0 Einleitung

Gegenstand dieses Textes sind numerische Methoden zur Lösung von Problemen der linearen Optimierung, d. h. von sog. „linearen Programmen". Die „mathematische" Optimierung ist ein sehr großes Gebiet mit vielfältigen, verschiedenen Problemtypen aus praktisch allen Bereichen der Wissenschaften, Technik und Wirtschaft. Zur Abgrenzung der betrachteten Problemtypen gegen andere, ebenfalls sehr interessante Aufgabenstellungen, die Gegenstand eigener Abhandlungen sind, wird im Folgenden zunächst einmal kurz angesprochen, was NICHT Gegenstand dieses Textes ist.

0.1 Allgemeine Optimierungsaufgaben und thematische Abgrenzung

In abstrakter Formulierung lautet ein Optimierungsproblem in allgemeinem Kontext wie folgt:

- gegeben: Menge möglicher Aktivitäten und ein Bewertungsmaß;
- gesucht: „optimale" Aktivität.

Derartige Aufgabenstellungen treten typischerweise auf in den Wirtschaftswissenschaften (z. B.: Produktionsplanung, Güterverteilung, Finanzplanung, etc.), der Industrietechnik (z. B.: Prozesssteuerung, Funktionsoptimierung, etc.), dem Verkehrswesen (z. B.: Stromnetze, Flugpläne, etc.) und vielfach auch im Alltag (z. B.: Diätenplanung, Stundenplanentwurf, Bestellplanung, etc.).

Zur quantitativen Lösung solcher Probleme ist zunächst die präzise Formulierung derselben in mathematischer Sprache erforderlich. Ein erster Schritt in diese Richtung lautet wie folgt:

- gegeben: eine Vergleichsmenge M von Objekten (z. B.: Punkte des $\mathbb{R}^n$, stetige Funktionen über einem Bereich $\Omega \subset \mathbb{R}^n$, glatte Flächen im $\mathbb{R}^n$) und eine Abbildung $f : M \to \mathbb{R}$ (Bewertungsmaß);
- gesucht: ein „optimales" Element $x^* \in M$ mit der Eigenschaft $f(x^*) \le f(x)$ für alle $x \in M$.

Zur Abgrenzung der im Folgenden betrachten „linearen (finiten) Optimierung" von anderen, verwandten Aufgabenstellungen der (mathematischen) Optimierung seien letztere zunächst kurz gelistet und danach näher erläutert:

- „Variationsrechnung“ (Begründer: Euler[1] 1750);
- „Approximationstheorie“ (Begründer: Bernstein[2] 1912);
- „Kontrolltheorie“ (Begründer: Pontryagin[3] 1940);
- „Spieltheorie“ (Begründer: John von Neumann[4] 1944);
- „Lineare Programmierung“ (Begründer: Dantzig[5] 1947);
- „Nichtlineare Programmierung“ (Begründer: Kuhn[6]/Tucker[7] 1951).

0.2 Variationsrechnung

In der klassischen „Variationsrechnung“ besteht die Vergleichsmenge M aus Funktionen. Wir erläutern dies anhand der Aufgabe der sog. „Brachistochrone” (griechisch: brachistos = kürzeste, chronos = Zeit), die auf Jacob Bernoulli[8] (1696) zurückgeht,

[1]Leonhard Euler (1707–1783): geb. in Basel; universeller Mathematiker und Physiker; bedeutendster und produktivster Mathematiker seiner Zeit; wirkte in Berlin und St. Petersburg; Arbeiten zu allen mathemischen Gebieten seiner Zeit.

