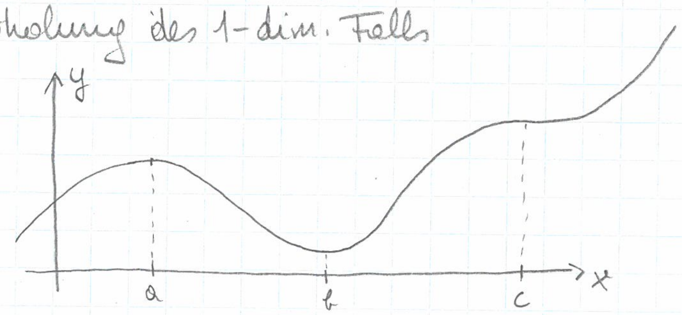
Nicht-lineare Optimierung (mit Nebenbedingungen)

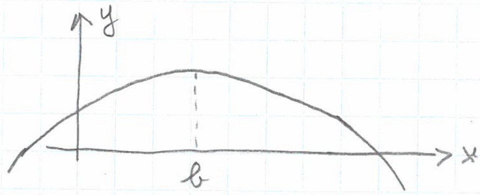
▷ Wiederholung des 1-dim. Falls



lokale Optima: Bei a ist ein lokales Maximum, b lokales Minimum, c Sattelpunkt.

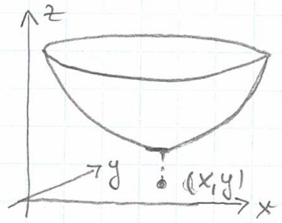
An allen drei Stellen gilt $y'(x) = 0$ bzw. $dy = 0$.
Solche Stellen heißen kritische Punkte oder Extremstellen.

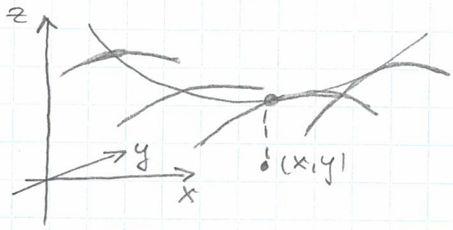
globale Optima:



▷ 2-dim. Fall

Beispiele:





(x,y) ist ein kritischer Punkt der Fkt. $z(x,y)$ wenn

$z'(x,y) = 0$ gleichbedeutend mit $dz = 0$ bei (x,y)

$\mathrm{grad}(z) = 0$ – " –

$\nabla z = 0$ – " –

$\frac{\partial z}{\partial x} = 0$ und $\frac{\partial z}{\partial y} = 0$ – " –

Woher weiß man, ob ein kritischer Pkt. ein lokales Minimum oder ein lokales Maximum oder ein Sattelpunkt ist?

1-dim: $y''(x) \lesseqgtr 0$ 2-dim.: $\underbrace{H(x,y)}_{\text{Matrix}} \lesseqgtr \underbrace{0}_{\text{Zahl}}$? ↯

Die Hessematrix $H(x,y)$ ist symmetrisch, d.h. $H(x,y)^T = H(x,y)$, weil $\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial z}$.

Jede symmetrische Matrix hat reelle Eigenwerte (und die Eigenvektoren zu unterschiedlichen Eigenwerten sind orthogonal.)

Wenn alle Eigenwerte von $H(x,y)$ größer Null sind, dann heißt $H(x,y)$ positiv definit bei (x,y).

Analog: EW ≥ 0 .... positiv semidefinit

EW < 0 .... negativ definit

EW ≤ 0 .... negativ semidefinit

Rest ..... indefinit

Optimalitätskriterien:

Notwendige Bedingung: Falls z bei (x,y) ein lokales Minimum (Maximum) hat, dann ist $\nabla z(x,y) = 0$ und $H(x,y)$ positiv semidefinit (negativ semidefinit).

Hinreichende Bedingung: Falls $\nabla z(x,y) = 0$ und $H(x,y)$ positiv definit (negativ definit), dann hat z bei (x,y) ein lokales Minimum (Maximum)

▷ n-dim. Fall: völlig analog

$$z'(x_1,\dots,x_n) = \left(\frac{\partial z}{\partial x_1}, \dots, \frac{\partial z}{\partial x_n}\right) = \operatorname{grad}(z)(x_1,\dots,x_n)^T$$

$$H(x_1,\dots,x_n) = \begin{pmatrix} \frac{\partial^2 z}{\partial x_1^2} & \cdots & \frac{\partial^2 z}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 z}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 z}{\partial x_n^2} \end{pmatrix} (x_1,\dots,x_n)$$

▷ Warum sind die Eigenwerte der Hessematrix entscheidend?

