

Wir beginnen unseren Exkurs in die kontinuierliche Optimierung mit drei kurzen Einführungsbeispielen. Beim ersten Beispiel liegt ein eindimensionales Problem vor. Anschließend betrachten wir ein mehrdimensionales Optimierungsproblem. Hier werden wir feststellen, dass sich die Methoden aus dem eindimensionalen Beispiel einfach übertragen lassen. Zum Abschluss betrachten wir ein Problem mit Nebenbedingungen.

2.1 Ausgleichskurve (Curve Fitting)

Das erste Beispiel begegnet uns häufig in den Ingenieur- und Naturwissenschaften, immer dann, wenn wir in Experimenten Messungen durchführen und diese anschließend auswerten wollen. Hierzu versuchen wir, eine passende Kurve möglichst gut durch die ermittelten Messwerte zu legen. Um dieses zu verdeutlichen, betrachten wir folgendes Problem:

Wir gehen von einer Kultur mit 100 Bakterien aus und beobachten ihr Wachstum, d. h. wir zählen jede Stunde die Anzahl $N(t)$ der vorhandenen Bakterien. Die so ermittelten Messwerte werden in Tab. 2.1 und Abb. 2.1 gezeigt. Natürlich haben wir uns das eine oder andere Mal beim Bakterienzählen verzählt, sodass unsere Daten einen gewissen Messfehler besitzen.

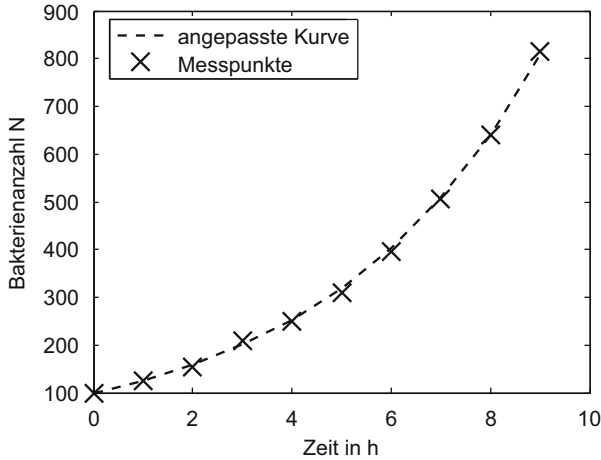
Die Biologie (vgl. z. B. Cypionka [1], Kap. 8) lehrt uns, dass eine Bakterienkultur, zumindest zeitweise, exponentiell wächst. Dieses wird durch die folgende Funktion beschrieben:

$$N(t) = N_0 \cdot e^{\mu \cdot t}, \quad (2.1)$$

wobei N_0 die Anfangspopulation zu Beginn der Beobachtung ($t = 0$) beschreibt, wir mit μ die Wachstumsrate bezeichnen und t die Zeit in Stunden angibt.

Tab. 2.1 Gemessene Bakterienanzahl

t in h	0	1	2	3	4	5	6	7	8	9
N(t)	100	125	155	210	250	310	395	505	640	815

**Abb. 2.1** Messwerte und gefittete Exponentialkurve

Da wir uns beim Wert für die Anfangspopulation sicher sind, können wir $N_0 = 100$ Bakterien setzen. Es bleibt also nur die Wachstumsrate μ als freier Parameter übrig, den wir nun so bestimmen wollen, dass der zugehörige Funktionsgraph bestmöglich an die Messwerte angepasst wird.

Probleme dieser Art werden mit der Methode der kleinsten Quadrate (Least Squares) gelöst. Da die gemessenen Größen möglichst nahe an den theoretisch vorhergesagten Größen auf der Ausgleichskurve liegen sollen, wird die quadratische Abweichung zu jedem Messzeitpunkt betrachtet. Es wird hierbei bewusst nicht der Betrag gewählt, da mit Quadraten einfacher weitergerechnet werden kann. Anschließend wird über alle Zeitpunkte summiert. Das Optimierungsziel besteht dann darin, den oder die Parameter zu finden, welche die Summe minimieren.

