

# Hinweise zur Mathematik

Max Albert\*

5. April 2021

Diese Skript enthält Zusammenfassungen zu einigen Themen, die für die Vorlesung „Mikroökonomische Theorie II“ und alle darauf aufbauenden Vorlesungen der Professur VWL VI relevant sind.

Alle Veranstaltungen setzen den Stoff der Mathematikveranstaltung des ersten Semesters als bekannt voraus. Etwaige Lücken in Ihren Kenntnissen sollten Sie dringend schließen. Weder in den Vorlesungen noch in den Übungen kann die notwendige Mathematik in ausreichendem Umfang erklärt werden. Was Sie dort zur Mathematik hören, dient nur dazu, Sie an bekannten Stoff zu erinnern oder auf dieser Grundlage neue Begriffe einzuführen. Sie müssen Ihre Lücken daher durch selbstständige Arbeit schließen.

Einige mathematische Konzepte sind von besonderer Bedeutung und werden immer wieder verwendet:

- Differentialrechnung, insbesondere partielle Ableitungen, totales Differential, Elastizitäten, Satz über implizite Funktionen und ihre Ableitungen, Höhenlinien, Ableitungen der Umkehrfunktion,
- Optimierung mit und ohne Nebenbedingung, Bedingungen erster und zweiter Ordnung, Lagrangeansatz,
- homogene und homothetische Funktionen,
- offene, abgeschlossene und konvexe Mengen sowie
- (Quasi-)Konkavität und (Quasi-)Konvexität von Funktionen.

Wir empfehlen Ihnen daher, sobald wie möglich zu überprüfen, ob Sie über ausreichende Kenntnisse zu diesen Stichwörtern verfügen.

Es ist manchmal sinnvoll, sich Funktionen grafisch zu veranschaulichen. Insbesondere bei Funktionen mit zwei unabhängigen Variablen kann dabei ein Mathematikprogramm sehr hilfreich sein. Beachten Sie dazu die Hinweise zu Literatur und Software und nützliche Links auf der Website der Professur für Statistik und Ökonometrie unter „Lehre“.

## 1 Allgemeine Notation

**Gleichheit, Definition, Identität:** Eine Gleichung legt eine Menge von Lösungswerten fest. Beispiel:  $x^2 - 1 = 0$  für  $x \in \mathbb{R}$  legt fest, dass  $x \in \{1, -1\}$ .

---

\*Ich danke Henrik Egbert, Lothar Grall, Matthias Greiff und Andreas Hildenbrand, die in früheren Jahren als Übungsleiter zur Verbesserung dieses Skripts wesentlich beigetragen haben.

Für Definitionen (Definitionsgleichungen) wird das Symbol  $:=$  oder  $=:$  verwendet; der definierte Ausdruck steht auf der Seite des Doppelpunkts.

Definitionen führen zu Identitäten; dies sind Gleichungen, die für alle Einsetzungen der Variablen gelten. Beispiel:  $h(x) := f[g(x)]$  definiert die Funktion  $h$  als Verkettung der Funktionen  $g$  und  $f$ , d.h.  $h := f \circ g$  mit  $\circ$  als dem Symbol für die Verkettung von Funktionen. Daraus ergibt sich beispielsweise die Identität  $f^{-1}[h(x)] \equiv g(x)$ . Diese Identität besagt, dass die Gleichung  $f^{-1}[h(x)] = g(x)$  für alle möglichen Werte von  $x$  erfüllt ist. Mit anderen Worten: Die Funktionen auf beiden Seiten sind gleich. Das kann man auch durch eine Gleichung für die Funktionen ausdrücken:  $f^{-1} \circ h = g$ .

Häufig verwendet man einfach das Gleichheitszeichen, ohne sich um den Unterschied zwischen Gleichungen, Definitionen und Identitäten zu kümmern. Das ist normalerweise unproblematisch, weil aus dem Kontext hervorgeht, was gemeint ist. Im nachfolgenden Text werden Definitionsgleichungen immer durch  $:=$  oder  $=:$  gekennzeichnet. Identitäten werden jedoch nur gelegentlich durch  $\equiv$  gekennzeichnet.

