Kuhn-Tucker-Bedingungen

Ist x_* ein lokales Minimum einer stetig differenzierbaren Funktion $f : \mathbb{R}^n \ni D \to \mathbb{R}$ auf der durch die Ungleichungen $g_k(x_1, \ldots, x_n) \geq 0$ definierten Menge D und sind die Gradienten (ebenfalls als stetig vorausgesetzt) der aktiven Gleichungen $g_k(x_*) = 0$, $k \in I$, linear unabhängig, dann existieren Lagrange-Multiplikatoren $\lambda_k \geq 0$, so dass

$$\operatorname{grad} f(x_*) = \sum_{k \in I} \lambda_k \operatorname{grad} g_k(x_*)\,.$$

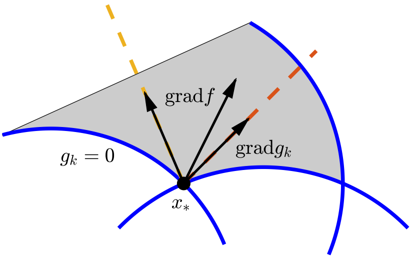
Für ein lokales Maximum ist entsprechend $\lambda_k \leq 0$.

Die Indexmenge I lässt sich auch implizit durch die Bedingungen

$$\lambda_k g_k(x_*) = 0$$

festlegen. Ist $g_k(x_*) > 0$, so folgt $\lambda_k = 0$, d.h. die nichttrivialen Multiplikatoren λ_k entsprechen den aktiven Nebenbedingungen ($g_k(x_*) = 0$).

Geometrisch bedeutet die Kuhn-Tucker-Bedingung für ein Minimum, dass der Gradient der Zielfunktion f in dem durch die Gradienten der aktiven Nebenbedingungen aufgespannten Kegel liegt. Dies ist in der Abbildung für bivariate Funktionen veranschaulicht. Bei einem Minimum x_* in einer Ecke liegt grad f in dem durch die relevanten Gradientenrichtungen (gestrichelt) begrenzten Sektor.



Eine Gleichungsnebenbedingung $h(x_1, \ldots, x_n) = 0$ kann durch die Ungleichungen $h \geq 0 \land -h \geq 0$ ersetzt werden. Damit ist das Kuhn-Tucker-Kriterium auch in diesem Fall anwendbar. Der entsprechende Term

$$\lambda_+ \operatorname{grad} h + \lambda_- \operatorname{grad}(-h) = \lambda \operatorname{grad} h$$

besitzt einen Lagrange-Multiplikator $\lambda = \lambda_+ - \lambda_-$, dessen Vorzeichen nicht eingeschränkt ist.

Beweis

Die inaktiven Nebenbedingungen ($g_k(x_*) > 0$, $k \notin I$) sind irrelevant, da sie in der Umgebung von x_* keine Einschränkung bedeuten.
$\rightsquigarrow$ Minimalität von f ebenfalls auf der kleineren Menge, die durch die Gleichungsbedingungen $g_k(x) = 0$ beschrieben wird

Satz über Lagrange-Multiplikatoren $\implies$ behauptete Identität mit $\lambda_k \in \mathbb{R}$

zu zeigen: $\lambda_k \geq 0$
Indirekter Beweis: Annahme $\lambda_\ell < 0$ für ein $\ell \in I$
lineare Unabhängigkeit der Gradienten $\implies$

$$\exists v : (\operatorname{grad} g_\ell(x_*))^{\mathrm{t}}\, v = 1, \quad (\operatorname{grad} g_k(x_*))^{\mathrm{t}}\, v = 0,\ k \in I \backslash \ell\,,$$

d.h. v liegt in der Tangentialebene der durch g_k, $k \neq \ell$, definierten Fläche S und hat eine nichttriviale Komponente in Richtung der Normalen der Fläche $S_\ell :\ g_\ell(x) = 0$

wähle eine Kurve $t \mapsto x(t) \in S$ mit Anfangspunkt $x(0) = x_*$ und Anfangsrichtung $x'(0) = v \quad \rightsquigarrow$

$$g_\ell(x(t)) = g_\ell(x_*) + \left((\operatorname{grad}\, g_\ell(x_*))^{\mathrm{t}}\, v\right)\, t + O(t^2) = 0 + t + O(t^2),\ \ t \to 0$$