Konkret für unser Beispiel bedeutet dieses, dass wir von allen Kurvenpunkten

$$N(t_i) = N_0 \cdot e^{\mu \cdot t_i} \quad (2.2)$$

zu den Zeiten $t_i = 0, 1, \dots, 9$ die Messpunkte $N_i = 100, 125, \dots, 815$ abziehen, dann die Differenzen quadrieren und anschließend die Summe bilden:

$$\min_{\mu \in \mathbb{R}} f(\mu) = (N_0 \cdot e^{\mu \cdot 0} - 100)^2 + (N_0 \cdot e^{\mu \cdot 1} - 125)^2 + \dots + (N_0 \cdot e^{\mu \cdot 9} - 815)^2.$$

Wenn wir das Summenzeichen verwenden, wird der Ausdruck etwas handlicher:

$$\min_{\mu \in \mathbb{R}} f(\mu) = \sum_{i=0}^9 (N_0 \cdot e^{\mu \cdot t_i} - N_i)^2.$$

Um auszudrücken, dass die Wachstumsrate μ der freie Optimierungsparameter ist, wird unter dem Minimum notiert, dass bzgl. μ optimiert wird. Weiter wird der erlaubte Definitionsbereich angegeben. Da wir hier alle reellen Zahlen ($\mu \in \mathbb{R}$) zulassen, handelt es sich um ein Optimierungsproblem ohne Nebenbedingungen. Wir können μ frei wählen.

Unser Ziel ist es, einen möglichst kleinen Wert für die Summe zu erhalten, damit die Abstände der Messpunkte zur Kurve möglichst klein sind. Daher bildet die Summe, die wir mit $f(\mu)$ bezeichnen, in diesem Beispiel die Zielfunktion, welche wir minimieren wollen.

Falls möglich, ist es immer eine gute Idee, die Zielfunktion zu plotten, um sich einen Überblick über das Problem und die möglichen Minima zu verschaffen. Abb. 2.2a zeigt einen Ausschnitt der Zielfunktion für unser Bakterienbeispiel. Wir beobachten deutlich ein (lokales) Minimum zwischen $\mu = 0,2$ und $\mu = 0,3$.

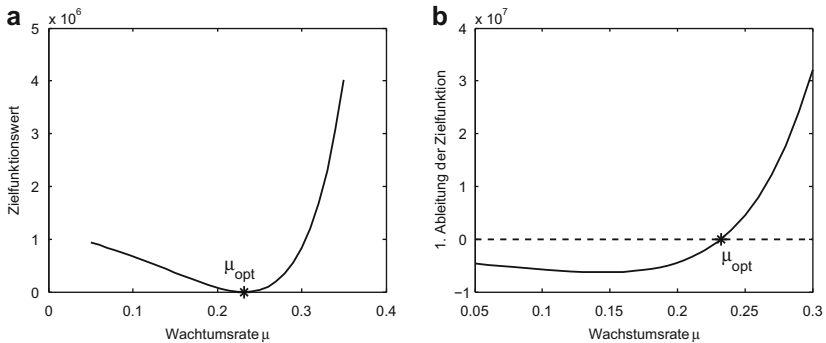


Abb. 2.2 Zielfunktion (a) und erste Ableitung der Zielfunktion (b)

Um jedoch den genauen Wert ablesen zu können, ist der Plot zu ungenau. Wir wollen daher den optimalen Parameter im Folgenden berechnen.

Erinnern wir uns an die Schule: Hier haben wir gelernt, dass bei einem lokalen Extremum (Minimum oder Maximum) die erste Ableitung der Zielfunktion null ist, da dann eine waagerechte Tangente, parallel zur y-Achse, vorliegt. Dieses ist auch beim lokalen Minimum unserer Zielfunktion in Abb. 2.2a zu beobachten.

Wir leiten daher unsere Zielfunktion $f(\mu)$ einmal ab. Wichtig ist, dass wir bzgl. des Optimierungsparameters μ differenzieren und nicht z.B. bzgl. t , da wir den Parameter μ bestimmen wollen und für t feste Werte vorgegeben sind (vgl. Tab. 2.1). Wir erhalten mit der Kettenregel folgende erste Ableitung:

$$f'(\mu) = \cdot \sum_{i=0}^9 \underbrace{N_0 \cdot t_i \cdot e^{\mu \cdot t_i}}_{\text{innere Abl.}} \cdot \underbrace{2 \cdot (N_0 \cdot e^{\mu \cdot t_i} - N_i)}_{\text{äußere Abl.}}.$$

Einen Plot der ersten Ableitung zeigt Abb. 2.2b. Tatsächlich beobachten wir zwischen $\mu = 0,2$ und $\mu = 0,3$ eine Nullstelle in der ersten Ableitung an der Stelle, wo wir das lokale Minimum vermuten.

Um nun alle möglichen Kandidaten für Minima zu bestimmen, setzen wir die erste Ableitung gleich null und berechnen alle Nullstellen. Da dieses hier per Hand nicht möglich ist, lassen wir den Computer für uns rechnen. Dieser liefert $\mu = 0,2322$. Die Abb. 2.2a bestätigt schließlich, dass auch tatsächlich ein Minimum vorliegt.

