätzung

$$\varepsilon^2 \sum_{k=0}^{\infty} \alpha_k^2 \leq \sum_{k=0}^{\infty} \alpha_k^2 \|\nabla f(x^k)\|^2 < \infty,$$

folgen, welche der vorausgesetzten Zoutendijk-Bedingung widerspricht.

(ii) Aus (3.29) können wir mit Aussage (v) aus Lemma 3.10 die folgende Abschätzung gewinnen:

$$f(x^k) - f(x^{k+1}) \geq \vartheta \alpha_k^2 \|\nabla f(x^k)\|^2 \geq 2\beta \vartheta \alpha_k^2 [f(x^k) - f(x^*)].$$

Auflösen dieser Abschätzung nach  $f(x^{k+1}) - f(x^*)$ , Mehrfachanwendung des erhaltenen Ergebnisses sowie Verwendung der Beziehung  $1 + x \leq e^x$ ,  $x \in \mathbb{R}$ , liefern

$$\begin{aligned} 0 \leq f(x^{k+1}) - f(x^*) &\leq (1 - 2\beta \vartheta \alpha_k^2) [f(x^k) - f(x^*)] \\ &\leq [f(x^0) - f(x^*)] \prod_{i=0}^k (1 - 2\beta \vartheta \alpha_i^2) \\ &\leq [f(x^0) - f(x^*)] \exp\left(-2\beta \vartheta \sum_{i=0}^k \alpha_i^2\right). \end{aligned}$$

Da die Zoutendijk-Bedingung  $\lim_{k \rightarrow \infty} \sum_{i=0}^k \alpha_i^2 = \infty$  impliziert, können wir damit

$$\lim_{k \rightarrow \infty} f(x^k) = f(x^*)$$

schließen. Die Konvergenz  $x^k \rightarrow x^*$  ( $k \rightarrow \infty$ ) ist nun eine Konsequenz von Lemma 3.10 (iii). ■

**Bemerkung 3.17** *Im Fall, dass (V1)–(V4) erfüllt sind, wird die Konvergenz des Modellalgorithmus 3.6 vollständig durch die Zoutendijk-Bedingung beschrieben. Denn dann gilt auch die Umkehrung der Aussage (ii) von Satz 3.16, dass die Konvergenz von  $\{x^k\}$  die Zoutendijk-Bedingung impliziert (vgl. [WaWe77]).*

Der Winkel  $\sphericalangle(u, v)$  zwischen zwei Vektoren  $u, v \in \mathbb{R}^n \setminus \{0\}$  wird bekanntlich über die Beziehung

$$\cos(\sphericalangle(u, v)) = \frac{u^T v}{\|u\| \|v\|}$$

definiert. Somit ergibt sich für  $\alpha_k$  aus (3.28)

$$\alpha_k = -\frac{\nabla f(x^k)^T p^k}{\|\nabla f(x^k)\| \|p^k\|} = \cos(\sphericalangle(-\nabla f(x^k), p^k)) > 0.$$

Die Zoutendijk-Bedingung impliziert demzufolge, dass die Konstante  $\alpha_k$  für  $k \rightarrow \infty$  nicht zu schnell gegen 0 bzw. der Winkel zwischen  $-\nabla f(x^k)$  und  $p^k$  nicht zu schnell gegen  $90^\circ$  streben darf (siehe Abb. 3.4).

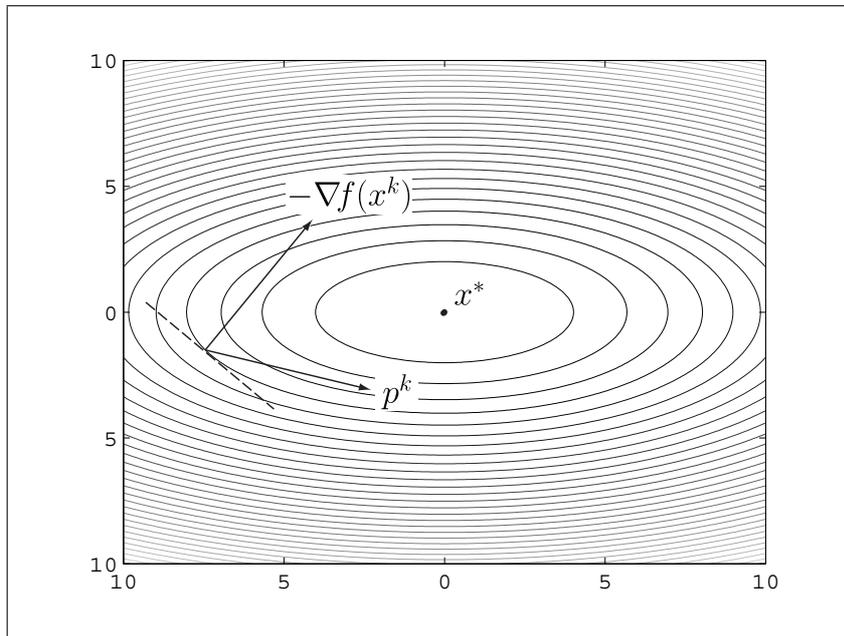


Abb. 3.4: Winkel zwischen  $-\nabla f(x^k)$  und  $p^k$

Die Zoutendijk-Bedingung ist offenbar für das Gradientenverfahren mit  $\alpha_k = 1$  erfüllt und allgemeiner für jedes Verfahren erfüllt, für das mit einem  $\sigma > 0$  gilt:

$$\alpha_k \geq \sigma, \quad k \in \mathbb{N}_0. \quad (3.32)$$

Verfahren mit der Eigenschaft (3.32) bezeichnet man als *gradientenähnliche Verfahren*. Demnach ist Satz 3.16 insbesondere für gradientenähnliche Verfahren relevant. Zu diesen Verfahren gehören, wie in Aufgabe 3.19 gezeigt werden soll, die wichtigen Verfahren, bei denen die Richtung

$$p^k := -H_k \nabla f(x^k) \quad (3.33)$$

mit einer symmetrischen, positiv definiten Matrix  $H_k \in \mathbb{R}^{n \times n}$  gewählt wird (vgl. Beispiel 3.3) und bei denen für alle  $k$  die Bedingung

$$m \|x\|^2 \leq x^T H_k x \leq M \|x\|^2, \quad x \in \mathbb{R}^n, \quad (3.34)$$

mit Konstanten  $m, M > 0$  erfüllt ist.

**Bemerkung 3.18** *Sind die  $H_k$ 's für alle  $k$  symmetrische Matrizen, welche der Bedingung (3.34) genügen und damit positiv definit sind, so folgt*

$$0 < m \leq \lambda_{\min}(H_k) \leq \lambda_{\max}(H_k) = \|H_k\| \leq M, \quad (3.35)$$

wobei  $\lambda_{\min}(H_k)$  der kleinste und  $\lambda_{\max}(H_k)$  der größte Eigenwert von  $H_k$  ist (vgl. Lemma 2.19 und Bemerkung 2.20). Demnach impliziert die Bedingung (3.34), dass die kleinsten Eigenwerte der  $H_k$ 's „gleichmäßig von Null weg beschränkt“ sind. Man sagt in diesem Fall auch, dass die  $H_k$ 's gleichmäßig positiv definit sind.

Verfahren des zuletzt genannten Typs sind gradientenähnlich.

**Aufgabe 3.19** *Zeigen Sie, dass Verfahren mit einer Richtungswahl wie in (3.33) unter der Voraussetzung, dass die  $H_k$ 's symmetrische Matrizen sind, welche für alle  $k$  die Bedingung (3.34) erfüllen, gradientenähnlich sind.*

Für solche Verfahren können wir nun - sogar in Verbindung mit einer semieffizienten Schrittweitenregel - zu einer stärkeren Konvergenzaussage als der in Satz 3.16 gelangen. Denn letztere ist sehr schwach, da sie nicht ausschließt, dass die durch das Verfahren erzeugte Iteriertenfolge  $\{x^k\}$  Häufungspunkte besitzt, die nicht kritische Punkte von  $(P)$  sind.

**Satz 3.20** *Es seien (V1)–(V3) erfüllt, und der Modellalgorithmus 3.6 sei mit einer semieffizienten Schrittweitenregel mit Konstante  $\vartheta > 0$  versehen. Weiter gelte für alle  $k$*

$$p^k := -H_k \nabla f(x^k), \quad (3.36)$$

wobei die  $H_k$ 's symmetrische Matrizen seien, welche mit Konstanten  $m, M > 0$  für alle  $k$  der folgenden Bedingung genügen:

$$m \|x\|^2 \leq x^T H_k x \leq M \|x\|^2, \quad x \in \mathbb{R}^n. \quad (3.37)$$

Bricht dann der Algorithmus nicht nach endlich vielen Schritten ab, so hat man:

- (i) Jeder Häufungspunkt von  $\{x^k\}$  ist kritische Lösung von (P).
- (ii) Besitzt (P) genau eine kritische Lösung  $x^*$  in  $N_0$ , so gilt  $\lim_{k \rightarrow \infty} x^k = x^*$ .
- (iii) Ist auch (V4) erfüllt und  $x^*$  die dann existierende eindeutige Lösung von (P), so folgt  $\lim_{k \rightarrow \infty} x^k = x^*$  und gelten mit Konstanten  $\nu \in (0, 1)$  und  $c > 0$  die Abschätzungen

$$0 < f(x^{k+1}) - f(x^*) \leq \nu [f(x^k) - f(x^*)], \quad k \in \mathbb{N}_0, \quad (3.38)$$

und

$$\|x^k - x^*\| \leq c (\sqrt{\nu})^k, \quad k \in \mathbb{N}_0, \quad (3.39)$$

wobei  $\nu$  definiert ist durch

$$\nu := 1 - 2\beta\vartheta \min(m, m^2/M^2).$$

**Beweis.** Gemäß (2.17) und Bemerkung 3.18 haben wir

$$\|H_k\| = \lambda_{\max}(H_k) \leq M. \quad (3.40)$$

Damit erreichen wir unter Verwendung von (3.26)

$$\begin{aligned} f(x^k) - f(x^{k+1}) &\geq \vartheta \min \left( -\nabla f(x^k)^T p^k, \left\{ \frac{\nabla f(x^k)^T p^k}{\|p^k\|} \right\}^2 \right) \\ &= \vartheta \min \left( \nabla f(x^k)^T H_k \nabla f(x^k), \left\{ \frac{\nabla f(x^k)^T H_k \nabla f(x^k)}{\|H_k \nabla f(x^k)\|} \right\}^2 \right) \\ &\geq \vartheta \min \left( m \|\nabla f(x^k)\|^2, \left\{ \frac{m \|\nabla f(x^k)\|^2}{M \|\nabla f(x^k)\|} \right\}^2 \right) \\ &= \vartheta \min \left( m, \frac{m^2}{M^2} \right) \|\nabla f(x^k)\|^2. \end{aligned} \quad (3.41)$$

Da Aussage (iv) von Lemma 3.14 den Grenzwert

$$\lim_{k \rightarrow \infty} [f(x^k) - f(x^{k+1})] = 0$$

impliziert, schließen wir aus (3.41) die Konvergenz  $\|\nabla f(x^k)\| \rightarrow 0$  ( $k \rightarrow \infty$ ). Aussagen (i) und (ii) folgen nun mit Lemma 3.14 (v). Der Beweis der Aussage (iii) wird als Aufgabe 3.21 gestellt. ■

**Aufgabe 3.21** Man beweise Satz 3.20 (iii).

Die zweite Abschätzung in (3.38) ist möglicherweise pessimistisch. Die Abschätzungen in (3.38) und (3.39) zeigen aber, dass der Modellalgorithmus unter den Voraussetzungen und Spezifikationen von Satz 3.20 (iv) bezüglich der Folge der Funktionswerte  $\{f(x^k)\}$  mindestens  $Q$ -linear und bezüglich der Iteriertenfolge mindestens  $R$ -linear konvergiert.