**Klammerung:** Für verschachtelte Ausdrücke werden zur leichteren Lesbarkeit manchmal verschiedene Klammern verwendet, z.B. bei Summen wie  $a + \{b + [c + (d + e)]\}$  oder bei verschachtelten Funktionen wie  $h\{g[x, f(z)]\}$ . Diese Klammern dürfen nicht (und können eigentlich auch nicht) mit Mengenklammern (s. Bsp. oben:  $x \in \{1, -1\}$ ) oder Klammern für die Angabe von Intervallen (z.B. das links offene, rechts geschlossene Intervall  $]0, 1[$ ) verwechselt werden.

**Rundung:** Es wird durchgehend die mathematische und nicht die kaufmännische Rundung verwendet. Das ist wichtig für Klausuren, in denen häufig nach gerundeten Werten gefragt wird.

Beispiele: Wenn man  $\frac{1}{3}$  auf zwei Stellen hinter dem Dezimalpunkt angibt, geht man von drei Stellen aus, also  $0.333\dots$ , und rundet ab, weil die dritte Stelle kleiner als 5 ist:  $\frac{1}{3} \approx 0.33$ . Bei  $\frac{2}{3}$  erhält man  $0.666\dots$  und rundet auf, weil die dritte Stelle größer als 5 ist:  $\frac{2}{3} \approx 0.67$ . Ist die dritte Stelle gleich 5, dann wird aufgerundet, wenn danach noch Stellen ungleich Null folgen:  $\frac{66.502}{100} \approx 0.666$ . Sind aber alle folgenden Stellen gleich Null (d.h. ist die zu runde Zahl genau in der Mitte zwischen den möglichen gerundeten Werten), dann wird so gerundet, dass die zweite Stelle entweder eine gerade Zahl oder 0 ist:  $\frac{33.5}{100} \approx 0.34$  und  $\frac{66.5}{100} \approx 0.66$ .

Rechnen Sie so weit wie möglich mit den *exakten* Werten, auch wenn gerundete Zwischenergebnisse verlangt werden. Frühzeitiges Runden führt u.U. dazu, dass Sie die geforderte Genauigkeit verfehlen. Beispiel:  $\frac{1}{3} \cdot \frac{9}{5} \approx 0.33 \cdot 1.8 \approx 0.59$  statt  $\frac{1}{3} \cdot \frac{9}{5} = \frac{3}{5} = 0.60$ .

**Variablen- und Funktionssymbole:** Häufig wird, um die Zahl der verwendeten Symbole klein zu halten, dasselbe Symbol für eine Variable und eine Funktion verwendet. Beispiel:  $y = y(x, z)$ . Wenn das bei längeren Rechnungen zur Verwirrung führt, empfiehlt es sich, ein eigenes Symbol für die Funktionen einzuführen.

## 2 Differentialrechnung

**Ableitungen:** Für die Ableitung einer Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  schreiben wir  $f'(x)$ ; die Ableitung an einer bestimmten Stelle wie z.B.  $x = 7$  lautet dann  $f'(7)$ .

Für eine Funktion  $g: \mathbb{R}^2 \rightarrow \mathbb{R}$  findet man für die erste partielle Ableitung nach der ersten Variablen am häufigsten folgende Schreibweisen:

$$\frac{\partial g(x, y)}{\partial x} \text{ oder } g_x(x, y) \text{ oder } D_1 g(x, y)$$

Für diese Veranstaltung ist die letztgenannte Schreibweise verbindlich.

Die Ableitung an einer bestimmten Stelle wie z.B.  $(x, y) = (13, 5)$  lautet  $D_1g(13, 5)$ . Der Index am  $D$ -Operator gibt die Nummer der Variablen an, nach der abgeleitet wird; es spielt damit keine Rolle, welcher Buchstabe für diese Variable verwendet wird. Beachten Sie, dass  $D_1g(13, 5)$  eine Zahl ist, während es sich bei  $D_1g(x, y)$ ,  $D_1g(13, y)$  und  $D_1g(x, 5)$  um (verschiedene) Funktionen handelt.

Diese Schreibweise für partielle Ableitungen erscheint Ihnen vielleicht zunächst etwas ungewohnt. Sie ist aber eine große Erleichterung beim Schreiben (insbesondere in Word oder Powerpoint). Wenn Sie beispielsweise die in der Ökonomie immer noch gebräuchliche Schreibweise mit dem  $\partial$ -Symbol verwenden, müssten Sie anstelle von  $D_1g(13, 5)$  folgenden Ausdruck schreiben:

$$\left. \frac{\partial g(x, y)}{\partial x} \right|_{(x, y) = (13, 5)}$$

Gelegentlich hat eine Funktion keinen eigenen Namen, z.B.  $g[f(x), x]$ . Um die Ableitung dieses Ausdrucks nach  $x$  zu bezeichnen, besteht die beste Möglichkeit darin, der Funktion einen Namen zu geben, also z.B.  $k(x) := g[f(x), x]$  zu definieren, und dann die gesuchte Ableitung mit  $k'(x)$  zu bezeichnen. Es gilt natürlich  $k'(x) = D_1g[f(x), x] f'(x) + D_2g[f(x), x]$ .