2.2 Standortplanung

Im zweiten Beispiel stellen wir uns eine Supermarktkette vor, die mehrere Läden in unterschiedlichen Orten betreibt. Für diese Läden soll ein Lager gebaut werden, aber wo? Ziel ist es, das Lager möglichst zentral zwischen die Supermärkte zu platzieren, damit die Entfernung zu diesen möglichst klein ist. Es liegt also ein Optimierungsproblem vor.

Als konkretes Beispiel gehen wir von drei Supermärkten aus, die sich an den folgenden Standorten befinden (vgl. Abb. 2.3b):

$$P_1 = (0, 0), \quad P_2 = (4, -2) \quad \text{und} \quad P_3 = (3, 4). \quad (2.3)$$

Im Optimierungsproblem suchen wir nun nach einem Standort $P = (x, y)$ zwischen den Supermärkten, sodass der Abstand zu allen drei Punkten P_i möglichst klein wird. Um dieses Problem mathematisch zu modellieren, berechnen wir

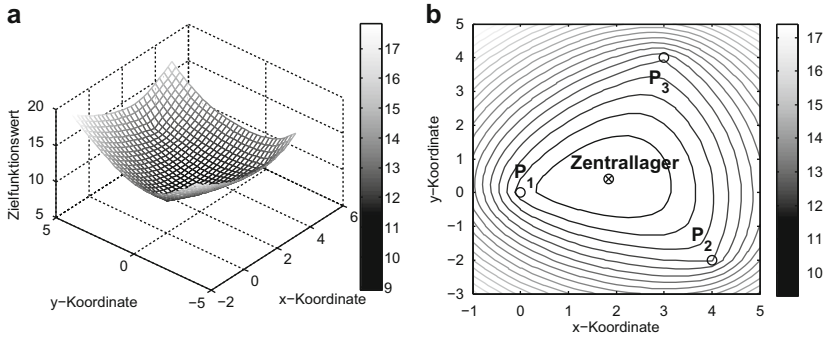


Abb. 2.3 Oberflächen- (a) und Konturenplot (b) der Zielfunktion

zunächst den euklidischen Abstand der drei Supermärkte zum gesuchten Punkt $P = (x, y)$. Diesen erhalten wir nach dem Satz von Pythagoras:

$$d_i := \sqrt{(x - p_x^i)^2 + (y - p_y^i)^2}, \quad i = 1, 2, 3,$$

wobei wir mit p_x^i und p_y^i die Koordinaten der Punkte P_i bezeichnet haben, an denen die Supermärkte liegen.

Wir wollen alle Abstände gleichzeitig minimieren, daher addieren wir sie. Setzen wir unsere Punkte ein, erhalten wir das Optimierungsproblem:

$$\min_{(x, y) \in \mathbb{R}^2} f(x, y) = \underbrace{\sqrt{x^2 + y^2}}_{=d_1} + \underbrace{\sqrt{(x - 4)^2 + (y + 2)^2}}_{=d_2} + \underbrace{\sqrt{(x - 3)^2 + (y - 4)^2}}_{=d_3}.$$

Wir suchen die beiden Koordinaten $x \in \mathbb{R}$ und $y \in \mathbb{R}$, was wir wieder unter dem Minimum notieren. Da sich aus der Problemstellung zunächst keine Einschränkungen an x und y ergeben, liegt ein Optimierungsproblem ohne Nebenbedingungen vor. Natürlich sagt uns die Intuition, dass die optimale Lösung irgendwo zwischen den Punkten P_i liegt, um das Problem jedoch so einfach wie möglich zu halten, geht diese Beobachtung nicht in die mathematische Formulierung ein.

Wir können wieder die Zielfunktion $f(x, y)$ plotten. Abb. 2.3a zeigt einen Oberflächenplot und in Abb. 2.3b werden die Konturen gezeigt. Zusätzlich haben wir die Positionen P_i der Supermärkte zur besseren Orientierung eingezeichnet.

Betrachten wir die Zielfunktion, so fällt auf, dass diese genau ein Minimum besitzt. Wie aber können wir dieses berechnen? Gehen wir zurück zum eindimensionalen Fall: Hier haben wir nach Nullstellen der ersten Ableitung gesucht, da dann die Tangente waagrecht ist. Wir versuchen dieses Kriterium auf den zweidimensionalen Fall zu übertragen. Dabei ersetzen wir die Ableitung $f'(\mu)$ durch den Gradienten $\nabla f(x, y)$. Wenn der Gradient verschwindet, erhalten wir eine waagerechte Tangentialebene, parallel zur $x - y$ -Ebene. Wir haben also ein Kriterium für ein Extremum gefunden.