Taylorreihe bei einem (lokalen) Optimum

$$z = z_0 + \underbrace{\left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right)}_{= \nabla z^T = 0} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} + \frac{1}{2} (\Delta x, \Delta y) \underbrace{\begin{pmatrix} \frac{\partial^2 z}{\partial x^2} & \frac{\partial^2 z}{\partial x \partial y} \\ \frac{\partial^2 z}{\partial y \partial x} & \frac{\partial^2 z}{\partial y^2} \end{pmatrix}}_{H} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} + \text{Fehler}$$

H mit EW λ_1 und λ_2 und orthogonalen EV v_1 und v_2

Jeder Vektor $\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$ läßt sich als Linearkombination der EV schreiben:

$$\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = c_1 v_1 + c_2 v_2 \quad \text{mit Koeff. } c_1 \text{ und } c_2 \text{ aus } \mathbb{R}.$$

Dann wird die Taylorreihe zu

$$\begin{aligned} z &= z_0 + \frac{1}{2} (c_1 v_1 + c_2 v_2)^T H (c_1 v_1 + c_2 v_2) + \ldots \\ &= z_0 + \frac{1}{2} (c_1 v_1^T + c_2 v_2^T)(c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2) + \ldots \\ &= z_0 + \frac{1}{2} (c_1^2 \lambda_1 v_1^T v_1 + 0 + 0 + c_2 \lambda_2 v_2^T v_2) + \ldots \\ &= z_0 + \frac{1}{2} (\lambda_1 \underbrace{c_1^2 \|v_1\|^2}_{\geq 0} + \lambda_2 \underbrace{c_2^2 \|v_2\|^2}_{\geq 0}) + \ldots \end{aligned}$$

Wenn beide EW positiv $\Rightarrow$ $z > z_0$ für kleine $\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$: Minimum

— " — negativ $\Rightarrow$ $z < z_0$ — " — : Maximum

▷ Eine symmetrische Matrix (wie die Hessematrix) ist genau dann positiv definit, wenn alle führenden Hauptminoren positiv sind:

Bsp.: $H = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$ hat die führenden Hauptminoren

$\det(1)$, $\det\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, $\det\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$

Allg.:

Wenn alle führend. Hauptm. von H > 0, dann ist H pos. def.

$H \geq 0$ — H pos. sem. def.

$-H > 0$ — H neg. def.

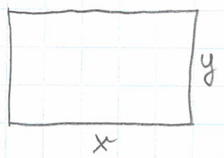
$-H \geq 0$ — H neg. sem. def.

H pos. def. $\Leftrightarrow$ $-H$ neg. def

H pos. sem. def. $\Leftrightarrow$ $-H$ neg. sem. def.

mit Nebenbedingungen:

Bsp: Finde jenes Rechteck mit Umfang $u = 12\,m$, das maximalen Flächeninhalt hat.



$z = x y \qquad u = 2x + 2y = 12$

$x + y = 6$

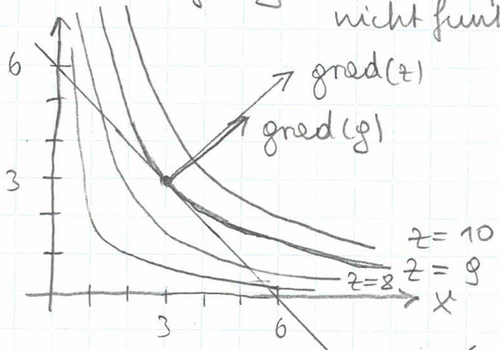
Nichtlineares Programm (NLP):
$$\boxed{\begin{array}{ll} \max & z = x y \\ \text{s.t.} & x + y = 6 \end{array}}$$

▷ 1. Lösungsweg: $z = x y = x(6-x) = 6x - x^2$

$$dz = 6 - 2x \stackrel{!}{=} 0$$

$\underline{x = 3}$ und somit $\underline{y = 3}$. $\underline{z^* = 9}$

▷ 2. Lösungsweg (der auch funktioniert, wenn der 1. Lösungsweg nicht funktioniert.)



$g := x + y - 6 = 0$ .... Nebenbed. als Nullstellengebilde von g.

$x + y = 6 \iff x + y - 6 = 0$

Die Gradienten der Zielfunktion $z = xy$ und der Nebenbedingungsfunktion $g = x + y - 6$ sind parallel am Optimum:

$$\text{grad}(z) = \lambda\, \text{grad}(g)$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Rightarrow \begin{array}{l} x = \lambda \\ y = \lambda \end{array} \Rightarrow \underline{x = y}$$

Einsetzen in $x + y - 6 = 0$

$$x + x - 6 = 0$$

$$2x = 6$$

$$\underline{x = 3,\ y = 3,\ z^* = 9}$$

Allgemein:

Falls $z(x_1,\dots,x_n)$ unter den Nebenbedingungen

$$g_1(x_1,\dots,x_n)=0$$
$$g_2(x_1,\dots,x_n)=0$$
$$\vdots$$
$$g_m(x_1,\dots,x_n)=0$$

ein lokales Optimum bei $x^*=(x_1^*,\dots,x_n^*)$ hat, dann gibt es $\lambda_1,\lambda_2,\dots,\lambda_m\in\mathbb{R}$, die sogenannten Lagrange-Multiplikatoren, so dass

$$\boxed{\begin{aligned}\operatorname{grad}(z)(x^*) &= \lambda_1\operatorname{grad}(g_1)(x^*)+\lambda_2\operatorname{grad}(g_2)(x^*)+\dots+\\ &\quad\dots+\lambda_m\operatorname{grad}(g_m)(x^*).\end{aligned}}$$

Bemerkung: Siehe Literatur für Kriterien für Maximum bzw. Minimum und die Verallgemeinerung auf Ungleichungsnebenbedingungen.