Alternativ wird manchmal auch  $(g[f(x), x])' = D_1g[f(x), x] f'(x) + D_2g[f(x), x]$  geschrieben. Mit einer einzigen unabhängigen Variablen ist das möglich. Bei mehreren unabhängigen Variablen gibt es jedoch keine vernünftige Alternative zur Einführung eines neuen Funktionsnamens. Beispiel: Wir betrachten  $g[f(y), x, y]$ . Um jetzt die Ableitung dieses Ausdrucks nach  $y$  anzugeben, muss man  $k(x, y) := g[f(y), x, y]$  definieren und kann dann die gewünschte Ableitung eindeutig mit  $D_2k(x, y)$  bezeichnen. Es gilt natürlich

$$D_2k(x, y) = D_1g[f(y), x, y] f'(y) + D_3g[f(y), x, y] .$$

Wird eine Funktion zweimal abgeleitet, z.B. zuerst nach der ersten und dann nach der zweiten Variable, schreibt man  $D_2[D_1g(x_1, x_2)]$  oder bequemer  $D_{1,2}g(x_1, x_2)$ . Mehrfache Ableitungen nach derselben Variablen lassen sich auch durch einen hochgestellten *geklammerten* Index kennzeichnen, also  $D_1^{(2)}g(x_1, x_2)$  statt  $D_{1,1}g(x_1, x_2)$ . Diese Schreibweise werden wir nicht verwenden.

Im Einklang mit dieser Schreibweise kann man im übrigen auch  $Df(x)$  statt  $f'(x)$  schreiben und  $D^{(2)}f(x)$  statt  $f''(x)$ . Auch davon werden wir keinen Gebrauch machen.

**Ableitung und Näherungsrechnung:** Wir betrachten eine Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$ , die den Zusammenhang zwischen einer unabhängigen Variable  $x$  und einer abhängigen Variable  $y$  beschreibt, also  $y = f(x)$ . Die Ableitung der Funktion ist

$$f'(a) := \lim_{\Delta x \rightarrow 0} \frac{f(a + \Delta x) - f(a)}{\Delta x} ,$$

falls der Limes existiert. Dabei ist  $\Delta x$  einfach eine Variable; häufig wird  $h$  statt  $\Delta x$  verwendet. Die Verwendung von  $\Delta x$  statt  $h$  erinnert daran, dass diese Variable die Abweichung der Variable  $x$  vom Ausgangspunkt  $a$  angibt.

Die erste Näherung für die Veränderung von  $f$  an der Stelle  $a$  wird durch die lineare Funktion  $f'(a)\Delta x$  beschrieben, wobei  $\Delta x$  die unabhängige Variable ist. Es gilt  $\Delta y := f(a + \Delta x) - f(a) \approx f'(a)\Delta x$ . Aus der Definition der Ableitung folgt, dass die Näherung für hinreichend kleine Absolutwerte von  $\Delta x$  immer besser wird in dem Sinne, dass der Fehler der Näherung  $\Delta y$  im Verhältnis zu

$\Delta x$  immer geringer wird:

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y - f'(a)\Delta x}{\Delta x} \equiv \lim_{\Delta x \rightarrow 0} \frac{f(a + \Delta x) - f(a) - f'(a)\Delta x}{\Delta x} \equiv$$

$$\lim_{\Delta x \rightarrow 0} \frac{f(a + \Delta x) - f(a)}{\Delta x} - f'(a) \equiv 0$$

Anstelle der unabhängigen Variablen  $h$  oder  $\Delta x$  verwenden wir im folgenden  $dx$  und schreiben für die Näherungsformel  $dy = f'(a)dx$ , womit daran erinnert werden soll, dass  $dy$  für hinreichend kleine Werte von  $|dx|$  eine gute Näherung für die Veränderung von  $y$  ist. Auch in dieser Schreibweise sind  $dy$  und  $dx$  nichts anderes als abhängige und unabhängige Variablen einer linearen Funktion. Wir verwenden also  $\frac{dy}{dx}$  nicht als Symbol für die Ableitung.