[2]Sergei Natanovich Bernstein (1880–1968): Russischer Mathematik jüdischer Herkunft; nach Studium in Paris wirkte er in Kharkov und Leningrad; fundamentale Beiträge zur Theorie partieller Differentialgleichungen, Differentialgeometrie, Wahrscheinlichkeitstheorie und vor allem zur Approximationstheorie („Bernstein-Polynome“)

[3]Lev Semenovich Pontryagin (1908–1988): Russischer Mathematiker; durch einen Unfall mit 14 Jahren erblindet wurde er dennoch einer der bedeutensten Mathematiker des 20. Jahrhunderts; Prof. am Steklov Institut der Univ. Moskau; wichtige Beiträge u. a. zur Algebraischen Topologie und Differential Topologie sowie zur Kontrolltheorie („Pontryaginsches Maximumprinzip“ und „bang-bang“-Prinzip).

[4]John von Neumann (1903–1957): US-amerikanischer Mathematiker ungarischer Abstammung; wirkte hauptsächlich am Institute for Advanced Studies in Princeton (zus. mit A. Einstein u.a.) und gilt als mathematisches Genie; lieferte fundamentale Beiträge zu den mathematischen Grundlagen der Quantenmechanik, zur Operatortheorie, zur Spieltheorie, zur Gruppentheorie und zur Theorie der partiellen Differentialgleichungen; Pionier der Automatentheorie und theoretischen Informatik.

[5]George Bernard Dantzig (1914–2005): US-amerikanischer Mathematiker; entwickelte 1947 den Simplex-Algorithmus wärend seiner Tätigkeit in einem Forschungslabor der U.S. Air Force; s. sein Buch „Lineare Programmierung und Erweiterungen“, Springer 1966 (Übersetzung aus dem Englischen); seit 1966 Prof. an der Stanford University, USA.

[6]Harold William Kuhn (1925–2014): US-amerikanischer Mathematiker kanadischer Herkunft; bekannt insbesondere durch die Karush-Kuhn-Tucker-Bedingungen, die in einem Tagungsgsbeitrag mit A. W. Tucker präsentiert wurden: „Nonlinear programming“, Proceedings of 2nd Berkeley Symposium. Berkeley, University of California Press. pp. 481492; John von Neumann Preis 1980 mit D. Gale und A. W. Tucker.

[7]Albert William Tucker (1905–1995): Kanadischer Mathematiker; 1933–1974 Prof. an der Princeton University (NJ, USA); Beiträge zur linearen Optimierung und Spieltheorie („Karush-Kuhn-Tucker (KKT) Optimalitätsbeding“).

[8]Bernoulli: Schweizer Mathematiker Familie; Jakob Bernoulli (1655–1705) lehrte in Basel; verwendete bereits die vollständige Induktion; Entdecker der “Bernoulli-Zahlen” und Mitbegründer der Wahrscheinlichkeitsrechnung; sein jüngerer Bruder Johann Bernoulli (1667–1748) wirkte zuletzt in Basel und galt nach dem Tode seines Bruders Jakob als führender Mathematiker seiner Zeit; er leistete Beiträge über Reihen und Differentialgleichungen; sein Sohn Daniel Bernoulli (1700–1782) setzte diese Arbeiten fort; er wirkte in St. Petersburg und Basel und leistete wichtige Beiträge zur Hydromechanik und Gasdynamik.

Aufgabenstellung: *Man verbinde zwei Punkte* $A = (x_0, y_0)$, $B = (x_1, y_1)$ *in der Ebene* $\mathbb{R}^2$ *mit* $y_0 > y_1$ *durch eine Kurve, längs derer ein der Schwerkraft unterworfener (reibungslos) gleitender Massepunkt in kürzester Zeit von* A *nach* B *gelangt.*

Dabei bedeutet „Kurve“ den Graph $\{(x, y(x)) \in \mathbb{R}^2,\, x_0 \le x \le x_1\}$ einer stetig differenzierbaren Funktion $y = y(x)$. Zur Lösung dieses Optimierungsproblems verwenden wir die Bezeichnungen $s = s(x)$ für die Bogenlänge und $v(s)$ für die Geschwindigkeit des Massepunkts (mit Masse m) entlang der Kurve:

$$
\begin{aligned}
s(x) &= \int_{x_0}^{x} \sqrt{1 + y'(x)^2}\, dx, \quad \frac{ds}{dx}(x) = \sqrt{1 + y'(x)^2}, \\
v(s) &= \frac{ds}{dt}(x) = \frac{ds}{dx}\frac{dx}{dt}(x) = \sqrt{1 + y'(x)^2}\,\frac{dx}{dt}(x), \quad dt = \frac{\sqrt{1 + y'(x)^2}}{v(x)}\, dx.
\end{aligned}
$$