### 3.7 Lösungen der Aufgaben

Aufgabe 3.5 Für  $p \in \mathbb{R}^n$  erhält man mit der Cauchy-Schwarz-Ungleichung für den Zielfunktionswert von Problem (3.8) die folgende Abschätzung nach unten.

$$\begin{aligned}\nabla f(x)^T p &= (\nabla f(x)^T A^{-1}) A p = \langle A^{-1} \nabla f(x), p \rangle_A \\ &\geq - \|A^{-1} \nabla f(x)\|_A \|p\|_A.\end{aligned}$$

Dabei wird die untere Schranke für  $p^*$  angenommen, denn es gilt  $\|p^*\|_A = 1$  und

$$\begin{aligned}\nabla f(x)^T p^* &= - \frac{\nabla f(x)^T A^{-1} \nabla f(x)}{\|A^{-1} \nabla f(x)\|_A} = - \frac{(A^{-1} \nabla f(x))^T A (A^{-1} \nabla f(x))}{\|A^{-1} \nabla f(x)\|_A} \\ &= - \frac{\|A^{-1} \nabla f(x)\|_A^2}{\|A^{-1} \nabla f(x)\|_A} = - \|A^{-1} \nabla f(x)\|_A \|p^*\|_A.\end{aligned}\quad (3.42)$$

Somit ist  $p^*$  eine Lösung des Problems (3.8). Da  $A^{-1}$  nichtsingulär und  $\nabla f(x) \neq 0$  ist, folgt weiter  $A^{-1} \nabla f(x) \neq 0$  und damit nach (3.42)  $\nabla f(x)^T p^* < 0$ . Die Eindeutigkeit von  $p^*$  erschließt man wie für den in Bemerkung 3.4 betrachteten Fall  $A := I$ .

Aufgabe 3.15 Gemäß Lemma 3.14 (iv) konvergiert die Folge der Funktionswerte  $\{f(x^k)\}$  und damit auch jede ihrer Teilfolgen gegen ein  $\hat{f}$ . Wenn  $\{x^{k_i}\}$  und  $\{x^{\ell_i}\}$  Teilfolgen von  $\{x^k\}$  sind, welche gegen  $x^*$  bzw.  $x^{**}$  konvergieren, folgt daher aus der Stetigkeit von  $f$

$$\lim_{i \rightarrow \infty} f(x^{k_i}) = f(x^*) = \hat{f} = \lim_{i \rightarrow \infty} f(x^{\ell_i}) = f(x^{**}).$$

Aufgabe 3.19 Für alle  $k$  folgt mit (3.34) und (3.35)

$$\begin{aligned}\alpha_k &= - \frac{\nabla f(x^k)^T p^k}{\|\nabla f(x^k)\| \|p^k\|} = \frac{\nabla f(x^k)^T H_k \nabla f(x^k)}{\|\nabla f(x^k)\| \|H_k \nabla f(x^k)\|} \\ &\geq \frac{m \|\nabla f(x^k)\|^2}{M \|\nabla f(x^k)\|^2} \geq \frac{m}{M} > 0.\end{aligned}$$

Aufgabe 3.21 Sind (V1)–(V4) erfüllt, so besitzt  $(P)$  einen eindeutigen kritischen Punkt  $x^*$  in  $N_0$ , welcher die eindeutige Lösung von  $(P)$  ist (Bemerkung 3.7). Nach Aussage (ii) gilt somit  $\lim_{k \rightarrow \infty} x^k = x^*$ . Weiter ist nach Lemma 3.14 (ii)  $x^k \in N_0$  für alle  $k$  und hat man somit nach Aussage (v) von Lemma 3.10

$$\|\nabla f(x^k)\|^2 \geq 2\beta [f(x^k) - f(x^*)] \geq 0.$$

Damit liefert die Abschätzung in (3.41)

$$f(x^k) - f(x^{k+1}) \geq 2\beta\vartheta \min\left(m, \frac{m^2}{M^2}\right) [f(x^k) - f(x^*)] \geq 0$$

bzw.

$$-2\beta\vartheta \min\left(m, \frac{m^2}{M^2}\right) [f(x^k) - f(x^*)] \geq f(x^{k+1}) - f(x^k).$$

Addition von  $f(x^k) - f(x^*)$  auf beiden Seiten liefert die Ungleichung in (3.38).

Mit Lemma 3.10 (iii) und (3.38) schließt man

$$\begin{aligned} \frac{\beta}{2} \|x^k - x^*\|^2 &\leq f(x^k) - f(x^*) \leq \nu [f(x^{k-1}) - f(x^*)] \\ &\leq \dots \leq \nu^k [f(x^0) - f(x^*)]. \end{aligned}$$

Letzteres impliziert  $\nu > 0$  und die Abschätzung in (3.39) mit. Dass  $\nu < 1$  ist, folgt wegen  $2\beta\vartheta \min(m, m^2/M^2) > 0$ .

## 4 Schrittweitenregeln

Die Konstruktion von konkreten Verfahren zur Lösung des unrestringierten Optimierungsproblems

$$(P) : \quad \text{Minimiere } f(x) \text{ über alle } x \in \mathbb{R}^n,$$

welche vom Typ des Modellalgorithmus 3.6 sind, erfordert die Festlegung einer Regel zur Bestimmung einer Abstiegsrichtung in der aktuellen Iterierten und einer Regel zur Berechnung der Schrittweite für diese Richtung. Da nun bekannt ist, welche Bedingung die von einer bestimmten Regel erzeugten Schrittweiten im Hinblick auf die Konvergenz eines Verfahrens erfüllen sollten (Definition 3.13), ist es möglich, diese beiden Schritte zur Spezifikation eines Verfahrens voneinander zu trennen.

In diesem Kapitel wollen wir einige bekannte Regeln für die Schrittweitenbestimmung vorstellen und zeigen, dass diese effizient oder semieffizient sind. Sofern erforderlich, werden wir auch Verfahren zu ihrer Berechnung diskutieren. Wir beginnen in Abschnitt 4.1 mit Regeln, welche die sog. *exakten Schrittweiten*, die *Minimumschrittweiten* und die *Curry-Schrittweite*, beinhalten. In Abschnitt 4.2 werden wir die *Armijo-Schrittweite* einführen, welche unter allen Schrittweiten die am einfachsten zu berechnende, aber im Unterschied zu allen anderen hier behandelten auch nur semieffizient ist. Anschließend werden wir in Abschnitt 4.3 die *Wolfe-Powell-Schrittweitenregel* sowie einen Algorithmus für deren Berechnung angeben. Zuletzt werden wir in Abschnitt 4.4 eine wichtige Variante dieser Regel, die *strenge Wolfe-Powell-Schrittweitenregel*, untersuchen.

Wir setzen hier generell die Bedingungen (V1)–(V3) voraus und betrachten wieder einen Punkt  $x \in N_0$  und eine Abstiegsrichtung  $p$  in  $x$ , für welche eine Schrittweite bestimmt werden muss, d. h., wir betrachten ein Paar von Vektoren

$$x \in N_0, \quad p \in \mathbb{R}^n \text{ mit } \nabla f(x)^T p < 0. \quad (4.1)$$

Insbesondere ist also  $\nabla f(x) \neq 0$  und  $x$  somit kein kritischer Punkt.

### 4.1 Exakte Schrittweiten

#### 4.1.1 Existenz und Effizienz

Eine naheliegende Schrittweitenregel ist die, einen globalen Minimalpunkt des eindimensionalen Optimierungsproblems

$$\inf_{t \in [0, \infty)} f(x + tp)$$

als Schrittweite zu wählen. Denn ein globaler Minimalpunkt von  $f(x + tp)$  liefert den größten Fortschritt bei der Reduzierung des Zielfunktionswertes von  $f$  bei  $x$  in Richtung  $p$ . Wir definieren also:

**Definition 4.1** Jedes  $t_M := t_M(x, p)$ , für welches

$$f(x + t_M p) = \inf_{t \in [0, \infty)} f(x + tp)$$

gilt, heißt Minimumschrittweite.

Es ist jedoch nur für einfache Funktionen wie z. B. konvexe Funktionen und dann zumeist auch nur näherungsweise möglich, eine Minimumschrittweite zu berechnen. Für alle anderen Funktionen ist es daher praxisnäher, den kleinsten, positiven, kritischen Punkt der Funktion  $\varphi(t) := f(x + tp)$  als Schrittweite zu akzeptieren. Diese Überlegung motiviert die folgende Definition.