Die Idee lässt sich auch auf Funktionen mehrerer Variabler ausdehnen. Wir betrachten eine Funktion  $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ , die den Zusammenhang zwischen der unabhängigen Variablen  $x_1$  und  $x_2$  und der abhängigen Variable  $y$  beschreibt, also  $y = g(x_1, x_2)$ . Die erste Näherung für die Veränderung von  $y$  an der Stelle  $(a_1, a_2)$  wird durch die lineare Funktion  $D_1g(a_1, a_2)dx_1 + D_2g(a_1, a_2)dx_2$ , das sog. totale Differential, beschrieben. Die Näherung ist desto besser, je kürzer der Vektor  $(dx_1, dx_2)$ , d.h. desto kleiner  $\sqrt{(dx_1)^2 + (dx_2)^2}$  ist. Wir schreiben die Näherungsformel (das sog. totale Differential) auch mit  $dy$  als unabhängiger Variable, um daran zu erinnern, dass  $dy$  für kleine Veränderungen eine gute Näherung für die Veränderung von  $y$  ist:  $dy = D_1g(a_1, a_2)dx_1 + D_2g(a_1, a_2)dx_2$ .

**Ableitung einer impliziten Funktion:** Wir betrachten eine stetig differenzierbare Funktion  $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ . Sei  $g(a_1, a_2) = y_0$ , wobei  $y_0$  eine Konstante ist. Sei  $U_1$  eine offene Umgebung von  $a_1$  und  $U_2$  eine offene Umgebung von  $a_2$ . Zu jedem  $x_2 \in U_2$  möge genau ein  $x_1 \in U_1$  mit  $g(x_1, x_2) = y_0$  existieren. Mit anderen Worten: Wir können  $x_1$  als Funktion von  $x_2$  auffassen, also  $x_1 = h(x_2)$ , wobei die sogenannte implizite Funktion  $h: U_2 \rightarrow U_1$  dadurch definiert ist, dass sie  $g[h(x_2), x_2] \equiv y_0$  für alle  $x_2 \in U_2$  erfüllt. Die Existenz einer solchen Umkehrfunktion erfordert  $D_2g(a_1, a_2) \neq 0$ .

Wenn  $h$  an der Stelle  $(a_1, a_2)$  differenzierbar ist, gilt

$$D(g[h(a_2), a_2]) = D_1g[a_1, a_2]h'(a_2) + D_2g[a_1, a_2] = 0$$

und somit

$$h'(a_2) = -\frac{D_2g[a_1, a_2]}{D_1g[a_1, a_2]}.$$

Die Näherungsformel für die Veränderung von  $x_1$  gemäß der impliziten Funktion  $x_1 = h(x_2)$  ist also

$$dx_1 = -\frac{D_2g[a_1, a_2]}{D_1g[a_1, a_2]}dx_2.$$

Folgende Form der Berechnung findet sich häufig: Man geht vom totalen Differential

$$dy = D_1g(a_1, a_2)dx_1 + D_2g(a_1, a_2)dx_2$$

von  $g$  an der Stelle  $a_1, a_2$  aus, also von der Näherungsformel für die Veränderung von  $y$  (s.o.). Wenn sich  $x_1$  und  $x_2$  ausgehend von  $(x_1, x_2) = (a_1, a_2)$  so verändern, dass die Funktion  $g$  näherungsweise auf dem Niveau  $y_0$  konstant bleibt, muss

$$dy = D_1g(a_1, a_2)dx_1 + D_2g(a_1, a_2)dx_2 = 0$$

gelten. Daraus ergibt sich die obige Näherungsformel, die aber in diesem Fall meist umgeformt wird:

$$\frac{dx_1}{dx_2} = - \frac{D_2 g(a_1, a_2)}{D_1 g(a_1, a_2)}.$$

Auf der rechten Seite steht die Ableitung der impliziten Funktion, die bei dieser Herleitung allerdings keinen Namen erhält.

Mit Hilfe des totalen Differentials lässt sich schnell und intuitiv rechnen. Allerdings kann man bei komplizierteren Berechnungen leicht Fehler machen, weil die impliziten Funktionen nicht ausdrücklich eingeführt werden, so dass nicht klar ist, welche Variable als Funktion welcher anderen Variable betrachtet wird.