$\implies \quad x(t)$ für hinreichend kleines $t > 0$ zulässig:

$$g_\ell(x(t)) \geq 0, \quad g_k(x(t)) = 0,\ k \in I \setminus \ell$$

Konstruktion von v und Identität für grad $f \quad \implies$

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} f(x(t))_{|t=0} &= (\operatorname{grad} f(x(t)))^{\mathrm{t}} x'(t)_{|t=0} = (\operatorname{grad}\, f(x_*))^{\mathrm{t}} v \\ &= \sum_{k \in I} \lambda_k (\operatorname{grad} g_k(x_*))^{\mathrm{t}} v = \lambda_\ell (\operatorname{grad} g_\ell(x_*))^{\mathrm{t}} v = \lambda_\ell < 0 \end{aligned}$$

$\implies \quad$ Abnahme von f entlang der Kurve, im Widerspruch zur Minimalität von $f(x_*)$

Beispiel

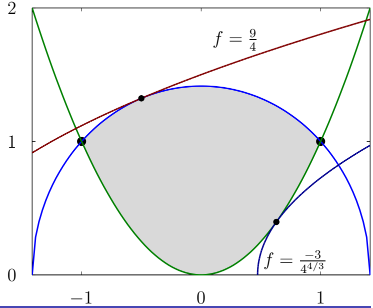
Extrema der Funktion

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

auf der durch

$$\begin{array}{rcccc} g_1(x,y) & = & y - x^2 & \geq & 0 \\ g_2(x,y) & = & 2 - x^2 - y^2 & \geq & 0 \end{array}$$

definierten Menge D



Kuhn-Tucker-Bedingungen

$$\operatorname{grad} f = (-1, 2y)^{\mathrm{t}} = \lambda \underbrace{(-2x, 1)^{\mathrm{t}}}_{\operatorname{grad} g_1} + \varrho \underbrace{(-2x, -2y)^{\mathrm{t}}}_{\operatorname{grad} g_2}$$

$$\lambda \underbrace{(y - x^2)}_{g_1} = 0, \quad \varrho \underbrace{(2 - x^2 - y^2)}_{g_2} = 0$$

mit Lagrange-Multiplikatoren λ und ϱ gleichen Vorzeichens

lineare Unabhängigkeit der Gradienten der aktiven Nebenbedingungen für alle zulässigen Punkte (nicht Null und an den Schnittpunkten der Randkurven nicht parallel)
$\implies$ Notwendigkeit der Kuhn-Tucker-Bedingung für alle Extrema
verschiedene Fälle je nachdem welche Nebenbedingungen aktiv sind
(i) Keine Nebenbedingung aktiv:

$$\implies \quad \lambda = \varrho = 0, \quad (-1, 2y) = (0, 0)$$

nicht erfüllbar, keine Extrema von f im Inneren von D
(ii) g_1 aktiv, d.h. $y - x^2 = 0$ und $g_2(x, y) > 0$:

$$\implies \quad \varrho = 0, \quad (-1, 2y) = \lambda(-2x, 1)$$

$\implies$ $\lambda = 2y = 2x^2 \geq 0$ und

$$-1 = (2x^2)(-2x) \quad \iff \quad x = 4^{-1/3}, \quad y = 4^{-2/3}$$

$\implies$ Kuhn-Tucker-Bedingung für ein lokales Minimum

(iii) g_2 aktiv, d.h. $2 - x^2 - y^2 = 0$, $g_1(x,y) > 0$:

$$\implies \quad \lambda = 0, \quad (-1, 2y) = \varrho(-2x, -2y)$$

$\implies$ $\varrho = -1 < 0$ ($y = 0$ wegen $(\pm\sqrt{2}, 0) \notin D$ nicht möglich) und

$$x = -1/2, \quad y = \sqrt{7}/2$$

$\implies$ Kuhn-Tucker-Bedingung für ein lokales Maximum

(iv) g_1 und g_2 aktiv, d.h. $y - x^2 = 0$ und $2 - x^2 - y^2 = 0$:
$\implies$ $(x,y) = (1,1)$ oder $(x,y) = (-1,1)$
Einsetzen in die Gleichung für grad f $\rightsquigarrow$

$$(-1, 2) = \lambda(-2, 1) + \varrho(-2, -2) \quad \implies \quad \lambda = 1, \quad \varrho = -1/2$$

bzw.

$$(-1, 2) = \lambda(2, 1) + \varrho(2, -2) \quad \implies \quad \lambda = 1/3, \quad \varrho = -5/6$$

keine Extremstellen wegen der verschiedenen Vorzeichen der Lagrange-Multiplikatoren

Existenz von Extrema auf der kompakten zulässigen Menge $\implies$
Minimum im Fall (ii) bei $(4^{-1/3}, 4^{-2/3})$,
Maximum im Fall (iii) bei $(-1/2, \sqrt{7/2})$

alternative Typbestimmung durch Vergleich der Funktionswerte

$$f(4^{-1/3}, 4^{-2/3}) = -3\sqrt[3]{2}/8 < 9/4 = f(-1/2, \sqrt{7}/2)$$

Beispiel

Minimierung einer Funktion $f : \mathbb{R}^n \to \mathbb{R}$ auf einem Rechteck

$$D : \; a_k \leq x_k \leq b_k, \quad k = 1, \ldots, n$$

Kuhn-Tucker-Bedingung für ein lokales Minimum

$$\operatorname{grad} f(x_*) = \sum_{k=1}^{n} (\lambda_k - \varrho_k)\, e_k = (\lambda_1 - \varrho_1, \ldots, \lambda_n - \varrho_n)^{\mathrm{t}}$$
$$\lambda^{\mathrm{t}} (x_* - a) = 0, \; \varrho^{\mathrm{t}} (b - x_*) = 0, \quad \lambda_k, \; \varrho_k \geq 0$$

mit e_k dem k-ten Einheitsvektor

$\pm\infty$ als Intervallgrenzen zugelassen

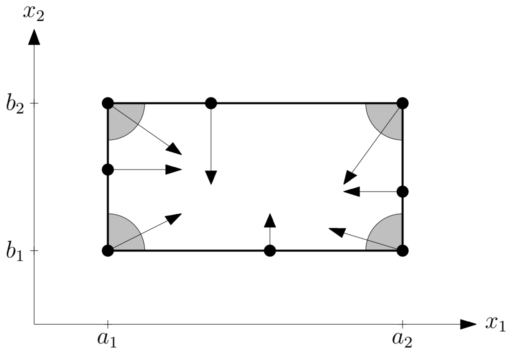
z.B. $a_k = -\infty \quad \Longrightarrow \quad \lambda_k = 0$

$\rightsquigarrow$ keine Einschränkung in den Kuhn-Tucker-Bedingungen

(i) $a_k < x_{*,k} < b_k$:
Lagrange-Multiplikatoren λ_k und ϱ_k Null und damit auch die k-te Komponente g_k von $\operatorname{grad} f(x_*)$

(ii) Eine der Ungleichungen für x_k aktiv:
$\rightsquigarrow$ entsprechender Lagrange-Multiplikator bestimmt Vorzeichen von g_k:

$$\begin{aligned} x_{*,k} = a_k \quad &\Longrightarrow \quad \varrho_k = 0,\ g_k = \lambda_k \geq 0 \\ x_{*,k} = b_k \quad &\Longrightarrow \quad \lambda_k = 0,\ g_k = -\varrho_k \leq 0 \end{aligned}$$



mögliche Richtungen von grad $f(x_*)$ für lokale Minima von bivariaten Zielfunktionen auf einem Rechteck $[a_1, b_1] \times [a_2, b_2]$