**Definition 4.2** Die Curry-Schrittweite  $t_C := t_C(x, p)$  ist definiert durch

$$t_C(x, p) := \inf \left\{ \bar{t} \in [0, \infty) \mid \left. \frac{d}{dt} f(x + tp) \right|_{t=\bar{t}} = 0 \right\}.$$

Minimumschrittweiten und die Curry-Schrittweite bezeichnet man auch als *exakte Schrittweiten*. Für diese können wir zeigen:

**Satz 4.3** Es seien (V1)–(V3) erfüllt. Für alle Paare  $x$  und  $p$  mit (4.1) existieren eine Minimumschrittweite und die Curry-Schrittweite, und diese sind positive Zahlen. Ferner sind die Curry- und jede Minimumschrittweite effiziente Schrittweiten mit der Konstanten

$$\vartheta_M = \vartheta_C := \frac{1}{2\gamma}.$$

**Beweis.** Wir weisen zunächst die Existenz von beiden Schrittweiten nach. Dazu sei  $\varphi(t) := f(x + tp)$ . Nach Lemma 3.9 ist  $\psi(t) = \varphi(0) - \varphi(t)$  auf  $(0, \hat{t})$  für ein  $\hat{t} := \hat{t}(x, p) > 0$  positiv und folglich

$$\varphi(0) > \varphi(t), \quad t \in (0, \hat{t}).$$

Ferner existiert nach diesem Lemma ein  $\kappa \geq \hat{t}$ , so dass

$$\varphi(0) < \varphi(t), \quad t \geq \kappa,$$

gilt. Zusammen erschließt man die Existenz eines  $t_M \in (0, \kappa)$  mit

$$\varphi(t_M) = \min_{t \in [0, \kappa]} \varphi(t) = \min_{t \in [0, \infty)} \varphi(t).$$

Die Menge aller kritischen Punkte von  $\varphi$  in  $[0, \kappa]$ , d. h. die Menge

$$K_\varphi := \{t \in [0, \kappa] \mid \varphi'(t) = 0\},$$

enthält  $t_M$  und ist somit nichtleer. Ferner ist  $K_\varphi$  kompakt, so dass ein kleinster kritischer Punkt  $t_C := \min_{t \in K_\varphi} t$  in  $[0, \kappa]$  existiert. Wegen  $\varphi'(0) = \nabla f(x)^T p < 0$  gilt  $t_C > 0$  und  $\varphi(t_C) < \varphi(0)$ . Mit  $x \in N_0$  hat man folglich auch  $(x + t_C p) \in N_0$ .

Als Nächstes wollen wir für jede Minimumschrittweite  $t_M$  und für die Curry-Schrittweite  $t_C$  eine Ungleichung des Typs (3.25) nachweisen. Mit (V3) erhalten wir

$$\begin{aligned} 0 &= \varphi'(t_C) = \nabla f(x)^T p + (\nabla f(x + t_C p) - \nabla f(x))^T p \\ &\leq \nabla f(x)^T p + \gamma t_C \|p\|^2. \end{aligned}$$

Mit  $\tilde{t} \in (0, \hat{t}]$  aus Lemma 3.12 impliziert dies

$$t_C \geq -\frac{1}{\gamma} \frac{\nabla f(x)^T p}{\|p\|^2} = \frac{\tilde{t}}{2} > 0. \quad (4.2)$$

Wegen  $\varphi'(0) < 0$  und weil  $t_C$  die kleinste positive Nullstelle von  $\varphi'$  ist, gilt weiter  $\varphi'(t) < 0$  für alle  $t \in [0, t_C)$ . Somit folgt aus (4.2), dass  $\varphi(\tilde{t}/2) \geq \varphi(t_C)$  ist, und demnach

$$f(x + t_M p) = \min_{t \in [0, \infty)} f(x + t p) \leq f(x + t_C p) \leq f(x + \frac{\tilde{t}}{2} p).$$

Dies impliziert schließlich mit (4.2) und mit Lemma 3.12 (ii) wegen  $\tilde{t}/2 \leq \hat{t}$

$$\begin{aligned} f(x) - f(x + t_M p) &\geq f(x) - f(x + t_C p) & (4.3) \\ &\geq f(x) - f(x + \frac{\tilde{t}}{2} p) \\ &\geq -\frac{\tilde{t}}{2} \nabla f(x)^T p - \frac{\tilde{t}^2}{4} \frac{\gamma}{2} \|p\|^2 \\ &= \frac{1}{\gamma} \left\{ \frac{\nabla f(x)^T p}{\|p\|} \right\}^2 - \frac{1}{2\gamma} \left\{ \frac{\nabla f(x)^T p}{\|p\|} \right\}^2 \\ &= \frac{1}{2\gamma} \left\{ \frac{\nabla f(x)^T p}{\|p\|} \right\}^2. \quad \blacksquare \end{aligned}$$

Wie man leicht verifiziert, ist die Funktion  $\varphi(t) := f(x + t p)$  konvex, wenn  $f$  konvex ist. Mit Korollar 2.50 und Satz 2.38 schließt man daher:

**Bemerkung 4.4** *Ist  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  konvex und sind die Voraussetzungen von Satz 4.3 erfüllt, so ist  $t_C := t_C(x, p)$  auch eine Minimumschrittweite und zwar unter allen Minimumschrittweiten die kleinste. Ist  $f$  strikt konvex, so existiert genau eine Minimumschrittweite  $t_M := t_M(x, p)$  und ist  $t_M = t_C$ .*

Für eine gleichmäßig konvexe, quadratische Funktion, für welche ja die Voraussetzungen (V1)–(V3) erfüllt sind (vgl. Beispiel 3.8), ist also die Minimumschrittweite eindeutig. Man kann sie in diesem Fall explizit angeben:

**Beispiel 4.5** Für die quadratische Funktion

$$f(x) := \frac{1}{2}x^T Qx + c^T x + \alpha, \quad x \in \mathbb{R}^n,$$

mit beliebiger symmetrischer Matrix  $Q$  hat man  $\nabla f(x) = Qx + c$  und demnach

$$\begin{aligned} 0 &= \frac{d}{dt} f(x + tp) = \nabla f(x + tp)^T p \\ &= [Q(x + tp) + c]^T p = (Qx + c)^T p + t p^T Qp \\ &= \nabla f(x)^T p + t p^T Qp. \end{aligned} \quad (4.4)$$

Ist  $Q$  positiv definit, so folgt für  $x$  und  $p$  wie in (4.1)

$$t_M = t_C = -\frac{(Qx + c)^T p}{p^T Qp} = -\frac{\nabla f(x)^T p}{p^T Qp} > 0. \quad (4.5)$$

Wir bemerken noch:

**Bemerkung 4.6** Für die quadratische Funktion

$$f(x) := \frac{1}{2}x^T Qx + c^T x + \alpha, \quad x \in \mathbb{R}^n,$$

mit positiv definiten Matrix  $Q$  ist  $t_M$  berechenbar und durch (4.5) gegeben. Sei nun  $p$  ein zu dem größten Eigenwert  $\lambda_{\max}(Q) = \gamma$  gehörender Eigenvektor von  $Q$  (vgl. (3.14)), so dass folgt

$$p^T Qp = p^T [\lambda_{\max}(Q) p] = \gamma \|p\|^2.$$

Weiter sei zu diesem  $p$  ein  $x \in N_0$  gewählt mit  $\nabla f(x)^T p < 0$ . (Zum Beispiel ist  $x := x^0$  für  $x^0 := Q^{-1}(-p - c)$  ein solches  $x$ , da man in diesem Fall  $\nabla f(x) = -p$  hat.) Für  $p$  und  $x$  erschließt man dann

$$\begin{aligned} f(x) - f(x + t_C p) &= -t_C (Qx + c)^T p - \frac{1}{2} t_C^2 p^T Qp \\ &= -t_C \nabla f(x)^T p - t_C^2 \frac{\gamma}{2} \|p\|^2 \\ &= \frac{1}{2\gamma} \left\{ \frac{\nabla f(x)^T p}{\|p\|} \right\}^2. \end{aligned}$$

Die Konstante  $\vartheta_M = \vartheta_C := 1/(2\gamma)$  in Satz 4.3 ist demnach „optimal“.

### 4.1.2 Die Methode vom Goldenen Schnitt

Numerisch ist die (näherungsweise) Berechnung einer Minimumschrittweite nur für konvexe bzw. für andere einfache Funktionen wie *unimodale Funktionen* realistisch möglich.

**Definition 4.7** Eine Funktion  $\varphi : [a, b] \rightarrow \mathbb{R}$  heißt *unimodal* (auf  $[a, b]$ ), falls genau ein  $t^* \in [a, b]$  existiert mit  $\varphi(t^*) = \min_{t \in [a, b]} \varphi(t)$  und falls  $\varphi$  auf  $[a, t^*]$  streng monoton fallend und auf  $[t^*, b]$  streng monoton wachsend ist.

Eine unimodale Funktion kann Sattelpunkte besitzen und muss somit nicht eine konvexe Funktion sein. Umgekehrt gilt zumindest:

**Beispiel 4.8** Jede strikt konvexe Funktion  $\varphi \in C[a, b]$  ist unimodal auf  $[a, b]$  (vgl. die Sätze 2.38 und 2.39).

Wir wollen im Folgenden ein *ableitungsfreies Verfahren* zur Berechnung des Minimalpunktes  $t^* \in [a, b]$  einer auf  $[a, b]$  unimodalen Funktion  $\varphi$  beschreiben. Ist  $[a_k, b_k]$  ein Teilintervall von  $[a, b]$  und  $t^* \in [a_k, b_k]$ , dann folgt gemäß Definition 4.7 für Punkte  $s_k$  und  $t_k$  mit  $a_k \leq s_k < t_k \leq b_k$

$$\begin{aligned} \varphi(s_k) > \varphi(t_k) &\Rightarrow \varphi(u) > \varphi(s_k), \quad u \in [a_k, s_k], \\ \varphi(s_k) \leq \varphi(t_k) &\Rightarrow \varphi(u) > \varphi(t_k), \quad u \in (t_k, b_k]. \end{aligned}$$

Demnach muss der Minimalpunkt  $t^*$  von  $\varphi$  im Fall  $\varphi(s_k) > \varphi(t_k)$  in  $[s_k, b_k]$  und im Fall  $\varphi(s_k) \leq \varphi(t_k)$  in  $[a_k, t_k]$  liegen. Es bietet sich also an, das folgende Intervall als neues Suchintervall zu wählen

$$[a_{k+1}, b_{k+1}] := \begin{cases} [s_k, b_k], & \text{falls } \varphi(s_k) > \varphi(t_k), \\ [a_k, t_k], & \text{falls } \varphi(s_k) \leq \varphi(t_k). \end{cases}$$

Dabei ist es erstrebenswert,  $s_k, t_k \in [a_k, b_k]$  mit  $s_k < t_k$  für jedes  $k$  so festzulegen, dass

- $b_k - s_k = t_k - a_k$ , d. h. die Länge des Intervalls  $[a_{k+1}, b_{k+1}]$  unabhängig vom Ausgang der Abfrage „ $\varphi(s_k) > \varphi(t_k)$ “ ist und
- beim Übergang von  $k$  zu  $k + 1$  nur eine neue Funktionsauswertung erforderlich ist.

