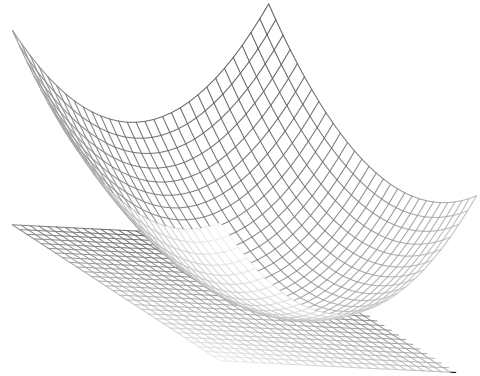


# Konvexe Optimierung

Kuhn, Tucker (1950), Karush (1939), John (1948)

Um eine Theorie der nichtlinearen Optimierung zu entwickeln, wird die Klasse der linearen Funktionen zunächst geringfügig erweitert. Hierbei rücken die konvexen Funktionen in den Vordergrund. Ihre Graphen verlaufen oberhalb ihrer Tangentialebenen. Eine Definition wird hier nicht benötigt. In Zweifelsfällen kann mit der Matrix der zweiten partiellen Ableitungen die Krümmung untersucht werden (Matrix ist positiv definit, d. h.  $f_{xx} f_{yy} - f_{xy}^2 > 0$ ,  $f_{xx} > 0$ , siehe Extrema von Funktionen mit zwei Variablen). In der  $xy$ -Ebene werden durch die Niveaulinien  $g(x, y) = 0$  konvexer Funktionen konvexe Bereiche festgelegt, d. h. zu je zwei Punkten enthalten diese Bereiche auch die Verbindungsstrecke.



Untersuchen wir das Minimum der Funktion

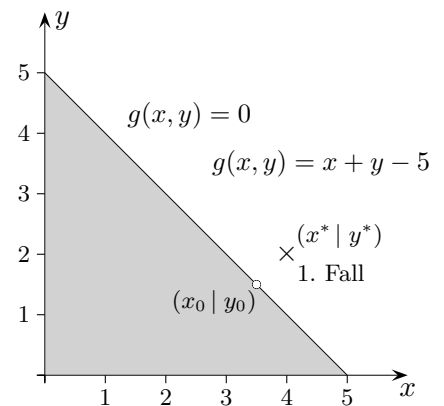
$$f(x, y) = (x - x^*)^2 + (y - y^*)^2$$

im Bereich  $x + y \leq 5$ ;  $x, y \geq 0$ .

Hierbei werden wir  $x^*, y^*$  variieren.

Es ist überraschend, dass alle denkbaren Fälle in einer handlichen Bedingung zusammengefasst werden können.

Der Bereich hat auf das Minimum nur dann einen Einfluss, wenn  $(x^* | y^*)$  außerhalb liegt. Diesen Fall betrachten wir zuerst.



1.  $(x^* | y^*) = (4 | 2)$

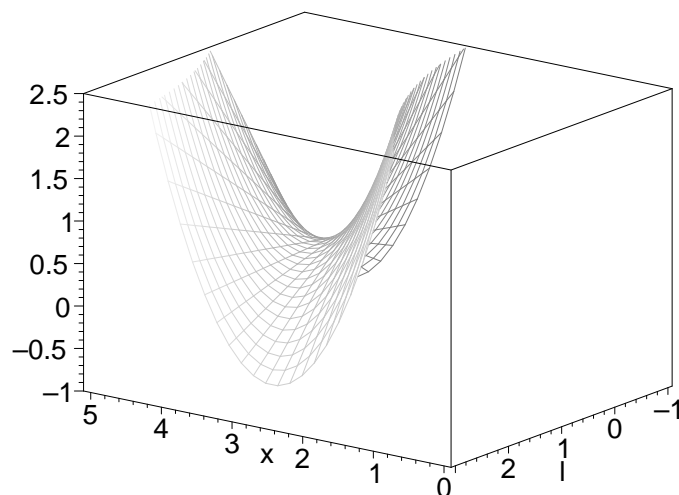
Das Minimum wird an der Stelle  $(\frac{7}{2} | \frac{3}{2})$  auf der Restriktionsgeraden  $g(x, y) = 0$  angenommen.

