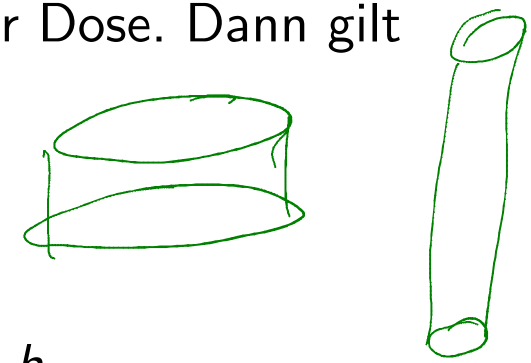
2.3 Extremalprobleme unter Nebenbedingungen

Frage: Welche Abmessungen sollte eine Metalldose haben, damit bei vorgegebenem Volumen der Materialverbrauch am geringsten ist?

Lösungsansatz: Sei $r > 0$ der Radius und $h > 0$ die Höhe der Dose. Dann gilt

$$\begin{aligned} V &= \pi r^2 h \\ O &= 2\pi r^2 + 2\pi r h \end{aligned}$$



Setze bei vorgegebenem Volumen $c \in \mathbb{R}_+$, und mit $x := r, y := h$,

$$\begin{aligned} f(x,y) &= 2\pi x^2 + 2\pi xy \\ g(x,y) &= \pi x^2 y - c = 0 \end{aligned}$$

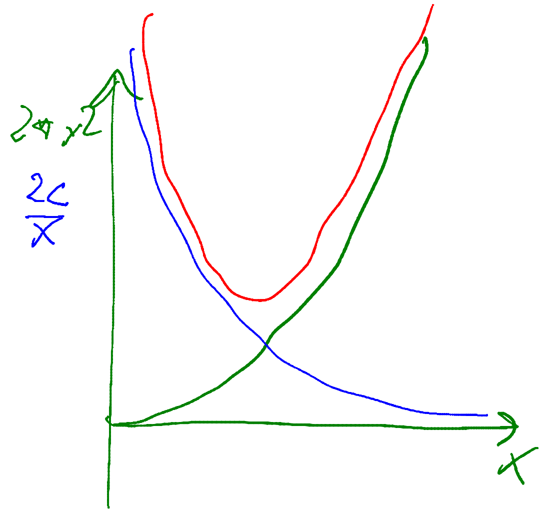
Bestimme das Minimum der Funktion $f(x,y)$ auf der Menge

$$G := \{(x,y) \in \mathbb{R}_+^2 \mid g(x,y) = 0\}$$

Lösung des restringierten Minimierungsproblems.

Aus $g(x,y) = \pi x^2 y - c = 0$ folgt

$$y = \frac{c}{\pi x^2} \quad = y(x)$$



Einsetzen in $f(x,y)$ ergibt $\quad f = f(x, y(x)) = h(x)$

$$h(x) := 2\pi x^2 + 2\pi x \frac{c}{\pi x^2} = 2\pi x^2 + \frac{2c}{x}$$

Bestimme das Minimum der Funktion $h(x)$:

$$h'(x) = 4\pi x - \frac{2c}{x^2} = 0 \quad \Rightarrow \quad 4\pi x = \frac{2c}{x^2} \quad \Rightarrow \quad \boxed{x = \left(\frac{c}{2\pi}\right)^{1/3}}$$

Hinreichende Bedingung $\quad O_{min} = 2\pi x_{min}^2 + 2\pi x_{min} y_{min} = 2\pi\left(\frac{c}{2\pi}\right)^{2/3} + 2\pi\left(\frac{c}{2\pi}\right)^{1/3} \frac{2c}{2\pi}\left(\frac{2\pi}{c}\right)^{2/3} = (2\pi)^{1/3} c^{2/3} \cdot 3$

$$h''(x) = 4\pi + \frac{4c}{x^3} \quad \Rightarrow \quad h''\left(\left(\frac{c}{2\pi}\right)^{1/3}\right) = 12\pi > 0$$

Allgemeine Formulierung des Problems.

Bestimme die Extremwerte der Funktion $\boxed{f : \mathbb{R}^n \to \mathbb{R}}$ unter den Nebenbedingungen Skalar

$$\mathbf{g}(\mathbf{x}) = 0$$

wobei $\boxed{\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m}$.