**Ableitung der Umkehrfunktion:** Wir betrachten eine umkehrbar eindeutige (bijektive) und differenzierbare Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$ . Die Umkehrfunktion lautet  $f^{-1}(z) =: g(z)$ . Es gilt  $g'(z) = (f^{-1})'(z) = 1/f'(x)$ , wobei jedoch  $z = f(x)$  oder äquivalent  $x = g(z)$  vorausgesetzt werden muss. Die Behauptung  $g'(z) = 1/f'(x)$  ohne die letztgenannte Einschränkung wäre falsch. Die Einschränkung kann auch berücksichtigt werden, indem man  $g'[f(x)] \equiv 1/f'(x)$  oder  $g'(z) \equiv 1/f'[g(z)]$  schreibt.

Eine Kurzform wie  $[f^{-1}]' = 1/f'$  ist irreführend, weil nicht erwähnt wird, an welchen Stellen die Funktionen ausgewertet werden müssen, damit die Kehrwertbeziehung gilt.

Beispiel: Sei  $f(x) = x^3$ . Dann gilt  $f^{-1}(z) = z^{1/3} =: g(z)$ . Es gilt nun  $g'(z) = [f^{-1}]'(z) = z^{-2/3}/3$  und  $f'(x) = 3x^2$ . Weiter gilt  $f(2) = 8$ ,  $g'(8) = [f^{-1}]'(8) = 1/12$  und  $f'(2) = 12$ , also

$$g'(8) = [f^{-1}]'(8) = 1/f'(2).$$

Es gilt jedoch nicht

$$[f^{-1}]'(2) = 1/f'(2),$$

wie es die gelegentlich zu findende Kurzform  $[f^{-1}]' = 1/f'$  nahezu legen scheint.

### 3 Optimierungsprobleme ohne Nebenbedingung

**Notation:** Soll eine Zielfunktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  durch geeignete Wahl der Instrumentvariablen  $x_1, x_2$  maximiert werden, schreiben wir das wie folgt:

$$\max_{x_1, x_2} f(x_1, x_2).$$

Meist werden wir Nebenbedingungen wie  $x_1, x_2 \geq 0$  (Nichtnegativitätsbedingungen) nicht explizit erwähnen, weil die entsprechenden Variablen bereits als positive Variablen definiert wurden. Soll die Nichtnegativität explizit erwähnt werden, bieten sich zwei Möglichkeiten:

$$\max_{x_1, x_2 \geq 0} f(x_1, x_2)$$

oder

$$\max_{x_1, x_2} \{f(x_1, x_2) : x_1, x_2 \geq 0\}.$$

**Lösung:** Das obige Maximierungsproblem führt zu den Bedingungen 1. Ordnung (kurz „FOC“ für „first-order conditions“)

$$D_j f(x_1, x_2) = 0, \quad j = 1, 2.$$

Wir betrachten meist Optimierungsprobleme, bei denen wir wissen (oft aufgrund von Vorüberlegungen anhand von Graphiken), dass jede Lösung der FOC auch eine Lösung des Problems ist, d.h.: die FOC sind hinreichende Bedingungen für eine Lösung.

Nichtnegativitätsbedingungen für die Instrumentvariablen können auf diese Weise nicht explizit berücksichtigt werden. Da wir es häufig mit solchen Nebenbedingungen zu tun haben, ist es möglich, dass wir nur Lösungen der FOC finden, die diese Nebenbedingungen verletzen. In solchen Fällen sind die FOC vielleicht hinreichende Bedingungen für eine Lösung, aber keine notwendigen Bedingungen: es gibt Lösungen, die die FOC nicht erfüllen.

Fälle, in denen die FOC nicht notwendig oder nicht hinreichend für eine Lösung sind, werden kurz in den Übungen zur „Mikroökonomischen Theorie II“ besprochen. In den zahlreichen Lehrbüchern zur Mathematik für Ökonomen finden Sie dazu mehr.

