

Kapitel 1.

Einführung

Die nichtlineare Optimierung ist ein bedeutendes Gebiet der Mathematik. Sie findet immer wieder Anwendungen in den schwierigen Problemen der Technik und der Wirtschaft. Es wurden viele Verfahren entwickelt, um nichtlineare Optimierungsprobleme zu lösen. In dieser Arbeit werden zwei Verfahren, das halbglatte Newton-Verfahren und das SQP-Verfahren, betrachtet und verglichen.

Das SQP-Verfahren gehört zu den bekanntesten Verfahren der nichtlinearen Optimierung. Es wurde schon seit den 60er Jahren entwickelt und wurde in vielen Optimierungsproblemen als Standardwerkzeug angewendet sowie weiterentwickelt. Das halbglatte Newton-Verfahren ist weniger bekannt als das SQP-Verfahren. Es basiert aber auf das bekannte Newton-Verfahren.

Bevor wir die beiden Verfahren näher betrachten, werden wir erst mal in diesem Kapitel die grundlegend Definitionen sowie die wichtigen Ergebnisse betrachten.

1.1. Allgemeine Optimierungsprobleme

Definition 1.1 Allgemein ist die Aufgabenstellung der Optimierung wie folgt definiert:

$$\min_{x \in \mathcal{F}} f(x) \tag{1.1}$$

Die Funktion $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ist die sogenannte Zielfunktion. D sei der Definitionsbereich von f . \mathcal{F} sei eine nichtleere Teilmenge von D , die man als Lösungsmenge bezeichnet. Alle Elemente von \mathcal{F} werden als zulässige Punkte bezeichnet. \mathcal{F} wird durch die sogenannten Nebenbedingungen definiert.

Ein einfaches Beispiel ist das Problem

$$\min_{x \in \mathbb{R}} (x - 1)^2.$$

Falls $\mathcal{F} = D$ gilt, dann bezeichnet man das Optimierungsproblem als unrestringiert. Es besitzt also keine Nebenbedingungen. Ansonsten heißt es ein restringiertes Optimierungsproblem.

Es ist üblich, dass man nur Minimierungsproblem betrachtet, weil ein Maximierungsproblem $\max g(x)$ zu dem Minimierungsproblem $\min f(x) := -g(x)$ äquivalent ist.

Definition 1.2 (Globale und lokale Lösung)

Ein Punkt $x^* \in \mathcal{F}$ heißt globale Lösung des Problems (1.1) oder globales Minimum von f , wenn

$$f(x^*) \leq f(x) \quad \forall x \in \mathcal{F} \quad (1.2)$$

gilt. Ein Punkt $x^* \in \mathcal{F}$ heißt strikte globale Lösung des Problems (1.1) oder striktes globales Minimum von f , wenn

$$f(x^*) < f(x) \quad \forall x \in \mathcal{F} \setminus \{x^*\} \quad (1.3)$$

gilt. Ein Punkt $x^* \in \mathcal{F}$ heißt lokale Lösung des Problems (1.1) oder lokales Minimum von f , wenn für eine Umgebung $U(x^*)$ von x^*

$$f(x^*) \leq f(x) \quad \forall x \in U(x^*) \cap \mathcal{F} \quad (1.4)$$

gilt. Ein Punkt $x^* \in \mathcal{F}$ heißt strikte lokale Lösung des Problems (1.1) oder striktes lokales Minimum von f , wenn für eine Umgebung $U(x^*)$ von x^*

$$f(x^*) < f(x) \quad \forall x \in U(x^*) \cap \mathcal{F} \setminus \{x^*\} \quad (1.5)$$

gilt. Eine Umgebung $U(x^*)$ von x^* ist einfach eine offene Menge, die x^* beinhaltet.

Wegen dieser Definitionen kommt die Unterscheidung zwischen der globalen und der lokalen Optimierung. Globale Optimierung versucht, globale Lösungen zu finden. Lokale Optimierung versucht dagegen, lokale Lösungen zu finden. Viele Verfahren finden lokale Lösungen, die jedoch nicht unbedingt globale Lösungen sind. Globale Lösungen sind nicht so einfach zu finden.