Die beiden an  $s_k$  und  $t_k$  gestellten Forderungen können, wie mit Aufgabe 4.10 gezeigt werden soll, in der Tat erfüllt werden, indem man jedes Intervall  $[a_k, b_k]$  mittels eines Goldenen Schnittes in zwei Intervalle teilt. Und zwar sagt man, dass ein Intervall  $[a, b]$  durch einen *Goldenen Schnitt* in zwei Intervalle  $[a, c]$  und  $[c, b]$  zerlegt wird, falls gilt:

$$\frac{\text{Länge des ganzen Intervalls}}{\text{Länge des längeren Teilintervalls}} = \frac{\text{Länge des längeren Teilintervalls}}{\text{Länge des kürzeren Teilintervalls}}. \quad (4.6)$$

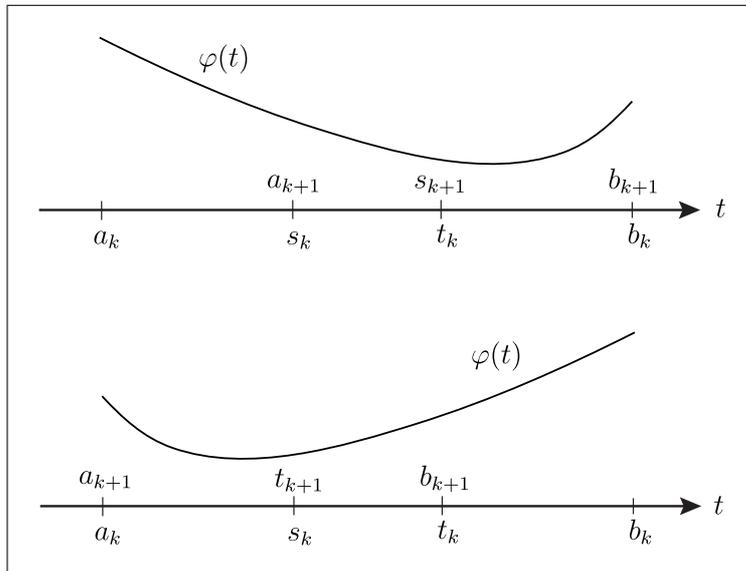


Abb. 4.1: Punktewahl bei Methode des Goldenen Schnittes

Der Punkt  $c$ , der eine solche Zerlegung liefert, lässt sich berechnen (s. Aufgabe 4.10). Und zwar erhält man, wenn  $[a, c]$  das längere Teilintervall ist,

$$c = a + F(b - a), \quad F := \frac{\sqrt{5} - 1}{2} \approx 0.618\,033\,989 \quad (4.7)$$

und, wenn  $[a, c]$  das kürzere Teilintervall ist,

$$c = a + (1 - F)(b - a), \quad 1 - F = \frac{3 - \sqrt{5}}{2} \approx 0.381\,966\,011. \quad (4.8)$$

Demnach wählt man für jedes  $k$  insbesondere (siehe Abb. 4.1)

$$s_k := a_k + (1 - F)(b_k - a_k), \quad t_k := a_k + F(b_k - a_k).$$

Damit haben wir die folgende Methode zur Minimierung einer unimodalen Funktion  $\varphi : [a, b] \rightarrow \mathbb{R}$  beschrieben.

**Algorithmus 4.9** (*Methode vom Goldenen Schnitt*)

(0) Wähle  $\varepsilon \in (0, b - a)$  und setze  $F := (\sqrt{5} - 1)/2$  sowie

$$a_0 := a, \quad b_0 := b, \quad s_0 := a + (1 - F)(b - a), \quad t_0 := a + F(b - a).$$

Berechne  $\varphi(s_0)$  und  $\varphi(t_0)$  und setze  $k := 0$ .

(1) Falls  $b_k - a_k \leq \varepsilon$  ist, stop! (Es gilt  $t^* \in [a_k, b_k]$ .)

(2) (i) Falls  $\varphi(s_k) > \varphi(t_k)$  ist, setze

$$a_{k+1} := s_k, \quad b_{k+1} := b_k, \quad s_{k+1} := t_k, \quad t_{k+1} := s_k + F(b_k - s_k)$$

und berechne  $\varphi(t_{k+1})$ .

(ii) Falls  $\varphi(s_k) \leq \varphi(t_k)$  ist, setze

$$a_{k+1} := a_k, \quad b_{k+1} := t_k, \quad s_{k+1} := a_k + (1 - F)(t_k - a_k), \quad t_{k+1} := s_k$$

und berechne  $\varphi(s_{k+1})$ .

(3) Setze  $k := k + 1$  und gehe nach (1).

**Aufgabe 4.10** Man zeige:

(a) Die Wahl von  $c$  in (4.7) und (4.8) zerlegt das Intervall  $[a, b]$  nach dem Goldenen Schnitt.

(b) Die beiden oben an  $s_k$  und  $t_k$  gestellten Forderungen führen auf die Setzungen in Schritt (2) von Algorithmus 4.9.

(c) Man weise nach, dass Algorithmus 4.9 für eine auf  $[a, b]$  unimodale Funktion  $\varphi$  nach endlich vielen Iterationen mit Schritt (1) abbricht.

(d) Wieviele Funktionsauswertungen benötigt Algorithmus 4.9 für die näherungsweise Bestimmung des Minimalpunktes einer unimodalen Funktion  $\varphi : [-1, 1] \rightarrow \mathbb{R}$ , wenn  $\varepsilon := 0.000\,001$  gewählt wird?

### 4.1.3 Anmerkungen

Das Verfahren vom Goldenen Schnitt oder andere ableitungsfreie Verfahren wie solche Verfahren, die  $f$  in jedem Schritt durch eine Funktion approximieren, welche  $f$  in bestimmten Punkten interpoliert, und die damit eine Näherung für den gesuchten Minimalpunkt berechnen ([Bre73]), benötigen nur Funktionswerte von  $f$  und konvergieren zum Teil unter schwachen Voraussetzungen superlinear. Allerdings ist z. B. die Voraussetzung der Unimodalität einer Funktion für die Konvergenz eines Verfahrens eine sehr gravierende Voraussetzung.

Zur Lösung des eindimensionalen Optimierungsproblems

$$\min_{t \in [0, \infty)} \varphi(t) := f(x + tp)$$

mit einer konvexen oder unimodalen Funktion  $\varphi$  kann man auch eine Nullstelle von  $\varphi'$  z. B. mit einem der aus der Numerischen Mathematik bekannten Verfahren, wie z. B. der *Regula Falsi* oder dem *Sekanten-Verfahren*, bestimmen, welche

nur Funktionswerte von  $\varphi'$  und damit nur den Gradienten von  $f$  verwenden. Das Newton-Verfahren benötigt dagegen  $\varphi''$  und damit im Hinblick darauf, dass  $\varphi$  nur eine Funktion in einer Veränderlichen ist, numerisch häufig zu teure Auswertungen der Hesse-Matrix von  $f$ . Im nichtkonvexen Fall muss man aber bei solchen Vorgehensweisen noch sicherstellen, dass es sich bei der gefundenen Nullstelle tatsächlich um einen globalen Minimierer von  $\varphi$  handelt.

Für die Berechnung der Curry-Schrittweite, also der kleinsten positiven Nullstelle von  $\varphi'(t) = \nabla f(x+tp)^T p$ , müssen keine speziellen Forderungen an  $f$  gestellt werden. Man beginnt normalerweise mit einer Einschachtelungsprozedur, bei der man mittels eines Vergleichs von Funktionswerten von  $\varphi'$  versucht, ein hinreichend kleines Intervall zu finden, in dem sich die gesuchte Nullstelle befindet. Anschließend kann man jedes Verfahren zur Bestimmung einer Nullstelle einer Funktion in einer Veränderlichen auf einem Intervall anwenden, wie z. B. eines der oben genannten.

Die Berechnung exakter Schrittweiten für eine allgemeine nichtlineare Funktion bedingt, dass zunächst eine bzw. die Lösung der entsprechenden eindimensionalen Aufgabe eingeschachtelt wird und dass das zu deren Bestimmung eingesetzte Verfahren mit ausreichender Geschwindigkeit konvergiert. Überdies können exakte Schrittweiten im Allgemeinen nur näherungsweise berechnet werden und muss darauf vertraut werden, dass das entsprechende Verfahren zur Lösung von Problem  $(P)$  auch mit diesen Näherungen konvergent ist. Aus all diesen Gründen verwendet man zumeist andere, leichter umzusetzende Schrittweitenregeln, für die auch Theorie und Praxis nicht auseinanderfallen. Einige solcher Regeln werden wir im Folgenden beschreiben.

## 4.2 Die Armijo-Schrittweite

Eine sehr populäre, da sehr leicht berechenbare, Schrittweite ist die folgende.

**Definition 4.11** Seien  $\eta \in (0, 1)$  und  $\zeta \in (0, 1/2)$  gegebene Zahlen. Ferner sei  $q := q(x, p)$  die kleinste Zahl aus  $\mathbb{N}_0$  derart, dass die Ungleichung

$$f(x) - f(x + \eta^q p) \geq -\zeta \eta^q \nabla f(x)^T p \quad (4.9)$$

erfüllt ist. Dann heißt  $t_A = t_A(x, p) := \eta^q$  Armijo-Schrittweite.

Die Armijo-Schrittweite ist also die größte Zahl aus der Menge

$$\{1, \eta, \eta^2, \eta^3, \eta^4, \dots\}, \quad (4.10)$$

für welche die Ungleichung in (4.9) erfüllt ist. Da  $1 \geq \eta \geq \eta^2 \geq \eta^3 \geq \dots$  gilt, muss man mit 1 beginnend und abnehmender Größe nur rechnerisch überprüfen, für welche Zahl  $\eta^q$  die Ungleichung (4.9) zum ersten Mal erfüllt ist. Die Berechnung der Armijo-Schrittweite ist also trivial.

**Beispiel 4.12** Sei  $f$  die quadratische Funktion

$$f(x) := \frac{1}{2} x^T Q x + c^T x + \alpha, \quad x \in \mathbb{R}^n,$$

mit positiv definiten Matrix  $Q$  und seien  $x$  und  $p$  Punkte wie in (4.1) gegeben. Im Hinblick auf die Ungleichung (4.9) stellen wir fest, dass

$$\begin{aligned} f(x) - f(x + tp) &= -t(Qx + c)^T p - \frac{1}{2} t^2 p^T Q p \\ &\geq -\zeta t \nabla f(x)^T p. \end{aligned}$$

genau dann gilt, wenn

$$t \leq -2(1 - \zeta) \frac{\nabla f(x)^T p}{p^T Q p} = 2(1 - \zeta) t_C$$

ist, wobei die Identität  $\nabla f(x) = Qx + c$  und die Curry-Schrittweite  $t_C$  für  $f$  aus (4.5) verwendet wurden. Die Armijo-Schrittweite  $t_A$  ist also die größte Zahl  $\eta^q$  aus der Menge (4.10), welche der Ungleichung

$$\eta^q \leq 2(1 - \zeta) t_C. \quad (4.11)$$

genügt. Für  $\zeta \in (0, 1/2)$  ist sie kleiner als  $2t_C$ .