Die Nebenbedingungen lauten also

$$\begin{array}{rcl} g_1(x_1, \dots, x_n) & = & 0 \\ & \vdots & \\ g_m(x_1, \dots, x_n) & = & 0 \end{array}$$

m Gleichungen

Alternativ: Bestimme die Extremwerte der Funktion $f(\mathbf{x})$ auf der Menge

$$G := \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$$

Die Lagrange–Funktion und das Lagrange–Lemma.

Wir definieren die Lagrange–Funktion

$$F(\mathbf{x}) := f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}) = G(x, \lambda)$$

und suchen die Extremwerte von $F(\mathbf{x})$ für festes $\lambda = (\lambda_1, \ldots, \lambda_m)^T$.

Die Zahlen λ_i, $i = 1, \ldots, m$ nennt man Lagrange–Multiplikatoren.

Satz: (Lagrange–Lemma) Minimiert (bzw. maximiert) $\mathbf{x}^0$ die Lagrange–Funktion $F(\mathbf{x})$ (für ein festes λ) über D und gilt $\mathbf{g}(\mathbf{x}^0) = 0$, so liefert $\mathbf{x}^0$ das Minimum (bzw. Maximum) von $f(\mathbf{x})$ über $G := \{\mathbf{x} \in D \,|\, \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$.

Beweis: Für ein beliebiges $\mathbf{x} \in D$ gilt nach Vorausetzung Minimum $F(x^0) \le F(x)$

$$F(x^0) = f(\mathbf{x}^0) + \lambda^T \underbrace{\mathbf{g}(\mathbf{x}^0)}_{=0} \le f(\mathbf{x}) + \lambda^T \underbrace{\mathbf{g}(\mathbf{x})}_{=0} = F(x)$$

Wählt man speziell $\mathbf{x} \in G$, so ist $\mathbf{g}(\mathbf{x}) = \mathbf{g}(\mathbf{x}^0) = \mathbf{0}$, also auch $f(\mathbf{x}^0) \le f(\mathbf{x})$.

F max/min und $\varphi(x) = 0 \implies f$ max/min unter $\varphi(x) = 0$

f extr. $+ \varphi(x) = 0 + \operatorname{Rang}(\nabla \varphi) = m \implies \operatorname{grad} F = 0$

Eine notwendige Bedingung für lokale Extrema.

Sind f und g_i, $i = 1, \ldots, m$, $\mathcal{C}^1$–Funktionen, so ist eine notwendige Bedingung für eine Extremstelle $\mathbf{x}^0$ von $F(\mathbf{x})$ gegeben durch

n Gleichungen

m Gleichungen

$$\left[\operatorname{grad} F(\mathbf{x}) = \operatorname{grad} f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \operatorname{grad} g_i(\mathbf{x}) = \mathbf{0}\right]$$

$\left.\begin{matrix} x_1, \ldots, x_n & n \\ \lambda_1, \ldots, \lambda_m & m \end{matrix}\right\} m+n$

Zusammen mit den Nebenbedinungen $\boxed{\mathbf{g}(\mathbf{x}) = 0}$ ergibt sich ein (nichtlineares) Gleichungssystem mit $(n+m)$ Gleichungen und $(n+m)$ Unbekannten $\mathbf{x}$ und λ.