Aus dem Energieerhaltungssatz (Newtonsches Gesetz), d. h. der Erhaltung der Gesamtenergie = potentielle Energie + kinetische Energie in der Zeit,

$$
v|_{t=0} = 0, \quad mgy(x) + \frac{m}{2} v(x)^2 = mgy_0,
$$

folgt dann

$$
v(x) = \sqrt{2g(y_0 - y(x))}.
$$

Damit lässt sich die Bestimmung der Kurve mit minimaler Fallzeit wie folgt formulieren:

$$
T(y) := \int_{x_0}^{x_1} \frac{\sqrt{1 + y'(x)^2}}{\sqrt{2g(y_0 - y(x))}}\, dx \;\to\; \min! \tag{0.2.1}
$$

wobei zum Vergleich alle Funktionen

$$
y \in M := \{z \in C^1[x_0, x_1],\, z(x_0) = y_0,\, z(x_1) = y_1\}
$$

zugelassen sind. Eine Lösung dieses nichtlinearen „Variationsproblems“ kann mit Hilfe von Techniken der „Variationsrechnung“ und der Theorie von Differentialgleichungen explizit angegeben werden.

Abbildung 1: *Darstellung des Brachistochronen-Problems*

0.3 Approximationstheorie

In der Approximationstheorie geht es um die „beste“ Approximation einer allgemeinen, komplizierten Funktion durch einfacher strukturierte wie z. B. Polynome, $p(x) = c_0+c_1x+c_2x^2+\cdots+c_mx^m$, rationale Funktionen $r(x) = p(x)/q(x)$ oder auch trigonometrische Polynome $t(x) = a_0 + a_1\sin(x) + b_1\cos(x) + a_2\sin(2x) + b_2\cos(2x) + \cdots + a_m\sin(mx) + b_m\cos(mx)$. Dies ist notwendig, um mit solchen Funktionen auf Computern arbeiten zu können. Dabei kommt es sehr auf die Wahl des Qualitätsmaßes für den Grad an „Approximation“ an. Ein standard Beispiel ist gegeben durch die Taylor[9]-Entwicklung

$$f(x) \approx t_f^m(x) = f(x_0) + \frac{1}{1!}f'(x_0)(x-x_0) + \cdots + \frac{1}{m!}f^{(m)}(x_0)(x-x_0)^m$$

einer hinreichend oft differenzierbaren Funktion f, welche offenbar Lösung des folgenden Optimierungsproblems ist:

$$F(t_f^m - f) := \sum_{j=0}^{m} |t_f^{m(j)}(x_0) - f^{(j)}(x_0)| = \min_{p\in P_m} \Big\{ \sum_{j=0}^{m} |p^{(j)}(x_0) - f^{(j)}(x_0)| \Big\}. \tag{0.3.2}$$

Allgemein sei $f \in C[a,b]$ eine gegebene stetige Funktion auf einem Intervall $I = [a,b] \subset \mathbb{R}^1$. Wir suchen nach einer „besten Approximation“ $p \in P_m$ zu f im Raum P_m der Polynome vom maximalen Grad m, so dass die Differenz $f-p$ in einem vorgeschriebenen Sinn minimal wird. Praxisrelevante Beispiele für solche Qualitätsmaße sind das der sog. „Gauß[10]-Approximation“ und der „Tschebyscheff[11]-Approximation“ sowie der einfachen „Interpolation“:

$$\begin{aligned}
&a) \quad F(f-p) := \Big(\int_a^b |f(x)-p(x)|^2\,dx \Big)^{1/2} \quad \text{(Gauß-Approximation)},\\
&b) \quad F(f-p) := \max_{x\in[a,b]} |f(x)-p(x)|, \quad \text{(Tschebyscheff-Approximation)},\\
&c) \quad F(f-p) := \sum_{i=0}^{m} |f(x_j)-p(x_j)|, \quad a \le x_0 < x_2 < \cdots < x_m \le b \quad \text{(Interpolation)}.
\end{aligned}$$