- im Innern: keine Einschränkung für grad f
- auf einer Kante: grad f orthogonal
- an einer Ecke: grad f zeigt ins Innere

Beispiel

Lineares Programm:
Minimierung einer linearen Funktion

$$c^{\mathrm{t}}x = \sum_{k=1}^{n} c_k x_k \to \min$$

unter den linearen Nebenbedingungen

$$Ax \geq b$$

mit einer $m \times n$-Matrix A

Kuhn-Tucker-Bedingungen für ein lokales Minimum $x_* \in \mathbb{R}^n$

$$c^{\mathrm{t}} = \lambda^{\mathrm{t}} A, \quad \lambda^{\mathrm{t}}(Ax_* - b) = 0$$

mit $\lambda_k \geq 0$ (entsprechend $\lambda_k \leq 0$ für ein Maximum) z.B.

$$x + y \to \min, \qquad x \geq 2, \qquad y \geq 1, \qquad x + 2y \geq 8\,,$$

d.h.

$$c = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 1 \\ 8 \end{pmatrix}$$

$\Longrightarrow$ Kuhn-Tucker-Bedingungen

$$1 = \lambda_1 + \lambda_3, \quad 1 = \lambda_2 + 2\lambda_3, \quad \lambda_1[x-2] + \lambda_2[y-1] + \lambda_3[x+2y-8] = 0$$

mit $\lambda_k, [\ldots] \geq 0$

Bedingungen $\Longrightarrow$ genau einer der Lagrange-Multiplikatoren λ_k gleich null

$\rightsquigarrow$ drei Fälle (jeweils zwei Nebenbedingungen aktiv)

(i) $\quad \lambda_1 = 0$:

$\implies \quad \lambda_3 = 1,\ \lambda_2 = -1$

kein Extremum, da verschiedene Vorzeichen

(ii) $\quad \lambda_2 = 0$:

$\implies \quad \lambda_3 = 1/2,\ \lambda_1 = 1/2$

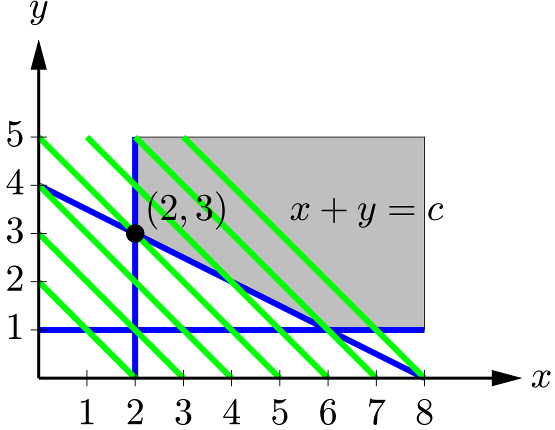
aktive Nebenbedingungen $x = 2,\ x + 2y = 8 \quad \rightsquigarrow \quad (x, y) = (2, 3)$

Kuhn-Tucker-Bedingungen für ein Minimum erfüllt

(iii) $\quad \lambda_3 = 0$:

aktive Nebenbedingungen $x = 2,\ y = 1 \quad \rightsquigarrow$ Punkt ausserhalb des zulässigen Bereichs

Beschränkheit von f nach unten auf dem zulässigen Bereich $\implies$ globales Minimum bei $(2, 3)$



geometrische Konstruktion der Lösung:
Niveaulinie der Zielfunktion berührt den zulässigen Bereich in $(2, 3)$
Zielfunktion steigt (fällt), wenn die Niveaugeraden den zulässigen Bereich schneiden (nicht schneiden)