Dies kann mit der Lagrange-Funktion  $L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$  errechnet werden,  $\lambda = 1$ ,  $f(x_0, y_0) = \frac{1}{2}$ .  $g(x, y) = -x - y + 5$  hätte  $\lambda = -1$  zur Folge.

Das Ergebnis wird durch eine - wie wir noch sehen werden typische - Ungleichung bestätigt:

$$\underbrace{L\left(\frac{7}{2}, \frac{3}{2}, \lambda\right)}_{\frac{1}{2}} \leq \underbrace{L\left(\frac{7}{2}, \frac{3}{2}, 1\right)}_{\frac{1}{2}} \leq \underbrace{L(x, y, 1)}_{\frac{1}{2}} = (x - 4)^2 + (y - 2)^2 + x + y - 5 = \left(x - \frac{7}{2}\right)^2 + \left(y - \frac{3}{2}\right)^2 + \frac{1}{2}$$

Um die Ungleichung zu veranschaulichen, stellen wir  $L(x, \frac{3}{2}, \lambda)$  grafisch dar.



$(x_0, y_0, \lambda_0)$  mit  $x_0, y_0, \lambda_0 \geq 0$  heißt Sattelpunkt von  $L$ , falls für  $x, y, \lambda \geq 0$  gilt:  $L(x_0, y_0, \lambda) \leq L(x_0, y_0, \lambda_0) \leq L(x, y, \lambda_0)$ .

Diese Definition und die Theorie werden in naheliegender Weise verallgemeinert, um sie auch in unanschaulichen Situationen anwenden zu können. In der Ökonomie werden diese Sattelpunkte als Gleichgewichtszustände interpretiert.

Falls bei unserem Beispiel  $(x^* | y^*)$  innerhalb des Bereichs ( $g(x, y) < 0$ ) oder auf dem Rand (z.B.  $g(x, y) = 0$  oder  $x = 0$ ) liegt, ist für  $\lambda = 0$  die Sattelpunktbedingung erfüllt.

2.  $(x^* | y^*) = (1 | 2)$

Das Minimum  $(x_0, y_0)$  wird an derselben Stelle angenommen. Es ist dann:

$$\underbrace{L(x_0, y_0, \lambda)}_{-2\lambda} \leq L(x_0, y_0, 0) \leq L(x, y, 0).$$

In die linke Abschätzung gehen wesentlich  $\lambda \geq 0$  und  $g(x_0, y_0) < 0$  ein.

Die Bedeutung der Sattelpunkte für Optimierungsprobleme kann nun erahnt werden. Der folgende Satz ist leicht einzusehen. Die Konvexität wird erst im Beweis der Umkehrung benötigt. Für differenzierbare Funktionen wird die Sattelpunktbedingung im *Kuhn-Tucker-Satz* durch äquivalente Bedingungen ersetzt.

Ist  $(x_0, y_0, \lambda_0)$  ein Sattelpunkt der Lagrange-Funktion  $L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$ , so löst  $(x_0, y_0)$  die Aufgabe  $f(x, y) \rightarrow \text{Min}, g(x, y) \leq 0, x, y \geq 0$ .

Sei also  $f(x_0, y_0) + \lambda g(x_0, y_0) \leq f(x_0, y_0) + \underbrace{\lambda_0 g(x_0, y_0)}_{= 0} \leq f(x, y) + \underbrace{\lambda_0 g(x, y)}_{\leq 0 \text{ (Voraussetzung)}}$ .

Die Behauptung folgt aus:

Der linken Ungleichung ist  $\lambda g(x_0, y_0) \leq \lambda_0 g(x_0, y_0)$  zu entnehmen.

Dies ist für  $\lambda \geq 0$  nur möglich, falls  $g(x_0, y_0) \leq 0$  ist.