Unter Verwendung der natürlichen Monombasis $\{1, x, x^2, \ldots, x^m\}$ von P_m und der allgemeinen Darstellung $p(x) = \sum_{i=0}^m \alpha_i x^i$ der gesuchten Bestapproximation lautet die Approximationsaufgabe in algebraischer Formulierung wie folgt:

$$F\Big(f - \sum\nolimits_{i=0}^{m} \alpha_i x^i\Big) \ \to\ \min! \tag{0.3.3}$$

[9]Brook Taylor (1685–1731): Englischer Mathematiker und Schüler Newtons; die nach ihm benannte Reihenentwicklung war im Kern bereits Gregory, Newton, Leibniz und Johann Bernoulli bekannt.

[10]Carl Friedrich Gauß (1777–1855): Bedeutender deutscher Mathematiker, Astronom und Physiker; wirkte in Göttingen; fundamentale Beiträge zur Arithmetik, Algebra und Geometrie, Begründer der modernen Zahlentheorie, Bestimmung von Planetenbahnen durch „Gauß-Ausgleich“, Arbeiten zum Erdmagnetismus und Konstruktion eines elektromagnetischen Telegraphen.

[11]Pafnuty Lvovich Tschebyscheff (russ.: Chebyshev) (1821–1894): Russischer Mathematiker; Prof. in St. Petersburg; Beiträge zur Zahlentheorie, Wahrscheinlichkeitstheorie und vor allem zur Approximationstheorie; entwickelte eine allgemeine Theorie orthogonaler Polynome.

Die Gaußsche Approximationsaufgabe führt auf ein *lineares* Problem im $\mathbb{R}^m$, das mit Methoden der Numerischen Linearen Algebra gelöst werden kann, während die Tschebyscheffsche Approximationsaufgabe generisch *nichtlinear* und viel schwieriger zu lösen ist. Die Interpolationsaufgabe wiederum besitz eine eindeutige Lösung, die etwa in Langrangescher[12] Darstellung explizit angebbar ist.

0.4 Kontrolltheorie

In der Kontrolltheorie geht es um die „optimale" Steuerung der Lösung eines sog. „dynamischen Systems", z. B. einer gewöhnlichen Differential- oder Integralgleichung, zum Minimum eines gegebenen „Kostenfunktionals". Zu bestimmen ist also etwa eine „Steuerfunktion" $q = q(x)$ als Koeffizient in einer gewöhnlichen Differentialgleichung

$$u'(t) = f(t, u(t), q(t)), \quad t \in [0, T], \quad u(0) = u^0,$$

so dass z. B.

$$F(u, q) := \psi_1(u(T)) + \int_0^T \psi_2(u(x))\, dt + \int_0^T \psi_3(q(x))\, dt \ \to \ \min!$$

Lösungen derartiger oft hochgradig nichtlinearer Optimierungsaufgaben lassen sich in der Regel nur numerisch approximieren. Wir geben ein Beispiel aus den Wirtschaftswissenschaften (optimale Produktionsplanung):