Der leichten Berechenbarkeit der Armijo-Schrittweite steht entgegen, dass sie nur semieffizient ist.

**Satz 4.13** *Es seien (V1)–(V3) erfüllt und Zahlen  $\eta \in (0, 1)$  und  $\zeta \in (0, 1/2)$  gegeben. Für alle Paare  $x$  und  $p$  mit (4.1) existiert genau eine Armijo-Schrittweite, und diese ist eine semieffiziente Schrittweite mit der Konstanten*

$$\vartheta_A := \min \left( \zeta, \frac{2\eta\zeta(1-\zeta)}{\gamma} \right).$$

**Beweis.** Aus der Definition der Richtungsableitung erhält man für  $\zeta \in (0, 1/2)$

$$-\lim_{t \rightarrow 0^+} \frac{f(x+tp) - f(x)}{t} = -\nabla f(x)^T p > -\zeta \nabla f(x)^T p.$$

Folglich gilt für alle genügend kleinen  $t > 0$

$$f(x) - f(x+tp) > -\zeta t \nabla f(x)^T p.$$

Wegen  $\lim_{q \rightarrow \infty} \eta^q = 0$  kann demzufolge Ungleichung (4.9) für genügend großes  $q$  erfüllt werden, wobei  $q$  offenbar eindeutig ist. Also existiert die Armijo-Schrittweite. (Dies kann man sogar für  $\zeta \in (0, 1)$  schließen, aber z. B. für Satz 7.13 unten muss  $\zeta \in (0, 1/2)$  vorausgesetzt werden.)

Als Nächstes zeigen wir, dass die Armijo-Schrittweitenregel eine semieffiziente Schrittweitenregel ist. Ist  $q = 0$ , also  $t_A = 1$ , so folgt aus (4.9)

$$f(x) - f(x+t_A p) \geq -\zeta \nabla f(x)^T p. \quad (4.12)$$

Ist  $q > 0$ , so gilt zum einen natürlich für  $\eta^q = t_A$  die Ungleichung (4.9), und zum anderen weiß man, dass dann die Ungleichung (4.9) für  $\eta^{q-1} = \eta^{-1} t_A$  noch nicht erfüllt und daher Folgendes richtig ist:

$$f(x) - f(x + \eta^{-1} t_A p) < -\zeta \eta^{-1} t_A \nabla f(x)^T p. \quad (4.13)$$

Für  $\hat{t}$  aus Lemma 3.9 betrachten wir jetzt zuerst den Fall  $\eta^{-1} t_A \leq \hat{t}$ . Für diesen folgt mit (4.13) und Lemma 3.12

$$\begin{aligned} -\zeta \eta^{-1} t_A \nabla f(x)^T p &> f(x) - f(x + \eta^{-1} t_A p) \\ &\geq -\eta^{-1} t_A \nabla f(x)^T p - (\eta^{-1} t_A)^2 \frac{\gamma}{2} \|p\|^2. \end{aligned}$$

Division durch  $\eta^{-1} t_A$  und anschließendes Auflösen nach  $t_A$  liefert damit

$$t_A \geq -\frac{2\eta(1-\zeta)}{\gamma} \frac{\nabla f(x)^T p}{\|p\|^2} > 0.$$

Unter Verwendung von (4.9) erhalten wir also

$$f(x) - f(x + t_A p) \geq \frac{2\eta\zeta(1-\zeta)}{\gamma} \left\{ \frac{\nabla f(x)^T p}{\|p\|} \right\}^2. \quad (4.14)$$

Ist andererseits  $\eta^{-1}t_A \geq \hat{t}$ , so ergibt sich mit Lemma 3.12 und  $\tilde{t}$  von dort

$$t_A \geq \eta\hat{t} \geq \eta\tilde{t} = -\frac{2\eta}{\gamma} \frac{\nabla f(x)^T p}{\|p\|^2} > 0.$$

Mit (4.9) erhält man daher

$$f(x) - f(x + t_A p) \geq \frac{2\eta\zeta}{\gamma} \left\{ \frac{\nabla f(x)^T p}{\|p\|} \right\}^2. \quad (4.15)$$

Aus (4.12), (4.14) und (4.15) zusammen folgt das gewünschte Ergebnis. ■

**Aufgabe 4.14** Wählt man  $q := q(x, p)$  nur im Fall, dass die Armijo-Ungleichung (4.9) für  $q = 0$  nicht erfüllt ist, wie in Definition 4.11 und anderenfalls als die größte ganze Zahl  $q \leq 0$ , für welche die Armijo-Ungleichung (4.9) erfüllt, aber für  $q - 1$  nicht erfüllt ist, so bezeichnet man die entsprechende Schrittweite  $t_{AA} = t_{AA}(x, p) := \eta^q$  als Armijo-Schrittweite mit Aufweitung. (Es kann somit auch  $t_{AA} > 1$  gelten.) Man zeige unter Verwendung des Beweises von Satz 4.13, dass diese Schrittweite unter den Voraussetzungen des Satzes existiert und eindeutig ist und dass sie eine effiziente (!) Schrittweite ist mit der Konstanten

$$\vartheta_{AA} := \frac{2\zeta\eta(1-\zeta)}{\gamma}.$$

## 4.3 Wolfe-Powell-Schrittweiten

### 4.3.1 Definition und Effizienz

Im Fall der Schrittweitenregel von Wolfe und Powell muss neben einer Ungleichung vom Armijo-Typ wie in (4.9) eine weitere Ungleichung erfüllt werden.<sup>1</sup>

**Definition 4.15** Es seien Zahlen  $\tau \in (0, 1/2)$  und  $\sigma \in (\tau, 1)$  gegeben. Dann heißt jedes Element der Menge

$$T_{WP}(x, p) := \{t \in \mathbb{R}_+ \mid -\tau t \nabla f(x)^T p \leq f(x) - f(x + tp), \\ -\nabla f(x + tp)^T p \leq -\sigma \nabla f(x)^T p\}$$

Wolfe-Powell-Schrittweite.

<sup>1</sup>Wir folgen hier der Namensgebung z. B. in [Fle91] und [GeiKa99]. Man findet auch die Bezeichnungen Powell-Wolfe- ([Kos89]) und Powell-Schrittweitenregel ([Wer92], [Alt02]).

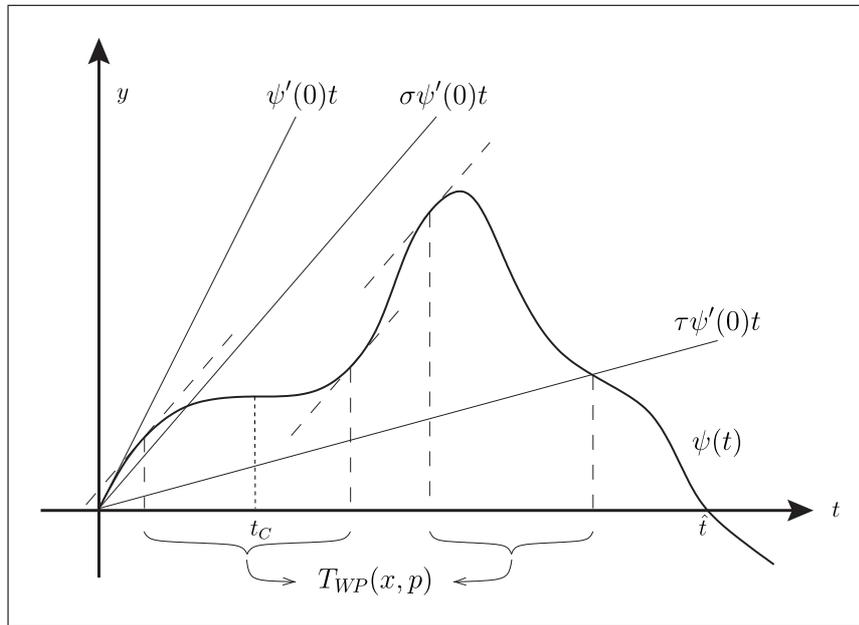


Abb. 4.2: Menge der Wolfe-Powell-Schrittweiten

Wir betrachten auch für diese Schrittweitenregel zuerst wieder das Beispiel einer gleichmäßig konvexen, quadratischen Funktion.

**Beispiel 4.16** *Es sei  $f$  die quadratische Funktion*

$$f(x) := \frac{1}{2}x^T Qx + c^T x + \alpha, \quad x \in \mathbb{R}^n, \quad (4.16)$$

wobei  $Q$  positiv definit sei. Weiter mögen  $x$  und  $p$  die Bedingungen in (4.1) erfüllen. Ähnlich wie in Beispiel 4.12 zeigt man unter Verwendung der Curry-Schrittweite  $t_C$  für  $f$ :

$$T_{WP}(x, p) = [(1 - \sigma)t_C, 2(1 - \tau)t_C]. \quad (4.17)$$

Aufgrund der Forderungen  $\tau \in (0, 1/2)$  und  $\sigma \in (\tau, 1)$  ist die Menge der Wolfe-Powell-Schrittweiten in diesem Fall ein ganzes Intervall, welches  $t_C$  in seinem Inneren enthält.

**Aufgabe 4.17** *Zeigen Sie die Gültigkeit der Beziehung (4.17) für quadratische Funktionen wie in (4.16) mit positiv definiter Matrix  $Q$ .*

Wir verwenden wieder die Funktion  $\psi$  mit

$$\psi(t) := f(x) - f(x + tp), \quad \psi'(t) = -\nabla f(x + tp)^T p, \quad \psi'(0) = -\nabla f(x)^T p > 0. \quad (4.18)$$

Damit entspricht die Menge  $T_{WP}(x, p)$  der Wolfe-Powell-Schrittweiten der Menge aller  $t \geq 0$ , welche den beiden Ungleichungen

$$\tau\psi'(0)t \leq \psi(t), \quad \psi'(t) \leq \sigma\psi'(0) \quad (4.19)$$

genügen (siehe Abb. 4.2). Wir zeigen als Nächstes unter den Standardvoraussetzungen, dass die Wolfe-Powell-Schrittweitenregel effizient ist und für jedes  $f$  die Wahl einer Schrittweite aus einem ganzen Intervall erlaubt.