**Parametrische Optimierungsprobleme:** Zielfunktionen enthalten in der Ökonomie so gut wie immer exogene Variablen oder Parameter.<sup>1</sup> Ein nichtökonomisches Beispiel:

$$\max_{x_1, x_2} f(x_1, x_2, a)$$

Ein solches Maximierungsproblem bestimmt, wenn es für alle möglichen Werte der Parameter eindeutig lösbar ist, Lösungsfunktionen  $x_j(a)$ ,  $j = 1, 2$ . Wir werden meist nur innere Lösungen betrachten und in diesem Bereich Differenzierbarkeit der Lösungsfunktionen unterstellen. Die Lösungsfunktionen erfüllen aufgrund ihrer Definition die Bedingungen 1. Ordnung:

$$D_j f [x_1(a), x_2(a), a] \equiv 0, \quad j = 1, 2.$$

**Optimalwertfunktionen:** Das gesuchte Maximum im o.a. parametrischen Optimierungsproblem hängt über die Lösungsfunktionen ebenfalls von den Parametern ab, wird also durch  $z(a) := f [x_1(a, c), x_2(a, c)]$  bestimmt. Häufig definiert man die sogenannte Optimalwertfunktion  $z$  (hier auch: Maximalwertfunktion) direkt durch

$$z(a) := \max_{x_1, x_2} f [x_1(a, c), x_2(a, c)].$$

Diese Definition hat den Vorteil, dass sie auch dann verwendbar ist, wenn es keine Lösungsfunktionen gibt, weil zu gewissen (oder allen) Werten von  $a$  mehrere Lösungen existieren (die dann natürlich alle den gleichen Maximalwert  $z(a)$  liefern müssen). Wir werden meist nur innere Lösungen betrachten und in diesem Bereich Differenzierbarkeit der Optimalwertfunktionen unterstellen.

**Einhüllendensatz:** Der Einhüllendensatz (Enveloppensatz, englisch *envelope theorem*) macht Aussagen über die erste Ableitung von Optimalwertfunktionen. Wenn wir die Definition  $z(a) := f [x_1(a, c), x_2(a, c)]$

---

<sup>1</sup>Ein Parameter ist eine Größe, die in einem Optimierungsproblem oder einer Gleichung bzw. einem Gleichungssystem unbestimmt bleibt oder konstant gehalten wird. In der Ökonomie bezeichnen wir Parameter auch als exogene Größen. Beispiel: Preise und Einkommen im Nutzenmaximierungsproblem der Haushaltstheorie würde man im Mathematiker-Jargon als Parameter des Problems bezeichnen, im Ökonomen-Jargon als exogene Größen.

auf beiden Seiten nach dem Parameter  $a$  ableiten, erhalten wir

$$\begin{aligned} z'(a) &\equiv D_1 f [x_1(a, c), x_2(a, c), a] x_1'(a) \\ &+ D_2 f [x_1(a, c), x_2(a, c), a] x_2'(a) \\ &+ D_3 f [x_1(a, c), x_2(a, c), a] \\ &\equiv D_3 f [x_1(a, c), x_2(a, c), a] . \end{aligned}$$

Der Grund dafür ist, dass die Lösungsfunktionen  $x_1(a)$ ,  $x_2(a)$  die Bedingung 1. Ordnung erfüllen:

$$D_j f [x_1(a), x_2(a), a] \equiv 0, \quad j = 1, 2 .$$

Damit entfallen die ersten beiden Summanden der Ableitung.

Das hier verwendete parametrische Maximierungsproblem ist nur ein Beispiel. Für die Anwendung auf andere Fälle ist es wichtig, die allgemeine Vorgehensweise zu verstehen. Der Einhüllendensatz sagt, dass die Ableitung der Optimalwertfunktion nach einem Parameter gleich der Ableitung der Zielfunktion nach dem Parameter ist, wobei man die Zielfunktion erst ableiten und danach die Lösungsfunktionen einsetzen kann (d.h.: man vernachlässigt die Ableitung der Lösungsfunktionen nach dem Parameter).

**Übungsbeispiel:** Wir betrachten

$$g(p) := \min_z z^2 + p^2 z .$$

Es existiert eine Lösungsfunktion  $z(p)$ . Der Einhüllendensatz liefert  $g'(p) \equiv 2pz(p)$  (Zielfunktion nach  $p$  ableiten, dann in die Ableitung Lösungsfunktion einsetzen). Wir können das überprüfen, indem wir die Lösungsfunktion und die Optimalwertfunktion berechnen:  $z(p) = -p^2/2$  und  $g(p) = -p^4/4$ , woraus wir  $g'(p) = -p^3 \equiv 2pz(p)$  erhalten.