Ein Industriebetrieb produziert ein Produkt P, dessen Produktion über die Zeit optimiert werden soll bei Minimierung der Produktionskosten und gleichzeitiger vollen Befriedigung der Nachfrage. Dabei bezeichne $x(t)$ den Lagerbestand, $r(t)$ die Produktionsrate und $d(t)$ die Nachfragerate. Die zeitliche Entwicklung des Lagerbestands ist dann beschrieben als Lösung der Anfagswertaufage

$$x'(t) = r(t) - d(t), \ \ t \geq 0, \quad x(0) = x^0 \text{ (gegeben)}.$$

Zu bestimmen ist nun eine stetige Funktion $r(t) \geq 0$ (optimale Produktionsrate), so dass auf dem Intervall $I = [0, T]$ gilt

$$x(t) = x(0) + \int_0^t \{r(s) - d(s)\}\, ds \geq 0,$$

und

$$F(r) := \int_0^T \{c[r(t)] + h[x(t)]\}\, dt \ \to \ \min!$$

wobei $c[r(t)]$ die Produktionskostenrate und $h[x(t)]$ die Lagerkostenrate bezeichnen.

[12] Joseph Louis de Lagrange (1736–1813): Französischer Mathematiker; 1766–1787 Direktor der mathem. Klasse der Berliner Akademie, dann Prof. in Paris; bahnbrechende Arbeiten zur Variationsrechnung, zur komplexen Funktionentheorie sowie zur Theoretischen Mechanik und Himmelsmechanik.

0.5 Spieltheorie

In der Spieltheorie geht es um die Entwicklung optimaler Strategien für die Handlungsweisen zweier oder mehrerer „Spieler“ zur Maximierung oder Minimierung der Ergebnisse der einzelnen Spieler unter Berücksichtigung „optimaler“ Aktionen der jeweils anderen Spieler. Zentrale Frage ist dabei die Existenz eines „Gleichgewichtszustands“, in dem sich alle Spieler „optimal“ verhalten. Dies soll anhand zweier Beispiele erläutert werden.

Beispiel 0.1 (aus den Politikwissenschaften: Wahlkampforganisation): Zwei Konkurrenten A und B um ein politisches Amt wollen unabhängig von einander ihre Werbeetats (Geld) so auf n verschiedene geographische Gebiete Ω_i, $i = 1, \dots, n$, verteilen, dass ihre jeweiligen Erfolgschancen in der Wahl maximiert werden. Dabei sind ihnen die Stragien des anderen Kandidaten unbekannt. Es bezeichnen a_i und b_i die von Kandidaten A bzw. B für Gebiet Ω_i, $i = 1, \dots, n$, aufgewendeten Werbemittel. Die Anzahl der noch unentschiedenen Wähler in Gebiet Ω_i sei u_i. Zur Modellierung des Problems sei angenommen, dass die auf A und B in Gebiet Ω_i aufgrund der Werbung entfallenden Stimmen gegeben sind durch

$$\Sigma_i(A) := \frac{a_i u_i}{a_i + b_i}, \quad \Sigma_i(B) := \frac{b_i u_i}{a_i + b_i}.$$

Dann ist die totale Stimmendifferenz als Resultat der Verteilung der Werbemittel

$$\Delta(A, B) = \sum_{i=1}^{n} \frac{a_i - b_i}{a_i + b_i} u_i. \tag{0.5.4}$$

Das Ziel von A und B ist es nun, $\Delta(A, B)$ zu maximieren bzw. zu minimieren. Die „Lösung“ dieser Aufgabe ist naturgemäß kompliziert.

Beispiel 0.2 (aus dem Alltag: Knobelspiel)): Zwei Spieler vergnügen sich mit dem Knobelspiel „Stein-Schere-Papier“. Dabei gelten die Wertigkeitsregeln

$$\text{Stein} > \text{Schere} > \text{Papier} > \text{Stein}$$

und

$$\text{gewonnen} = 1, \quad \text{unentschieden} = 0, \quad \text{verloren} = -1.$$

Dies führt auf die sog. „Auszahlungstafel“

	Stein	Schere	Papier
Stein	0	1	−1
Schere	−1	0	1
Papier	1	−1	0

Gesucht sind nun optimale Spielstrategien für beide Spieler. Wegen der Verfügbarkeit einer Auszahlungstafel wird dieses Spiel auch „Matrixspiel“ genannt.