Für  $\lambda = 0$  erhalten wir aus der linken Ungleichung  $0 \leq g(x_0, y_0)$ ,

insgesamt daher  $\lambda_0 g(x_0, y_0) = 0$ . □

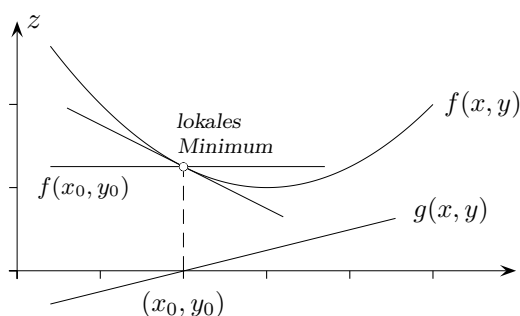
Die Umkehrung wird plausibel erscheinen.

Zu einer Lösung  $(x_0, y_0)$  der Aufgabe  $f(x, y) \rightarrow \text{Min}, g(x, y) \leq 0, x, y \geq 0, f, g$  konvex, gibt es ein  $\lambda_0 \geq 0$ , so dass  $(x_0, y_0, \lambda_0)$  Sattelpunkt der Lagrange-Funktion ist.

Liegt  $(x_0, y_0)$  innerhalb des Bereichs, so kann  $\lambda = 0$  gewählt werden.

Das lokale Minimum ist ein globales ( $f$  konvex):  $f(x_0, y_0) \leq f(x, y)$  für  $x, y \geq 0$ .

Der einzig interessante Fall wird durch die zweidimensionale Schnitt-Grafik beschrieben.



Das globale Minimum liegt außerhalb des Bereichs. Aus der Grafik ist ersichtlich, dass es ein  $\lambda_0 \geq 0$  gibt mit:  $f(x_0, y_0) \leq f(x, y) + \lambda_0 g(x, y)$  für  $x, y \geq 0$ . Rechts von  $(x_0, y_0)$  muss zu  $f(x, y)$  etwas addiert werden, links darf etwas subtrahiert werden ( $g(x, y) \leq 0$ ). Hier ist zu erkennen, wie relevant die Konvexität ist. Im Beweis des Satzes werden disjunkte konvexe Mengen durch eine Ebene getrennt.

Nun können die Kuhn-Tucker-Bedingungen ermittelt werden.

$f, g$  seien konvex und partiell differenzierbar.

$L(x, y, \lambda_0)$  ist wegen  $\lambda_0 \geq 0$  als Summe konvexer Funktionen konvex,  $x, y \geq 0$ ,

$$\implies L(x_0, y_0, \lambda_0) + (x - x_0) \cdot L_x(x_0, y_0, \lambda_0) + (y - y_0) \cdot L_y(x_0, y_0, \lambda_0) \leq L(x, y, \lambda_0)$$

Falls nun für  $(x_0, y_0, \lambda_0)$  gilt:

$$\begin{aligned} L_x &\geq 0, & L_y &\geq 0 \\ x_0 L_x &= 0, & y_0 L_y &= 0 \end{aligned} \quad \text{so folgt} \quad L(x_0, y_0, \lambda_0) \leq L(x, y, \lambda_0).$$

Mit  $\lambda_0 g(x_0, y_0) = 0$

$$g(x_0, y_0) \leq 0 \quad \text{folgt} \quad L(x_0, y_0, \lambda) \leq L(x_0, y_0, \lambda_0) = f(x_0, y_0) + \lambda_0 g(x_0, y_0).$$

Aus diesen Bedingungen ergibt sich somit ein Sattelpunkt und damit eine Lösung der Aufgabe  $f(x, y) \rightarrow \text{Min}, g(x, y) \leq 0, x, y \geq 0$ .

Umgekehrt folgen aus der Sattelpunktbedingung

$$L(x_0, y_0, \lambda) \leq L(x_0, y_0, \lambda_0) \leq L(x, y, \lambda_0), \quad x, y, \lambda \geq 0$$

die Kuhn-Tucker-Bedingungen.

$L_x \geq 0, L_y \geq 0$  ergeben sich aus  $L(x_0, y_0, \lambda_0) \leq L(x, y, \lambda_0)$ .