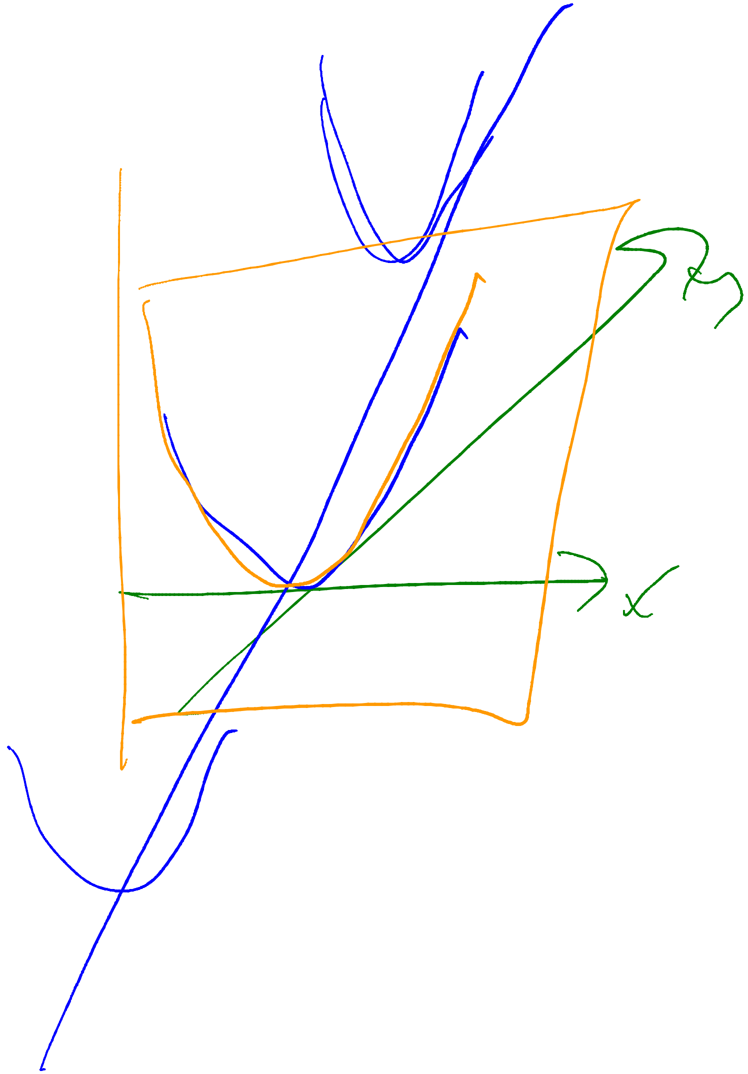
Die Lösungen $(\mathbf{x}^0, \lambda^0)$ sind die Kandidaten für die gesuchten Extremstellen, denn diese erfüllen die o.g. notwendige Bedingung.

Alternativ: Definiere eine Langrange–Funktion

$\operatorname{grad}_{(x,\lambda)} G(x,\lambda) = \begin{pmatrix} \operatorname{grad} f \\ g_1 \\ g_m \end{pmatrix} = 0$

$$G(\mathbf{x}, \lambda) := f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x})$$

und suche die Extremstellen von $G(\mathbf{x}, \lambda)$ bezüglich $\mathbf{x}$ **und** λ.



$f(x_i) - y = 0$

Einige Bemerkungen zu hinreichenden Bedingungen.

1. Man kann auch eine **hinreichende** Bedingung aufstellen:
   Sind die Funktionen f und $\mathbf{g}$ $\mathcal{C}^2$–Funktionen und ist die Hesse–Matrix $\mathbf{H}F(\mathbf{x}^0)$ der Lagrange–Funktion positiv (bzw. negativ) definit, so ist $\mathbf{x}^0$ tatsächlich ein strenges lokales Minimum (bzw. Maximum) von $f(\mathbf{x})$ auf G.

2. In den meisten Anwendungen ist die hinreichende Bedingung allerdings **nicht** erfüllt, obwohl $\mathbf{x}^0$ ein strenges lokales Extremum ist.

3. Insbesondere kann man aus der Indefinitheit der Hesse–Matrix $\mathbf{H}F(\mathbf{x}^0)$ **nicht** schließen, dass $\mathbf{x}^0$ kein Extremwert ist.

4. Ähnlich problematisch ist die hinreichende Bedingung, die man aus der Hesse–Matrix für die Lagrange–Funktion $G(\mathbf{x}, \lambda)$ bezüglich $\mathbf{x}$ **und** λ erhält.

$$f(x,y) = 2\pi x^2 + 2\pi xy$$

$$g(x,y) = \pi x^2 y - c = 0$$

$n = 2$

$m = 1$

$x = \left(\frac{c}{2\pi}\right)^{1/3}, \; y = 2x$

Lagrange Fkt. $F(x,y) = f(x,y) + \lambda g(x,y)$

$$\operatorname{grad} F(x,y) = \begin{pmatrix} 4\pi x + 2\pi y + 2\lambda\pi xy \\ 2\pi x + \lambda \pi x^2 \end{pmatrix} = 0$$

$$g(x,y) = \pi x^2 y - c = 0$$

x, y, λ

$0 = 2x + y - \frac{2}{x} xy = 2x - y = 0$

$y = 2x$

$\Rightarrow x(2 + \lambda x) = 0$

$x = 0$ ↯

$\lambda x = -2$

$\pi x^2 2x = c$

$x = \left(\frac{c}{2\pi}\right)^{1/3}$

$$HF = \begin{pmatrix} 4\pi + 2\lambda\pi y & 2\pi + 2\lambda\pi x \\ 2\pi + 2\lambda\pi x & 0 \end{pmatrix} = 2\pi \begin{pmatrix} 2 + \lambda y & 1 + \lambda x \\ 1 + \lambda x & 0 \end{pmatrix}.$$