## 4 Optimierungsprobleme mit Nebenbedingungen

**Notation:** Soll eine Zielfunktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  unter der Nebenbedingung  $g(x_1, x_2) = 0$  durch geeignete Wahl der Instrumentvariablen  $x_1, x_2$  maximiert werden, schreiben wir das wie folgt:

$$\max_{x_1, x_2} \{f(x_1, x_2) : g(x_1, x_2) = 0\}$$

Oft werden wir Nebenbedingungen wie  $x_1, x_2 \geq 0$  (Nichtnegativitätsbedingungen) nicht explizit erwähnen, weil die entsprechenden Variablen bereits als positive Variablen definiert wurden. Soll die Nichtnegativität explizit erwähnt werden, bieten sich zwei Möglichkeiten:

$$\max_{x_1, x_2 \geq 0} \{f(x_1, x_2) : g(x_1, x_2) = 0\}$$

oder

$$\max_{x_1, x_2} \{f(x_1, x_2) : g(x_1, x_2) = 0, x_1, x_2 \geq 0\}$$

**Lagrange-Ansatz:** Das Maximierungsproblem

$$\max_{x_1, x_2} \{f(x_1, x_2) : g(x_1, x_2) = 0\}$$

führt zur Lagrangefunktion

$$\mathcal{L}(x_1, x_2, \lambda) := f(x_1, x_2) - \lambda g(x_1, x_2)$$

und den FOC

$$D_j \mathcal{L}(x_1, x_2, \lambda) = 0, \quad j = 1, 2, 3.$$

Nebenbedingungen in Ungleichungsform (z.B. Nichtnegativitätsbedingungen) können im Lagrangeansatz nicht explizit berücksichtigt werden. Dazu gibt es eine Erweiterung, den sogenannten Kuhn-Tucker-Ansatz, der in der MSc-Veranstaltung „Mikroökonomie und Spieltheorie“ ausführlich behandelt wird.

Die Anmerkungen aus dem vorigen Abschnitt zu Fällen, in den die FOC nicht notwendig oder nicht hinreichend sind, gelten auch für den Lagrangeansatz.

**Parametrische Optimierungsprobleme:** Zielfunktion und/oder Nebenbedingungen enthalten in der Ökonomie so gut wie immer exogene Variablen oder Parameter. Ein nichtökonomisches Beispiel:

$$\max_{x_1, x_2} \{f(x_1, x_2, a) : g(x_1, x_2) - c = 0\}$$

Ein solches Maximierungsproblem bestimmt, wenn es für alle möglichen Werte der Parameter eindeutig lösbar ist, Lösungsfunktionen  $x_j(a, c), j = 1, 2$  (und im Rahmen des Lagrangeansatzes auch Lösungsfunktionen  $\lambda(a, c)$  für den Lagrangeparameter). Wir werden meist nur innere Lösungen betrachten und in diesem Bereich Differenzierbarkeit der Lösungsfunktionen unterstellen. Die Lösungsfunktionen erfüllen aufgrund ihrer Definition die Nebenbedingung:

$$g[x_1(a, c), x_2(a, c)] - c \equiv 0$$

Im Rahmen des Lagrangeansatzes erfüllen die Lösungsfunktionen  $x_j(a, c), j = 1, 2, \lambda(a, c)$  die FOC:

$$D_j \mathcal{L}[x_1(a, c), x_2(a, c), \lambda(a, c)] \equiv 0, \quad j = 1, 2, 3.$$

**Optimalwertfunktionen:** Das gesuchte Maximum im o.a. parametrischen Optimierungsproblem hängt über die Lösungsfunktionen ebenfalls von den Parametern ab, wird also durch  $z(a, c) := f[x_1(a, c), x_2(a, c)]$  bestimmt. Häufig definiert man die sogenannte Optimalwertfunktion  $z$  (hier auch: Maximalwertfunktion) direkt durch

$$z(a, c) := \max_{x_1, x_2} \{f(x_1, x_2, a) : g(x_1, x_2) - c = 0\}.$$

Diese Definition hat den Vorteil, dass sie auch dann verwendbar ist, wenn es keine Lösungsfunktionen gibt, weil zu gewissen (oder allen) Werten von  $a, c$  mehrere Lösungen existieren (die dann natürlich alle den gleichen Maximalwert  $z(a, c)$  liefern müssen). Wir werden meist nur innere Lösungen betrachten und in diesem Bereich Differenzierbarkeit der Optimalwertfunktionen unterstellen.