Lösungsansätze für solche Aufgaben der „Spieltheorie“ werden in diesem Text als Anwendungen der Lösungstheorie für „Lineare Optimierungsaufgaben“ (sog. „Lineare Programme“) hergeleitet werden. Andererseits können Aufgaben der „Linearen Programmierung“ als Spezialfälle der Spieltheorie interpretiert werden.

0.6 Lineare Programmierung

„Lineares Programm“ ist die historisch bedingte Bezeichnung für lineare Optimierungsaufgaben vom folgenden Typ:

Beispiel 0.3 (Produktionsplanung): Eine Fabrik kann zwei Typen A und B eines Produkts unter folgenden Bedingungen herstellen:

Produkt	Typ A	Typ B	maximal möglich
Stück pro Tag	x_1	x_2	100 Stück
Arbeitszeit pro Stück	4	1	160 Stunden
Kosten pro Stück	20	10	1100 EURO
Gewinn pro Stück	120	40	? EURO

Wie müssen x_1 und x_2 gewählt werden, damit der Gewinn maximal wird? Dabei muss offenbar der lineare Ausdruck

$$Q(x_1, x_2) := 120x_1 + 40x_2$$

zu einem Maximum gemacht werden unter den linearen Nebenbedingungen

$$\begin{aligned} x_1 + \quad x_2 &\leq \;\; 100 \\ 4x_1 + \quad x_2 &\leq \;\; 160 \;, \quad x_1 \geq 0\,, \quad x_2 \geq 0. \\ 20x_1 + 10x_2 &\leq 1100 \end{aligned} \tag{0.6.5}$$

Dies ist ein lineares Programm in sog. “Standardform”.

Beispiel 0.4 (Transportplanung): Die Produktion von 7 Zuckerfabriken soll so auf 300 Verbrauchsorte verteilt werden, dass der Bedarf befriedigt wird und die Transportkosten minimiert werden.

Fabrik	$F_j \; (j = 1, \dots, 7)$,	Verbrauchsort	$G_k \; (k = 1, \dots, 300)$
Produktion	a_j (pro Monat),	Verbrauch	r_k (pro Monat)

transportierte Menge $F_j \to G_k :\; x_{jk}$, Kosten c_{jk} (pro Einheit).

Es sei vorausgesetzt, dass Bedarf und Produktionsmenge gleich sind:

$$\sum_{k=1}^{300} r_k \;=\; \sum_{j=1}^{7} a_j .$$

Zu minimieren sind die Gesamtkosten

$$Q(x_{1,1},\ldots,x_{7,300}) \;:=\; \sum_{j=1}^{7}\sum_{k=1}^{300} c_{jk}x_{jk}$$

unter den Nebenbedingungen $x_{jk} \ge 0$ und

$$\sum_{j=1}^{7} x_{jk} \;=\; r_k \quad (k=1,\ldots,300), \qquad \sum_{k=1}^{300} x_{jk} \;=\; a_j \quad (j=1,\ldots,7).$$

Dies ist ein lineares Programm in sog. “kanonischer Form”.

Beispiel 0.5 (Approximationstheorie): Eine Funktion $f \in C[0,1]$ soll durch ein Polynom $p(x) = \sum_{i=1}^{6} x_i t^{i-1} \in P_5$ approximiert werden, so dass

$$x_0 := \max_{k=1,\ldots,10} |f(1/k) - p(1/k)|$$

minimal wird. Dies kann wie folgt als lineares Programm geschrieben werden:

$$\begin{aligned} x_0 + \sum_{i=1}^{6} x_i(1/k)^{i-1} &\ge f(1/k), \quad k=1,\ldots,10,\\ x_0 - \sum_{i=1}^{6} x_i u(1/k)^{i-1} &\ge -f(1/k), \quad k=1,\ldots,10. \end{aligned} \tag{0.6.6}$$

Dieses Beispiel zeigt, dass auch scheinbar *nichtlineare* Optimierungsaufgaben unter Umständen als *lineares* Programm formuliert werden können.