**Satz 4.18** *Es seien (V1)–(V3) erfüllt. Für alle Paare  $x$  und  $p$  mit (4.1) und jedes  $\tau \in (0, 1/2)$  und  $\sigma \in (\tau, 1)$  enthält die Menge  $T_{WP}(x, p)$  mindestens ein nichtleeres, nicht einpunktiges abgeschlossenes Intervall. Ferner ist jede Wolfe-Powell-Schrittweite  $t_{WP} := t_{WP}(x, p)$  eine effiziente Schrittweite mit der Konstanten*

$$\vartheta_{WP} := \frac{1}{\gamma} \min \left( \frac{1 - \sigma^2}{2}, 2\tau(1 - \tau) \right). \quad (4.20)$$

**Beweis.** Seien  $x$ ,  $p$ ,  $\sigma$  und  $\tau$  gegeben wie angenommen und sei

$$d(t) := \psi(t) - \tau\psi'(0)t, \quad d'(t) = \psi'(t) - \tau\psi'(0).$$

Dann ist  $d(0) = 0$  und mit der Curry-Schrittweite  $t_C$

$$d'(0) = (1 - \tau)\psi'(0) > 0, \quad d'(t_C) = -\tau\psi'(0) < 0.$$

Folglich besitzt  $d$  in  $(0, t_C]$  einen lokalen Maximalpunkt  $t_m$  mit  $d(t_m) > 0$  und  $d'(t_m) = 0$ , wobei die letztere Bedingung wegen  $\psi'(0) > 0$  liefert, dass

$$\psi'(t_m) = \tau\psi'(0) < \sigma\psi'(0) \quad (4.21)$$

gilt. Da also  $d(t_m) > 0$  und  $\psi'(t_m) < \sigma\psi'(0)$  für  $t_m \in (0, t_C]$  ist, existiert weiter ein nichtleeres, nicht einpunktiges abgeschlossenes Intervall  $I \subseteq (0, t_C]$ , so dass folgt:

$$d(t) > 0, \quad \psi'(t) < \sigma\psi'(0), \quad t \in I. \quad (4.22)$$

Letztere Ungleichungen implizieren offenbar die Inklusion  $I \subseteq T_{WP}(x, p)$ .

Für den Nachweis, dass die Wolfe-Powell-Schrittweitenregel  $t_{WP}$  effizient ist, untersuchen wir zunächst den Fall

$$t_{WP} \leq -\frac{2(1 - \tau)}{\gamma} \frac{\nabla f(x)^T p}{\|p\|^2} := t_r. \quad (4.23)$$

Nach der Definition von  $t_{WP}$  gilt  $(x + t_{WP}p) \in N_0$  mit  $x \in N_0$  und

$$-\nabla f(x + t_{WP}p)^T p \leq -\sigma \nabla f(x)^T p = -\nabla f(x)^T p + (1 - \sigma) \nabla f(x)^T p.$$

Unter Anwendung von Voraussetzung (V3) können wir daraus schließen:

$$-(1 - \sigma) \nabla f(x)^T p \leq [\nabla f(x + t_{WP} p) - \nabla f(x)]^T p \leq \gamma t_{WP} \|p\|^2.$$

Daher erhalten wir mit (4.23) und  $\hat{t}$  aus Lemma 3.12

$$0 < t_l := -\frac{1 - \sigma}{\gamma} \frac{\nabla f(x)^T p}{\|p\|^2} \leq t_{WP} \leq t_r \leq \hat{t}.$$

Nun verwenden wir, dass eine nach unten geöffnete Parabel wie die Parabel  $\Psi$  aus Lemma 3.12 ihr Minimum auf dem Intervall  $[t_l, t_r]$  in  $t_l$  oder  $t_r$  annimmt. Folglich liefert Anwendung von Lemma 3.12

$$\begin{aligned} f(x) - f(x + t_{WP} p) &\geq -t_{WP} \nabla f(x)^T p - t_{WP}^2 \frac{\gamma}{2} \|p\|^2 \\ &\geq \min_{t \in [t_l, t_r]} \left( -t \nabla f(x)^T p - t^2 \frac{\gamma}{2} \|p\|^2 \right) \\ &\geq \frac{1}{\gamma} \min \left( \frac{1 - \sigma^2}{2}, 2\tau(1 - \tau) \right) \left\{ \frac{\nabla f(x)^T p}{\|p\|} \right\}^2. \end{aligned}$$

Ist andererseits

$$t_{WP} \geq -\frac{2(1 - \tau)}{\gamma} \frac{\nabla f(x)^T p}{\|p\|^2} (> 0),$$

dann folgt direkt mit der Definition von  $t_{WP}$

$$f(x) - f(x + t_{WP} p) \geq -\tau t_{WP} \nabla f(x)^T p \geq \frac{2\tau(1 - \tau)}{\gamma} \left\{ \frac{\nabla f(x)^T p}{\|p\|} \right\}^2.$$

Fassen wir die in beiden Fällen gewonnenen Ungleichungen zusammen, so erhalten wir eine Ungleichung des Typs (3.25) mit der Konstante aus (4.20). ■

### 4.3.2 Numerische Berechnung

Die Wolfe-Powell-Schrittweitenregel ist auch numerisch realisierbar, da in ihrem Fall nicht wie bei den exakten Schrittweitenregeln ein einzelner Punkt, sondern gemäß Satz 4.18 nur ein  $t$  aus einem Intervall gefunden werden muss. Insbesondere kann eine Wolfe-Powell-Schrittweite in endlich vielen Schritten mit folgendem Algorithmus berechnet werden, wie der daran anschließende Satz beweist.

Die Idee dabei ist es, zunächst ein Intervall zu bestimmen, dessen linker Randpunkt die Armijo-Ungleichung, also die erste Ungleichung in  $T_{WP}(x, p)$  erfüllt und dessen rechter dies nicht tut. Wenn der linke Randpunkt des Intervalls auch der zweiten Ungleichung in  $T_{WP}(x, p)$  genügt, ist man fertig. Anderenfalls wird die Länge des Intervalls, ähnlich wie beim Bisektionsverfahren zur Bestimmung einer Nullstelle einer reellwertigen Funktion, so lange unter Beibehaltung der mit dem Intervall verbundenen Eigenschaften halbiert, bis der linke Randpunkt auch die zweite Ungleichung erfüllt.

**Algorithmus 4.19**

(0) Gib  $x \in N_0$  und  $p \in \mathbb{R}^n$  mit  $\nabla f(x)^T p < 0$ ,  $\tau \in (0, 1/2)$  und  $\sigma \in (\tau, 1)$  und setze  $k := 0$ .

(1) (i) Falls für  $t := 1$  die Armijo-Ungleichung

$$f(x) - f(x + tp) \geq -\tau t \nabla f(x)^T p \quad (4.24)$$

erfüllt ist, bestimme die kleinste Zahl  $b_k \in \{2, 2^2, 2^3, \dots\}$ , so dass (4.24) für  $t := b_k$  verletzt ist, und setze  $a_k := b_k/2$ .

(ii) Anderenfalls bestimme die größte Zahl  $a_k \in \{2^{-1}, 2^{-2}, 2^{-3}, \dots\}$ , so dass (4.24) für  $t := a_k$  erfüllt ist, und setze  $b_k := 2a_k$ .

(2) Falls für  $t := a_k$  die Ungleichung

$$-\nabla f(x + tp)^T p \leq -\sigma \nabla f(x)^T p \quad (4.25)$$

gilt, setze  $t_{WP} := a_k$ . *Stop!*

(3) Berechne

$$t_k := \frac{a_k + b_k}{2}.$$

Falls  $t_k$  die Bedingung (4.24) erfüllt, setze

$$a_{k+1} := t_k, \quad b_{k+1} := b_k.$$

Anderenfalls setze

$$a_{k+1} := a_k, \quad b_{k+1} := t_k.$$

(4) Setze  $k := k + 1$  und gehe nach (2)

**Satz 4.20** *Es seien (V1)–(V3) erfüllt. Dann bricht Algorithmus 4.19 nach endlich vielen Iterationen mit einer Wolfe-Powell-Schrittweite  $t_{WP} \in T_{WP}(x, p)$  ab.*

**Beweis.** Nach Lemma 3.9 ist  $f(x) < f(x + tp)$ ,  $t \geq \kappa$ , für ein  $\kappa > 0$ . Somit wird in Schritt (1) (i) nach endlich vielen Schritten ein  $t := b_0$ , wie angegeben, gefunden. Gemäß Satz 4.13 für  $\eta := 1/2$  und  $\zeta := \tau$  kann weiter in (ii) nach endlich vielen Schritten die Armijo-Schrittweite  $a_0$  bestimmt werden. Am Ende von Schritt (1) hat man somit in beiden Fällen für  $k = 0$

$$a_k < b_k, \quad t := a_k \text{ erfüllt (4.24),} \quad t := b_k \text{ erfüllt (4.24) nicht.} \quad (4.26)$$

Ist auch (4.25) für  $t := a_k$  richtig, so ist offenbar  $a_k \in T_{WP}(x, p)$  und bricht das Verfahren in (2) erfolgreich ab. Anderenfalls hat man zu Beginn von Schritt (3) die Situation in (4.26) und wird nun die Länge des Intervalls  $[a_k, b_k]$  in jedem Durchlauf von Schritt (3) halbiert, wobei entweder  $a_k$  vergrößert oder  $b_k$  verkleinert wird und  $a_k$  und  $b_k$  die Eigenschaften in (4.26) bewahren.

Würde Schritt (3) unendlich oft durchlaufen, so konvergierten die Folgen  $\{a_k\}$  und  $\{b_k\}$  aufgrund ihrer sich aus

$$a_0 \leq a_k \leq a_{k+1} \leq b_{k+1} \leq b_k \leq b_0$$

ergebenden Monotonie und der Tatsache, dass die Länge der Intervalle  $[a_k, b_k]$  gegen 0 geht, gegen dieselbe Zahl  $t^*$ . Aufgrund von (4.26) wäre dann weiter die Armijo-Ungleichung in (4.24) für  $t^*$  gleichzeitig erfüllt und „verletzt“, d. h., es wäre  $d(t^*) = 0$  für

$$d(t) := f(x) - f(x + tp) + \tau t \nabla f(x)^T p.$$

Es folgte außerdem  $d'(t^*) \leq 0$ , da anderenfalls

$$\frac{d(b_k) - d(t^*)}{b_k - t^*} > 0$$

und damit  $d(b_k) > 0$  für alle hinreichend großen  $k$  gelten würde, was aber im Widerspruch dazu stünde, dass  $b_k$  die Bedingung (4.24) verletzt. Wegen  $d'(t^*) \leq 0$  und  $\tau < \sigma$  hätte man dann jedoch

$$-\nabla f(x + t^*p)^T p \leq -\tau \nabla f(x)^T p < -\sigma \nabla f(x)^T p$$

und damit für alle hinreichend großen  $k$

$$-\nabla f(x + a_k p)^T p < -\sigma \nabla f(x)^T p.$$

Also wäre die Ungleichung (4.25) für  $t := a_k$  mit hinreichend großem  $k$  erfüllt, was aber der Annahme widerspricht, dass (3) unendlich oft durchlaufen wird. Demzufolge bricht der Algorithmus nach endlich vielen Durchgängen von Schritt (3) ab. ■

**Aufgabe 4.21** Gegeben sei die Funktion  $\varphi(t) := f(x + tp)$  mit<sup>2</sup>

$$\varphi(t) := 2.56t^2(6 - t)^2 + 0.04(1 - 2t)^2.$$

(a) Man ermittle die Armijo-Schrittweite aus Definition 4.11 für  $\zeta := 0.05$  und  $\eta = 0.25$ .