Definition 1.3 Das allgemeine nichtlineare Optimierungsproblem ist wie folgt definiert:

$$\min f(x) \quad (1.6)$$

$$\text{Nb. } g(x) \leq 0 \quad (1.7)$$

$$h(x) = 0 \quad (1.8)$$

Die Zielfunktion ist wieder die Funktion $f : D_f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$. Die Nebenbedingungen sind von den Funktionen $g : D_g \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^p$ und $h : D_h \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ abhängig.

D.h. die Menge \mathcal{F} sieht hier so aus:

$$\mathcal{F} := \{x | g(x) \leq 0, h(x) = 0\}. \quad (1.9)$$

Man kann hierbei den Unterschied zwischen der linearen Optimierung und der nichtlinearen Optimierung gut erkennen. Bei der linearen Optimierung muss die Zielfunktion linear sein (d.h. die Zielfunktion besitzt die Form $f(x) = c^T x$, $c \in \mathbb{R}^n$) und die Nebenbedingungen sind durch lineare Gleichungssysteme oder Ungleichungssysteme definiert. Bei der nichtlinearen Optimierung gibt es dagegen keine Einschränkung, wie die Zielfunktion und die Nebenbedingungen aussehen sollen.

1.2. Unrestringierte Optimierungsprobleme

Wir betrachten nun das unrestringierte Optimierungsproblem

$$\min_{x \in \mathbb{R}^n} f(x). \quad (1.10)$$

Wir nehmen hier der Einfachheit halber an, dass der Definitionsbereich D gleich \mathbb{R}^n sei.

Satz 1.1 (Notwendige Bedingung erster Ordnung)

Seien x^ eine lokale Lösung des Problems (1.10) und f einmal stetig differenzierbar in einer Umgebung von x^* , dann gilt*

$$\nabla f(x^*) = 0 \quad (1.11)$$

Diese Bedingung gilt aber nicht nur für lokales Minimum sondern auch für lokales Maximum von f .

Definition 1.4 (Stationärer Punkt)

f sei in x^ differenzierbar. x^* heißt ein stationärer Punkt von f , wenn x^* die notwendige Bedingung (1.11) erfüllt.*

Viele Optimierungsverfahren suchen normalerweise nach einem stationärem Punkt von f . Aber ein stationärer Punkt muss nicht ein globales oder lokales Minimum sein.

Satz 1.2 (Notwendige Bedingung zweiter Ordnung)

Seien x^* eine lokale Lösung des Problems (1.10) und f zweimal stetig differenzierbar in einer Umgebung von x^* , dann gilt (1.11) und

$$x^T f''(x^*)x \geq 0 \quad \forall x \in \mathbb{R}^n, \quad (1.12)$$

$f''(x^*)$ ist also positiv semidefinit.

Durch diese notwendige Bedingung können wir zwischen einem lokalen Minimum und Maximum unterscheiden.

Satz 1.3 (Hinreichende Bedingung zweiter Ordnung)

Sei f zweimal stetig differenzierbar in einer Umgebung von x^* . Die notwendige Bedingung (1.11) sei erfüllt und $f''(x^*)$ sei positiv definit, d.h.

$$x^T f''(x^*)x > 0 \quad \forall x \in \mathbb{R}^n. \quad (1.13)$$

Dann ist x^* eine strikte Lösung des Problems (1.10).

Diese hinreichende Bedingung benutzt man in der Regel erst dann, wenn man einen stationären Punkt findet.

Definition 1.5 (Abstiegsrichtung)

Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar in x . Ein Vektor $d \in \mathbb{R}^n \setminus \{0\}$ heißt Abstiegsrichtung von f in x , wenn

$$\nabla f(x)^T d < 0 \quad (1.14)$$

gilt.

Ist $\nabla f(x) \neq 0$, dann ist beispielsweise $d = -\nabla f(x)$ eine Abstiegsrichtung von f in x .