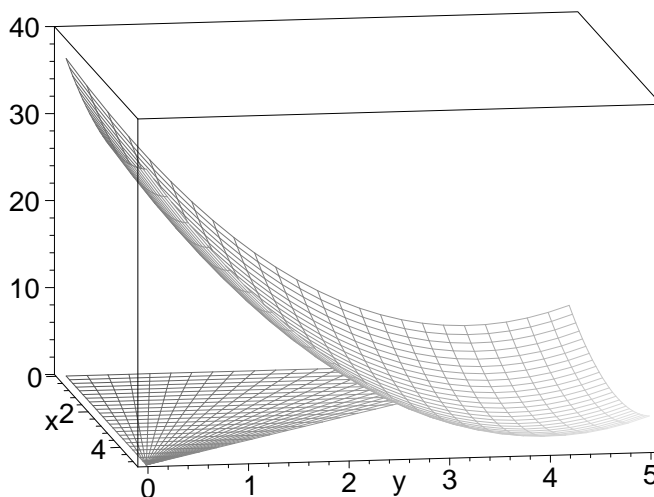
Aus dem Beweis auf der vorigen Seite entnehmen wir  $\lambda_0 g(x_0, y_0) = 0$  und  $g(x_0, y_0) \leq 0$ .

Wäre  $x_0 L_x > 0$  ( $y_0 L_y > 0$ ), so müssten beide Faktoren größer null sein und dies widerspräche  $L(x_0, y_0, \lambda_0) \leq L(x, y, \lambda_0)$ , da dann eine Stelle mit einem Funktionswert gefunden werden könnte, der kleiner als  $L(x_0, y_0, \lambda_0)$  wäre.

Damit wir bei der Anwendung des Kuhn-Tucker-Satzes nicht die Freude an der Mathematik verlieren, verwenden wir Maple. Zuerst wird gezeigt, wie Konvexität bequem nachgewiesen werden kann.

```
f := (x, y) -> 1/2 * (x - 3)^2 + 2 * (y - 4)^2;
fxx := diff(f(x, y), x$2);
fyy := diff(f(x, y), y$2);
fxy := diff(f(x, y), x$1, y$1);
fxx * fyy - fxy^2 > 0;
```

Ergebnis:  $fxx := 1$   
 $0 < 4$



```
g := (x, y) -> x + y - 5; # x + y ≤ 5
L := (x, y, u) -> f(x, y) + u * g(x, y);
Lx := diff(L(x, y, u), x);
Ly := diff(L(x, y, u), y);
Lu := diff(L(x, y, u), u);
s := solve({x * Lx = 0, y * Ly = 0, u * g(x, y) = 0, x >= 0, y >= 0, u >= 0}, {x, y, u});
```

Maple liefert:

```
s := {x = 0, u = 0, y = 0}, {u = 0, y = 0, x = 3},
      {x = 0, u = 0, y = 4}, {u = 0, x = 3, y = 4},
      {x = 7/5, u = 8/5, y = 18/5}
```

Nun sind die Ergebnisse einzeln ( $n$  von 1 bis 5) mit

$subs(s[n], Lx >= 0); subs(s[n], Ly >= 0); subs(s[n], g(x, y) <= 0);$

zu überprüfen.

```
n := 5:      entlarvt mit      0 ≤ 0
                                     0 ≤ 0
                                     0 ≤ 0      die Lösung {x = 7/5, u = 8/5, y = 18/5}
```

Mit Maple kann dieses Resultat auch ohne Umschweife erreicht werden:

```
s := solve({x * Lx = 0, y * Ly = 0, u * g(x, y) = 0, Lx >= 0, Ly >= 0, g(x, y) <= 0,
            x >= 0, y >= 0, u >= 0}, {x, y, u});
```

Die Änderung  $x + y \geq 5$  (d.h.  $g(x, y) = 5 - x - y$ ) des Bereichs führt zur erwarteten Lösung  $\{x = 3, y = 4, u = 0\}$ .