**Einhüllendensatz:** Der Einhüllendensatz lässt sich auch auf Optimierungsprobleme mit Nebenbedingungen anwenden. Man betrachtet die Lagrangefunktion des o.a. parametrisierten Optimierungsproblems jetzt als Funktion nicht nur der Instrumentvariablen und des Lagrangeparameters, sondern zusätzlich auch als Funktion der Parameter:

$$\mathcal{L}(x_1, x_2, \lambda, a, c) := f(x_1, x_2, a) - \lambda [g(x_1, x_2) - c]$$

Wenn man in die Lagrangefunktion die Lösungsfunktionen  $x_1(a, c)$ ,  $x_2(a, c)$ ,  $\lambda(a, c)$  einsetzt, erhält man die Optimalwertfunktion:

$$\begin{aligned} \mathcal{L}[x_1(a, c), x_2(a, c), \lambda(a, c), a, c] &:= \\ f[x_1(a, c), x_2(a, c), a] - \lambda(a, c) \{g[x_1(a, c), x_2(a, c)] - c\} &\equiv z(a, c) \end{aligned}$$

Der Grund dafür ist, dass die Lösungsfunktionen  $x_1(a, c)$ ,  $x_2(a, c)$  die Nebenbedingung erfüllen:  $g[x_1(a, c), x_2(a, c)] - c \equiv 0$ . Wenn wir die Identität auf beiden Seiten nach dem Parameter  $a$  ableiten, erhalten wir

$$\begin{aligned} D_1 z(a, c) &\equiv D_1 \mathcal{L}[x_1(a, c), x_2(a, c), \lambda(a, c), a, c] D_1 x_1(a, c) \\ &+ D_2 \mathcal{L}[x_1(a, c), x_2(a, c), \lambda(a, c), a, c] D_1 x_2(a, c) \\ &+ D_3 \mathcal{L}[x_1(a, c), x_2(a, c), \lambda(a, c), a, c] D_1 \lambda(a, c) \\ &+ D_4 \mathcal{L}[x_1(a, c), x_2(a, c), \lambda(a, c), a, c] \\ &\equiv D_4 \mathcal{L}[x_1(a, c), x_2(a, c), \lambda(a, c), a, c] . \end{aligned}$$

Der Grund dafür ist, dass die Lösungsfunktionen  $x_1(a, c)$ ,  $x_2(a, c)$ ,  $\lambda(a, c)$  die FOC erfüllen:

$$D_j \mathcal{L}[x_1(a, c), x_2(a, c), \lambda(a, c), a, c] \equiv 0, \quad j = 1, 2, 3.$$

Damit entfallen die ersten drei Summanden der Ableitung. Entsprechend erhält man

$$D_2 z(a, c) \equiv D_5 \mathcal{L}[x_1(a, c), x_2(a, c), \lambda(a, c), a, c] \equiv \lambda(a, c).$$

Das hier verwendete parametrische Maximierungsproblem ist nur ein Beispiel. Für die Anwendung auf andere Fälle ist es wichtig, die allgemeine Vorgehensweise zu verstehen. Der Einhüllendensatz sagt, dass die Ableitung der Optimalwertfunktion nach einem Parameter gleich der Ableitung der Lagrangefunktion nach dem Parameter ist, wobei man die Lagrangefunktion erst ableiten und danach die Lösungsfunktionen einsetzen kann (d.h.: man vernachlässigt die Ableitung der Lösungsfunktionen nach dem Parameter).

**Übungsbeispiel:** Wir betrachten

$$h(q) := \min_{z_1, z_2 \geq 0} \{z_1^2 + z_2^2 : z_1 + qz_2 = 5\} .$$

Die Lagrangefunktion ist

$$\mathcal{L}(z_1, z_2, \mu, q) := z_1^2 + z_2^2 - \mu [z_1 + qz_2 - 5] .$$

Es existieren Lösungsfunktionen  $z_1(q)$ ,  $z_2(q)$  und  $\mu(q)$ . Der Einhüllendensatz liefert  $h'(q) \equiv -\mu(q)z_2(q)$  (Lagrangefunktion nach  $q$  ableiten, dann in die Ableitung Lösungsfunktionen einsetzen). Wir können das überprüfen, indem wir die Lösungsfunktionen und die Optimalwertfunktion berechnen: Aus  $\mu(q) = 10/(1 + q^2)$ ,  $z_1(q) = 5/(1 + q^2)$ ,  $z_2(q) = 5q/(1 + q^2)$ ,  $h(q) = 25/(1 + q^2)$  erhalten wir  $h'(q) = -50q/(1 + q^2)^2 \equiv -\mu(q)z_2(q)$ .