$$\det HF = -2\pi (1 + \lambda x)^2 \leq 0$$

$$\operatorname{grad} f = 0 \quad \begin{pmatrix} 4\pi x + 2\pi y \\ 2\pi x \end{pmatrix} = 0 \Rightarrow \begin{matrix} x = 0 \\ y = 0 \end{matrix}$$

$$y \perp \operatorname{grad} g(x_0) = \begin{pmatrix} 2\pi xy \\ \pi x^2 \end{pmatrix} = \pi x \begin{pmatrix} 2y \\ x \end{pmatrix} = \pi x \begin{pmatrix} 4x \\ x \end{pmatrix} = \pi x^2 \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

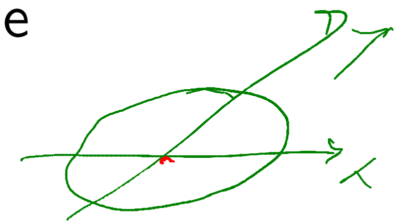
$$y = c \begin{pmatrix} -1 \\ 4 \end{pmatrix} \in T G(x_0)$$

$$\underbrace{(-1, 4)\, HF\left(\left(\frac{c}{2\pi}\right)^{1/3}, 2\left(\frac{c}{2\pi}\right)^{1/3}\right) \begin{pmatrix} -1 \\ 4 \end{pmatrix}}_{> 0}$$

Ein Beispiel zu restringierten Minimierungsproblems.

$n=2$

Gesucht seien die Extrema von $f(x, y) := xy$ auf der Kreisscheibe



$$K := \{(x, y)^T \mid x^2 + y^2 \leq 1\}$$

Da die betrachte Funktion f stetig und $K \subset \mathbb{R}^2$ kompakt ist, folgt aus der Min–Max–Eigenschaft die Existenz von globalen Maxima und Minima auf K.

Wir betrachten zunächst das Innere K^0 von K, also die offene Menge

Teil 1:

$$K^0 := \{(x, y)^T \mid x^2 + y^2 < 1\}$$

Innere

Die notwendige Bedingung für einen Extremwert lautet nun

$$\operatorname{grad} f = (y, x) = \mathbf{0}$$

Somit ist der Ursprung $\mathbf{x}^0 = \mathbf{0}$ Kandidat für ein (lokales) Extremum.

Fortsetzung des Beispiels.

Die Hesse–Matrix im Ursprung, gegeben durch

$$\mathbf{H}f(\mathbf{0}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

und ist indefinit. Daher ist $\mathbf{x}^0$ ein Sattelpunkt.

Teil 2:

Die Extrema der Funktion müssen also auf dem Rand liegen, der eine Gleichungsnebenbedingung darstellt:

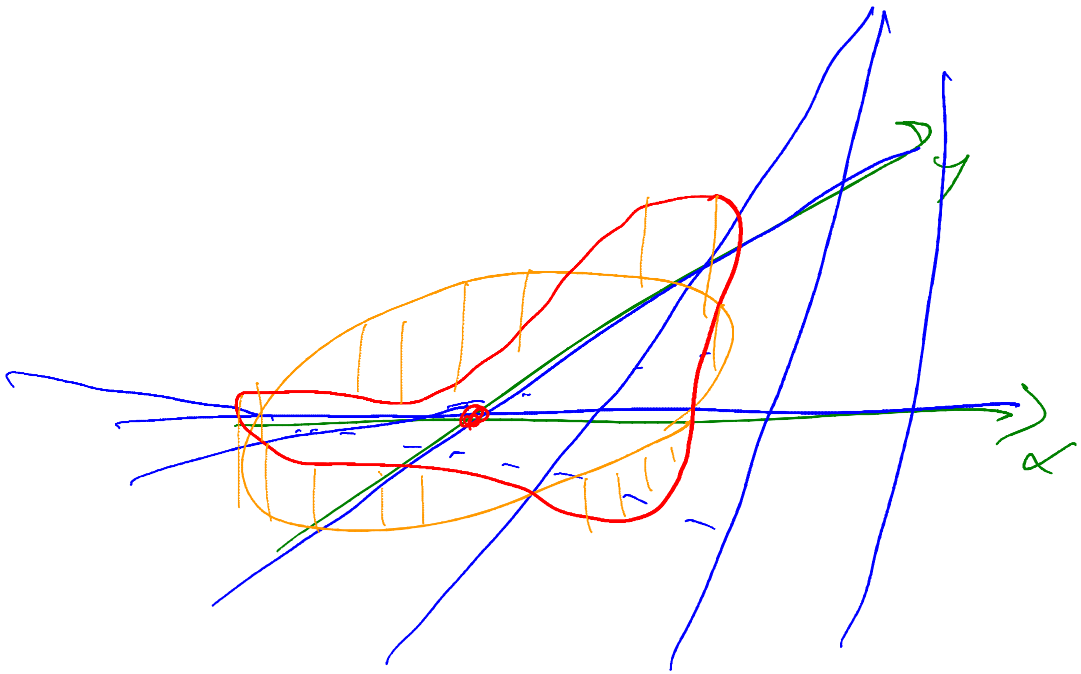
$$g(x, y) = x^2 + y^2 - 1 = 0$$

Wir suchen also die Extremwerte von $\underline{f(x, y) = xy}$ unter der Nebenbedingung $g(x, y) = 0$.

Die Lagrange–Funktion lautet

$$F(x, y) = xy + \lambda(x^2 + y^2 - 1)$$

$f(x,y) = x \cdot y$



Komplettierung des Beispiels.

$n=2 \quad m=1$

Damit ergibt sich das (nichtlineare) Gleichungssystem x, y, λ

$$\text{grad } F = 0 \quad \left\{ \begin{array}{rcl} y + 2\lambda x &=& 0 \\ x + 2\lambda y &=& 0 \end{array} \right. \qquad 2\lambda = -\frac{y}{x}$$

$$x - \frac{y}{x} y = \frac{1}{x}(x^2 - y^2) = 0 \qquad y = \pm x$$

$$g(x) = 0 \Rightarrow \quad x^2 + y^2 = 1 \qquad 2x^2 = 1 \qquad x = \pm\frac{1}{\sqrt{2}} \qquad y = \pm\frac{1}{\sqrt{2}}$$

mit den vier Lösungen

$$\lambda = \frac{1}{2} \quad : \quad \mathbf{x}^{(1)} = (\sqrt{1/2}, -\sqrt{1/2})^T \quad \mathbf{x}^{(2)} = (-\sqrt{1/2}, \sqrt{1/2})^T$$

$$\lambda = -\frac{1}{2} \quad : \quad \mathbf{x}^{(3)} = (\sqrt{1/2}, \sqrt{1/2})^T \quad \mathbf{x}^{(4)} = (-\sqrt{1/2}, -\sqrt{1/2})^T$$

Minima und Maxima lassen sich nun einfach aus den Funktionswerten ablesen

$$f(\mathbf{x}^{(1)}) = f(\mathbf{x}^{(2)}) = -1/2 \qquad f(\mathbf{x}^{(3)}) = f(\mathbf{x}^{(4)}) = 1/2$$

d.h. Minima sind $\mathbf{x}^{(1)}$ und $\mathbf{x}^{(2)}$, Maxima sind $\mathbf{x}^{(3)}$ und $\mathbf{x}^{(4)}$.

Lagrange–Multiplikatoren–Regel.

Satz: Seien $f, g_1, \ldots, g_m : D \to \mathbb{R}$ jeweils $\mathcal{C}^1$–Funktionen, und sei $\mathbf{x}^0 \in D$ ein lokales Extremum von $f(\mathbf{x})$ unter der Nebenbedingung $\mathbf{g}(\mathbf{x}) = \mathbf{0}$. m Gleichungen
$m < n$
Weiterhin gelte die Regularitätsbedingung

$$\text{rang}\left(\underbrace{\mathbf{J}\,\mathbf{g}(\mathbf{x}^0)}_{m \times n \text{ Matrix}}\right) = m$$

Dann existieren Lagrange–Multiplikatoren $\lambda_1, \ldots, \lambda_m$, so dass für die Lagrange Funktion

$$F(\mathbf{x}) := f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x})$$

die folgende **notwendige Bedingung erster Ordnung** gilt:

$$\text{grad}\, F(\mathbf{x}^0) = \mathbf{0}$$

Notwendige Bedingung zweiter Ordnung und hinreichende Bedingung.