(b) Man berechne mit Algorithmus 4.19 eine Wolfe-Powell-Schrittweite für  $\tau := 0.45$  und  $\sigma = 0.5$ .

---

<sup>2</sup>Man erhält  $\varphi$  hier mit der Rosenbrock-Funktion  $f(x) := 100(x_2 - x_1^2)^2 + (1 - x_1)^2$  und  $p := -\nabla f(x)$  für  $x := (1.2, 1.44)^T$ . Die Theorie für das Gradientenverfahren in Kapitel 5 lässt in diesem Fall nur einen geringen Abstieg hinsichtlich des Funktionswertes und somit eine ungewöhnlich kleine Schrittweite erwarten. (Letzteres macht die Aufgabe interessanter.)

## 4.4 Strenge Wolfe-Powell-Schrittweiten

In einigen Zusammenhängen hat sich die folgende Modifikation der Wolfe-Powell-Schrittweitenregel in der Praxis bewährt.

**Definition 4.22** *Es seien Zahlen  $\tau \in (0, 1/2)$  und  $\sigma \in (\tau, 1)$  gegeben. Dann heißt jedes Element der Menge*

$$T_{SWP}(x, p) := \{t \in \mathbb{R}_+ \mid -\tau t \nabla f(x)^T p \leq f(x) - f(x + tp), \quad (4.27)$$

$$|\nabla f(x + tp)^T p| \leq -\sigma \nabla f(x)^T p\} \quad (4.28)$$

strenge Wolfe-Powell-Schrittweite.

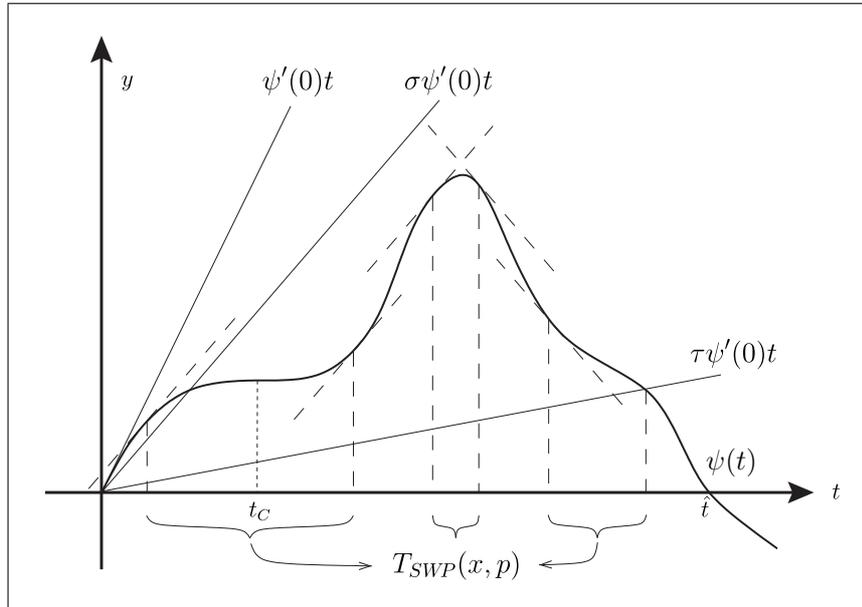


Abb. 4.3: Menge der strengen Wolfe-Powell-Schrittweiten

Zunächst betrachten wir auch hier wieder quadratische Funktionen:

**Beispiel 4.23** *Es sei  $f$  die quadratische Funktion*

$$f(x) := \frac{1}{2}x^T Qx + c^T x + \alpha, \quad x \in \mathbb{R}^n,$$

mit positiv definiten Matrix  $Q$ , und  $x$  und  $p$  mögen die Bedingungen in (4.1) erfüllen. Mit einer einfachen Rechnung ähnlich wie in Beispiel 4.12 zeigt man

$$T_{SWP}(x, p) = [(1 - \sigma) t_C, \min\{1 + \sigma, 2(1 - \tau)\} t_C], \quad (4.29)$$

wobei  $t_C$  wieder die Curry-Schrittweite für  $f$  ist. Für  $\tau \in (0, 1/2)$  und  $\sigma \in (\tau, 1)$  folgt also  $t_C \in T_{SWP}(x, p)$ .

Mit der Funktion  $\psi$  aus (4.18) ist die Menge  $T_{SWP}(x, p)$  der strengen Wolfe-Powell-Schrittweiten gerade die Menge aller  $t \geq 0$ , welche gleichzeitig die Ungleichungen

$$\tau\psi'(0)t \leq \psi(t), \quad |\psi'(t)| \leq \sigma\psi'(0) \quad (4.30)$$

erfüllen. Gegenüber der Wolfe-Powell-Schrittweitenregel werden bei der strengen Wolfe-Powell-Schrittweitenregel solche Schrittweiten ausgeschlossen, für welche die Steigung von  $\psi$  einen zu großen negativen Wert hat (vgl. Abb. 4.3).

Für die strenge Wolfe-Powell-Schrittweitenregel gilt ähnlich wie für die einfache:

**Satz 4.24** *Es seien (V1)–(V3) erfüllt. Für alle Paare  $x$  und  $p$  mit (4.1) und Zahlen  $\tau \in (0, 1/2)$  und  $\sigma \in (\tau, 1)$  enthält die Menge  $T_{SWP}(x, p)$  mindestens ein nichtleeres abgeschlossenes Intervall. Ferner ist jede strenge Wolfe-Powell-Schrittweite  $t_{SWP} := t_{SWP}(x, p)$  eine effiziente Schrittweite mit der Konstanten*

$$\vartheta_{SWP} := \frac{1}{\gamma} \min \left( \frac{1 - \sigma^2}{2}, 2\tau(1 - \tau) \right).$$

**Beweis.** Für den Nachweis der Existenz eines nichtleeren abgeschlossenen Intervalls  $J \subseteq T_{SWP}(x, p)$  kann man zunächst dem Beweis von Satz 4.18 folgen, wobei man zusätzlich die folgenden Beziehungen verwende:

$$-\tau\psi'(0) < 0 = d'(t_m) = \psi'(t_m) - \tau\psi'(0).$$

Für  $t_m \in (0, t_C]$  ergibt sich damit

$$d(t_m) > 0, \quad 0 < \psi'(t_m) < \sigma\psi'(0).$$

Dies wiederum impliziert die Existenz eines nichtleeren abgeschlossenen Intervalls  $J \subseteq (0, t_C]$ , so dass gilt:

$$d(t) > 0, \quad 0 < \psi'(t) < \sigma\psi'(0), \quad t \in J.$$

Damit folgt auch  $-\psi'(t) < \sigma\psi'(0)$  für alle  $t \in J$  und demnach  $J \subseteq T_{SWP}(x, p)$ . Der zweite Teil der Behauptung erschließt sich nun wegen  $T_{SWP}(x, p) \subseteq T_{WP}(x, p)$  aus Satz 4.18. ■

Ein Algorithmus zur Berechnung einer strengen Wolfe-Powell-Schrittweite ist in [GeiKa99] zu finden.

Die Ungleichungen in (4.30) zeigen, dass man für festes  $x$  und  $p$  durch die Wahl von hinreichend kleinen Zahlen  $\tau$  und  $\sigma$  eine beliebig kleine Umgebung eines isolierten kritischen Punktes von  $\psi(t)$  bzw.  $\varphi(t) := f(x + tp)$  erzeugen kann. Startet man also den Algorithmus zur Berechnung einer strengen

Wolfe-Powell-Schrittweite nahe genug bei der Curry-Schrittweite  $t_C$  (man verwende dazu ein Einschachtelungsverfahren), so erlaubt die strenge Wolfe-Powell-Schrittweitenregel die Wahl einer Schrittweite, welche der Curry-Schrittweite beliebig nahe kommt. Beispielsweise liefert (4.29) im gleichmäßig konvexen, quadratischen Fall für  $\tau = 0.05$  und  $\sigma = 0.1$  das Intervall  $T_{SWP}(x, p) = [0.9t_C, 1.1t_C]$ , während man für die Wolfe-Powell-Schrittweitenregel gemäß (4.17) das viel größere Intervall  $T_{WP}(x, p) = [0.9t_C, 1.9t_C]$  erhält. Es liegt somit nahe, die strenge Wolfe-Powell-Schrittweitenregel zu verwenden, wenn ein Verfahren relativ empfindlich auf starke Abweichungen von den exakten Schrittweiten reagiert. Wir verweisen diesbezüglich z. B. auf die Bemerkungen zu CG-Verfahren in Abschnitt 6.5.



## 4.5 Lösungen der Aufgaben

Aufgabe 4.10 (a) Sei  $c > (b+a)/2$ , also  $[a, c]$  das längere und  $[c, b]$  das kürzere der beiden Intervalle. Wir setzen  $c := a + F(b-a)$  und wollen  $F \in (1/2, 1]$  ermitteln. Bei einer Zerlegung von  $[a, b]$  nach dem Goldenen Schnitt muss gelten:

$$\frac{b-a}{c-a} = \frac{c-a}{b-c} \Leftrightarrow \frac{1}{F} = \frac{F}{1-F} \Leftrightarrow F^2 + F - 1 = 0.$$

Die beiden Lösungen der letzten Gleichung lauten

$$F_{1,2} = -\frac{1}{2} \pm \frac{1}{2}\sqrt{5}.$$

Da wir  $F \in (1/2, 1]$  suchen, erhalten wir

$$F := \frac{\sqrt{5}-1}{2} \approx 0.618\,033\,989.$$

Ist  $[a, c]$  das kürzere der beiden Intervalle, so muss aus Symmetriegründen (4.8) richtig sein.