Bestimmen wir für unser Beispiel den Gradienten der Zielfunktion, indem wir partiell bzgl. beider Koordinaten x und y ableiten und setzen die beiden Komponenten des Gradienten gleich null, ergeben sich zwei, nichtlineare Bestimmungsgleichungen. Diese sind leider wieder zu kompliziert, um sie mit der Hand zu lösen. Verwenden wir ein Computerprogramm, erhalten wir $x = 1,836$ und $y = 0,394$ als optimale Positionen für unser Zentrallager. Diese Lösung wird schließlich auch in Abb. 2.3b gezeigt.

2.3 Haltestellenplanung

Im letzten Beispiel betrachten wir nun ein Problem, bei dem wir Einschränkungen an die Optimierungsparameter vorgeben müssen.

Wir gehen wieder von unterschiedlichen Orten aus. Dieses Mal soll ein Bahnhof so gebaut werden, dass er für Alle möglichst günstig erreichbar ist. Da die Schienen schon verlegt sind, können wir den Bahnhof nicht an einen beliebigen Ort bauen. Er muss an den Gleisen liegen, d.h. wir haben eine Nebenbedingung, welche unsere optimale Lösung erfüllen muss.

Konkret betrachten wir wieder die drei Orte aus dem letzten Beispiel Gl. (2.3), um anschließend die Lösungen vergleichen zu können. Die Schienen können wir als Gerade modellieren, welche durch die Gleichung $y = x - 4$ beschrieben wird. Die Situation wird in Abb. 2.4 dargestellt.

Da wir als Standort für den Bahnhof wieder eine Position $P = (x, y)$ suchen, welche minimalen Abstand zu allen drei Orten besitzen soll, verändert sich unsere Zielfunktion nicht gegenüber dem letzten Beispiel. Wir müssen lediglich die Geradengleichung als Nebenbedingung hinzufügen und erhalten:

$$\min_{(x,y) \in \mathbb{R}^2} f(x, y) = \sum_{i=1}^3 \sqrt{(x - p_x^i)^2 + (y - p_y^i)^2}$$

sodass $y = x - 4$

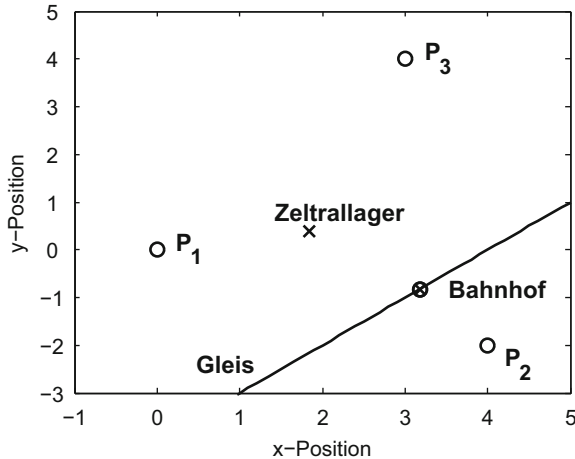


Abb. 2.4 Positionen P_i der drei Orte, optimaler Standort des Bahnhofs und des Zentrallagers aus Abschn. 2.2 zum Vergleich

Wie können wir dieses Optimierungsproblem lösen? Erinnern wir uns wieder an die Schule. Hier haben wir vielleicht sogenannte Extremwertaufgaben behandelt. Wir können die Nebenbedingung $y = x - 4$ in die Zielfunktion einsetzen und eliminieren so den Optimierungsparameter y . Es ergibt sich ein eindimensionales Optimierungsproblem ohne Nebenbedingung:

$$\min_{x \in \mathbb{R}} f(x) = \sum_{i=1}^3 \sqrt{(x - p_x^i)^2 + \underbrace{(x - 4 - p_y^i)^2}_{=y}}.$$

Dieses können wir analog zum letzten Beispiel in Abschn. 2.2 lösen, d.h. wir setzen den Gradienten null und lösen das nichtlineare Gleichungssystem mit dem Computer. Dieser liefert schließlich als optimale Position für unseren Bahnhof:

$$P = (3, 176, -0, 823).$$

Die Lösung haben wir in Abb. 2.4 durch einen Kreis mit Kreuz markiert. Zum Vergleich wird auch die Lösung des Standortproblems aus dem letzten Abschnitt gezeigt (Kreuz). Wir sehen, dass die Nebenbedingung erzwingt, dass der Bahnhof auf der Geraden platziert wird, d.h. die Nebenbedingung hat großen Einfluss auf die Lösung und verändert diese.

Mathematische Optimierung

Eine Einführung in die kontinuierliche Optimierung mit
Beispielen

Pieper, M.

2017, IX, 53 S. 20 Abb., Softcover

ISBN: 978-3-658-16974-9