Satz 1.4 Seien $f : \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar in x und d eine Abstiegsrichtung von f in x . Dann gibt es ein $\hat{\sigma} > 0$ mit

$$f(x + \sigma d) < f(x) \quad \forall \sigma \in]0, \hat{\sigma}[. \quad (1.15)$$

Die meisten Optimierungsverfahren sind iterativ. Sie fangen mit einem Anfangspunkt x^0 an und versuchen dann weitere Punkte $(x^1, x^2, \dots, x^k, \dots)$ zu finden, die besser als die vorherige sind. Viele iterative Verfahren zur Bestimmung einer lokalen Lösung sind

häufig Abstiegsverfahren. In der k -ten Iteration bestimmen sie zu einem Punkt x^k eine Abstiegsrichtung d^k und eine Schrittweite σ_k , sodass für $x^{k+1} := x^k + \sigma_k d^k$

$$f(x^{k+1}) < f(x^k) \quad (1.16)$$

gilt.

Ein Verfahren, welches die negative Gradienten als Abstiegsrichtungen verwendet, ist das Gradientenverfahren.

Verfahren 1.1 (Gradientenverfahren)

1. Wähle einen Startpunkt x^0 und ein Abbruchkriterium $\epsilon > 0$. Setze $k := 0$.
2. Ist $\|\nabla f(x^k)\| < \epsilon \Rightarrow STOP$.
3. Setze $d^k := -\nabla f(x^k)$.
4. Bestimme σ_k so, dass

$$f(x^k + \sigma_k d^k) < f(x^k + \sigma d^k) \quad \forall \sigma \geq 0. \quad (1.17)$$

5. Setze $x^{k+1} := x^k + \sigma_k d^k$ und $k := k + 1 \Rightarrow$ Gehe zu Schritt 2.

Verfahren 1.2 (Newton-Verfahren)

1. Wähle einen Startpunkt x^0 und ein Abbruchkriterium $\epsilon > 0$. Setze $k := 0$.
2. Ist $\|\nabla f(x^k)\| < \epsilon \Rightarrow STOP$.
3. Berechne die Lösung d des linearen Gleichungssystems

$$f''(x^k)d = -\nabla f(x^k). \quad (1.18)$$

Setze $d^k := d$.

4. Setze $x^{k+1} := x^k + d^k$ und $k := k + 1 \Rightarrow$ Gehe zu Schritt 2.

Man kann auch noch eine Schrittweitensteuerung durchführen, dann bekommt man ein Abstiegsverfahren, welches man als das gedämpfte Newton-Verfahren bezeichnet. Man ersetzt dazu den Schritt 4 in dem Newton-Verfahren mit den Schritten 4 und 5 in dem Gradientenverfahren.

1.3. Restringierte Optimierungsprobleme

Definition 1.6 Optimierungsproblem mit linearen Gleichungsnebenbedingungen

$$\min_{x \in \mathbb{R}^n} f(x) \quad (1.19)$$

$$\text{Nb. } Ax = b \quad (1.20)$$

Satz 1.5 Notwendige Bedingung erster Ordnung

Sei x^ lokale Lösung des linearen restringierten Optimierungsproblems und f sein in x^* differenzierbar. Dann gibt es ein $\lambda \in \mathbb{R}^m$ mit*

$$\nabla f(x^*) + A^T \lambda = 0. \quad (1.21)$$

Hat A einen vollen Rang, dann ist λ eindeutig zu bestimmen.

Diese Bedingung heißt die Multiplikatoren Regel von Lagrange. Man bezeichnet λ als die Lagrange-Multiplikator.

Satz 1.6 Hinreichende Bedingung zweiter Ordnung

Sei f in x^ zweimal stetig differenzierbar. Die notwendige Bedingung erster Ordnung sei erfüllt. Es gäbe eine Konstante $\alpha > 0$ mit*

$$d^T f''(x) d \geq \alpha \|d\|^2 \quad d \in \text{Kern } A. \quad (1.22)$$

Dann ist x^ eine strikte Lösung des linearen restringierten Problems.*