(b) Die erste Forderung impliziert, dass  $s_k$  und  $t_k$  symmetrisch zum Mittelpunkt von  $[a_k, b_k]$  liegen müssen, so dass mit einem  $\mathcal{F} \in (1/2, 1)$  gilt:

$$s_k := a_k + (1-\mathcal{F})(b_k - a_k), \quad t_k := a_k + \mathcal{F}(b_k - a_k). \quad (4.31)$$

Daraus ergibt sich die Länge des neuen Intervalls  $[a_{k+1}, b_{k+1}]$  mit

$$b_{k+1} - a_{k+1} = b_k - s_k = t_k - a_k = \mathcal{F}(b_k - a_k). \quad (4.32)$$

Nun sei zunächst  $\varphi(s_k) > \varphi(t_k)$  und somit  $[a_{k+1}, b_{k+1}] = [s_k, b_k]$ . Dann schließt man mit (4.31) und (4.32)

$$\begin{aligned} s_{k+1} &= a_{k+1} + (1-\mathcal{F})(b_{k+1} - a_{k+1}) \\ &= s_k + (1-\mathcal{F})\mathcal{F}(b_k - a_k) \\ &= a_k + (1-\mathcal{F})(b_k - a_k) + (1-\mathcal{F})\mathcal{F}(b_k - a_k) \\ &= a_k + (1-\mathcal{F}^2)(b_k - a_k). \end{aligned} \quad (4.33)$$

Die zweite Forderung kann erfüllt werden, wenn man  $s_{k+1} = t_k$  wählen kann, so dass im neuen Schritt nur eine neue Funktionsauswertung, nämlich  $\varphi(t_{k+1})$ , erforderlich ist. Nach (4.31) und (4.33) gilt dies genau dann, wenn man hat:

$$1 - \mathcal{F}^2 = \mathcal{F} \Leftrightarrow \mathcal{F}^2 + \mathcal{F} - 1 = 0.$$

Also ist

$$\mathcal{F} = \frac{\sqrt{5}-1}{2} \approx 0.618\,033\,989$$

die in diesem Fall einzig sinnvolle Lösung. Analog verfährt man im Fall  $\varphi(s_k) \leq \varphi(t_k)$ .

(c) Für die Länge des durch Algorithmus 4.9 bestimmten Intervalls  $[a_k, b_k]$  gilt offenbar

$$b_k - a_k = F(b_{k-1} - a_{k-1}) = F^k(b_0 - a_0) = F^k(b - a).$$

Folglich hat man  $\lim_{k \rightarrow \infty} (b_k - a_k) = 0$  und existiert demnach ein  $k_0$ , so dass  $b_k - a_k \leq \varepsilon$  für alle  $k \geq k_0$  folgt. Also bricht Algorithmus 4.9 spätestens mit  $k := k_0$  ab.

(d) Es ist  $b_k - a_k \leq \varepsilon$  genau dann erfüllt, wenn gilt:

$$\begin{aligned} F^k(b - a) \leq \varepsilon &\Leftrightarrow (0 <) \frac{b - a}{\varepsilon} \leq \left(\frac{1}{F}\right)^k \\ &\Leftrightarrow k \geq \frac{\ln(b - a) - \ln(\varepsilon)}{\ln(2) - \ln(\sqrt{5} - 1)}. \end{aligned}$$

Für  $[a, b] := [-1, 1]$  und  $\varepsilon := 0.000\,001$  errechnet man

$$k \geq \frac{\ln(2) + 6 \ln(10)}{\ln(2) - \ln(\sqrt{5} - 1)} \approx 30.15.$$

Folglich bricht Algorithmus 4.9 in diesem Fall bei  $k = 31$  ab. Er hat bis dahin 3 Funktionswertungen in der 0-ten und jeweils 1 Funktionsauswertung in den Iterationen 1 bis 30, also insgesamt 33 Funktionsauswertungen benötigt.

Aufgabe 4.14 Im Beweis von Satz 4.13 wurde anfangs gezeigt, dass die Ungleichung (4.9) immer für genügend großes  $q \in \mathbb{N}_0$  richtig ist und es somit ein kleinstes  $q \in \mathbb{N}_0$  gibt, für welches sie gilt. Es ist nun auszuschließen, dass sie nicht beliebig klein werden, d. h. die Ungleichung (4.9) nicht für jedes  $q \in \mathbb{Z}$  mit  $q < 0$  erfüllt sein kann. Dazu verwende man, dass nach Lemma 3.9 ein  $\kappa > 0$  existiert, so dass

$$x + tp \notin N_0, \quad f(x) < f(x + tp), \quad t \geq \kappa,$$

gilt, und dass man somit wegen  $0 < \eta < 1$  für alle hinreichend kleinen  $q < 0$  hat:

$$f(x) \leq f(x^0) < f(x + \eta^q p).$$

Für solche  $q$  ist die Ungleichung (4.9) nicht erfüllt, da für sie der Ausdruck auf der linken Seite von (4.9) eine negative und der auf ihrer rechten Seite eine positive

Zahl ist. Demzufolge existiert im Fall, dass die Ungleichung (4.9) für  $q = 0$  erfüllt ist, auch eine größte ganze Zahl  $q \leq 0$ , so dass sie für  $q$  erfüllt und für  $q - 1$  nicht erfüllt ist.

Für den Nachweis der Effizienz der aufgeweiteten Armijo-Schrittweite kann man nun dem Beweis der Semieffizienz der gewöhnlichen Armijo-Schrittweite folgen, wobei man den Fall  $q = 0$  nicht gesondert betrachten muss. (Dieser „verhindert“ ja durch den Term auf der rechten Seite in (4.12) eine Ungleichung vom Typ (3.25).) Streicht man also den Teil im Beweis von „Ist  $q = 0$ “ bis einschließlich „Ist  $q > 0$ “ nach Formel (4.12), so bleibt der Beweis auch für die aufgeweitete Armijo-Schrittweite richtig und folgt ihre Effizienz mit der angegebenen Konstante aus den Formeln (4.14) und (4.15).

Aufgabe 4.17 Es sei  $Q$  symmetrisch und positiv definit und

$$f(x) := \frac{1}{2}x^T Qx + c^T x + \alpha, \quad x \in \mathbb{R}^n.$$

Weiter mögen  $x$  und  $p$  die Bedingungen in (4.1) erfüllen. Die Ungleichung

$$\begin{aligned} -\tau t \nabla f(x)^T p &\leq f(x) - f(x + tp) = -t(Qx + c)^T p - \frac{1}{2}t^2 p^T Qp \\ &= -t \nabla f(x)^T p - \frac{1}{2}t^2 p^T Qp \end{aligned}$$

gilt genau dann, wenn

$$t \leq -2(1 - \tau) \frac{\nabla f(x)^T p}{p^T Qp} = 2(1 - \tau) t_C$$

ist, wobei (4.5) verwendet wurde. Weiter ist die Ungleichung (vgl. (4.4))

$$-\nabla f(x + tp)^T p = -\nabla f(x)^T p - t p^T Qp \leq -\sigma \nabla f(x)^T p$$

genau dann richtig, wenn gilt:

$$-(1 - \sigma) \frac{\nabla f(x)^T p}{p^T Qp} = (1 - \sigma) t_C \leq t.$$

Damit folgt die Darstellung von  $T_{WP}(x, p)$  in (4.17).

Aufgabe 4.21 (a) Es ist

$$\begin{aligned} \varphi(t) &:= f(x + tp) = 2.56 t^2 (6 - t)^2 + 0.04 (1 - 2t)^2, \\ \varphi'(t) &:= \nabla f(x + tp)^T p = 5.12 t (6 - t) (6 - 2t) - 0.16 (1 - 2t), \\ \psi(t) &:= \varphi(0) - \varphi(t) = 0.04 - \varphi(t), \\ \psi'(0) &= -\varphi'(0) = -\nabla f(x)^T p = 0.16. \end{aligned}$$

Damit bekommt man

$$\begin{aligned} d(t) &:= f(x) - f(x + tp) + \zeta t \nabla f(x)^T p \\ &= \psi(t) - \zeta t \psi'(0) \\ &= 0.04 - 2.56 t^2 (6 - t)^2 - 0.04 (1 - 2t)^2 - 0.16 \zeta t. \end{aligned}$$

Gesucht ist das größte  $t \in \{1, \eta, \eta^2, \eta^3, \dots\}$ , für welches  $d(t) \geq 0$  ist. Für  $\zeta := 0.05$  und  $\eta := 0.25$  berechnet man die Werte in Tabelle 4.1. (Es interessiert eigentlich nur das Vorzeichen von  $d(t)$  und nicht der genaue Wert.)

$q$	$t := 1/4^q$	$d(t)$
0	1	-64.008
1	0.25	-5.262
2	0.0625	-0.343 664
3	0.015 625	-0.020 047
4	0.003 906 25	-0.000 813 111
5	0.000 976 563	+0.000 060 423

Tabelle 4.1

Die gesuchte Schrittweite lautet somit  $t_A = 4^{-5} = 0.000\,976\,563$ .

(b) Für  $\tau = \zeta := 0.45$  bedeutet (4.24), dass  $d(t) \geq 0$  ist. Weiter hat man

$$\begin{aligned} d'(t) &= \psi'(t) - \zeta \psi'(0) = -\nabla f(x + tp)^T p + \zeta \nabla f(x)^T p \\ &= -\varphi'(t) - \zeta \psi'(0) \\ &= -5.12t(6 - t)(6 - 2t) + 0.16(1 - 2t) - 0.16\zeta. \end{aligned}$$

Für  $\sigma = \zeta := 0.5$  ist (4.25) gleichbedeutend mit  $d'(t) \leq 0$ . Wie in Teil (a) ermittelt man für  $\eta := 0.5$  die größte Zahl  $a_0 \in \{1, \eta, \eta^2, \eta^3, \dots\}$ , für welche  $d(t) \geq 0$  ist. Es ergibt sich in Schritt (1) (ii) von Algorithmus 4.19

$$a_0 = \left(\frac{1}{2}\right)^{11} = 0.000\,488\,281\,25, \quad b_0 = \left(\frac{1}{2}\right)^{10} = 0.000\,976\,562\,5$$

mit  $d(a_0) = 0.000\,020\,962 \geq 0$ . Beim ersten Durchlauf von Schritt (2) berechnet man

$$d'(a_0) = -0.010\,134\,279 \leq 0,$$

so dass  $t_{WP} := a_0 = 0.000\,488\,281\,25$  die berechnete Wolfe-Powell-Schrittweite ist.