Satz: 1) Ist $\mathbf{x}^0 \in D$ ein lokales Minimum von $f(\mathbf{x})$ unter der Nebenbedingung $\mathbf{g}(\mathbf{x}) = 0$, ist die Regularitätsbedingung erfüllt und sind $\lambda_1, \ldots, \lambda_m$ zugehörige Lagrange–Multiplikatoren, so ist die Hesse–Matrix $\mathbf{H}F(\mathbf{x}^0)$ der Lagrange–Funktion positiv semidefinit auf dem Tangentialraum

$$TG(\mathbf{x}^0) := \{\mathbf{y} \in \mathbb{R}^n \,|\, \operatorname{grad} g_i(\mathbf{x}^0) \cdot \mathbf{y} = 0 \text{ für } i = 1, \ldots, m\}$$

d.h., es gilt $\mathbf{y}^T\, \mathbf{H}F(\mathbf{x}^0)\, \mathbf{y} \geq 0$ für alle $\mathbf{y} \in TG(\mathbf{x}^0)$.

2) Ist für einen Punkt $\mathbf{x}^0 \in G$ die Regularitätsbedingung erfüllt, existieren Lagrange–Multiplikatoren $\lambda_1, \ldots, \lambda_m$, so dass $\mathbf{x}^0$ ein stationärer Punkt der zugehörigen Lagrange–Funktion ist, und ist die Hesse–Matrix $\mathbf{H}F(\mathbf{x}^0)$ positiv definit auf dem Tangentialraum $TG(\mathbf{x}^0)$, d.h., gilt

$$\mathbf{y}^T\, \mathbf{H}F(\mathbf{x}^0)\, \mathbf{y} > 0 \quad \forall\, \mathbf{y} \in TG(\mathbf{x}^0) \setminus \{0\},$$

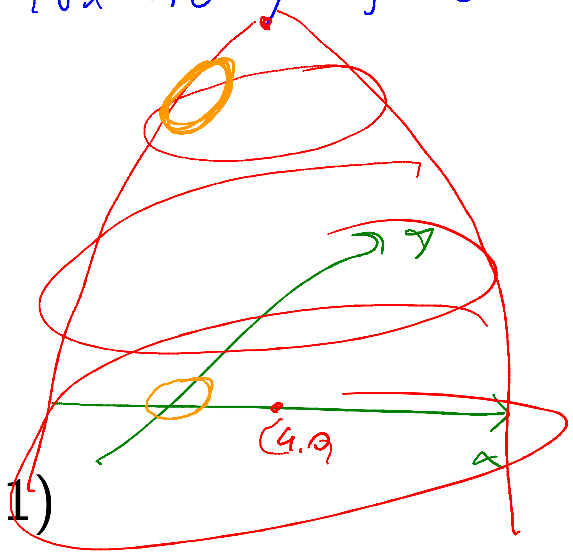
so ist $\mathbf{x}^0$ ein strenges lokales Minimum von $f(\mathbf{x})$ unter der Nebenbedingung $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.

Beispiel.

Man bestimme das globale Maximum der Funktion $n=2 \quad m=1$

$$f(x,y) = -x^2 + 8x - y^2 + 9 = -x^2+8x-16-y^2+9+16$$

$$= -(x-4)^2 - y^2 + 25$$



unter der Nebenbedingung

$$g(x,y) = x^2 + y^2 - 1 = 0$$

Die Lagrange–Funktion ist

$$F(x) = -x^2 + 8x - y^2 + 9 + \lambda(x^2 + y^2 - 1)$$

Aus der notwendigen Bedingung ergibt sich das nichtlineare System

$\operatorname{grad} F = 0$

$$\begin{cases} -2x + 8 = -2\lambda x \\ -2y = -2\lambda y \\ x^2 + y^2 = 1 \end{cases}$$

x, y, λ

$g(x) = 0$

Fortsetzung des Beispiels.

Aus der notwendigen Bedingung ergibt sich das nichtlineare System

$$\begin{aligned} -2x + 8 &= -2\lambda x \\ -2y &= -2\lambda y \\ x^2 + y^2 &= 1 \end{aligned}$$

$\lambda \neq 1$

$y = 0$

$x = \pm 1$

Aus der ersten Gleichung folgt $\lambda \neq 1$. Verwendet man dies in der zweiten Gleichung, so gilt $y = 0$. Aus der dritten Gleichung erkennt man sofort $x = \pm 1$. Demnach sind die beiden Punkte $\boxed{(x, y) = (1, 0)}$ und $\boxed{(x, y) = (-1, 0)}$ Kandidaten für das globale Maximum. Wegen

$$f(1, 0) = 16 \qquad f(-1, 0) = 0$$

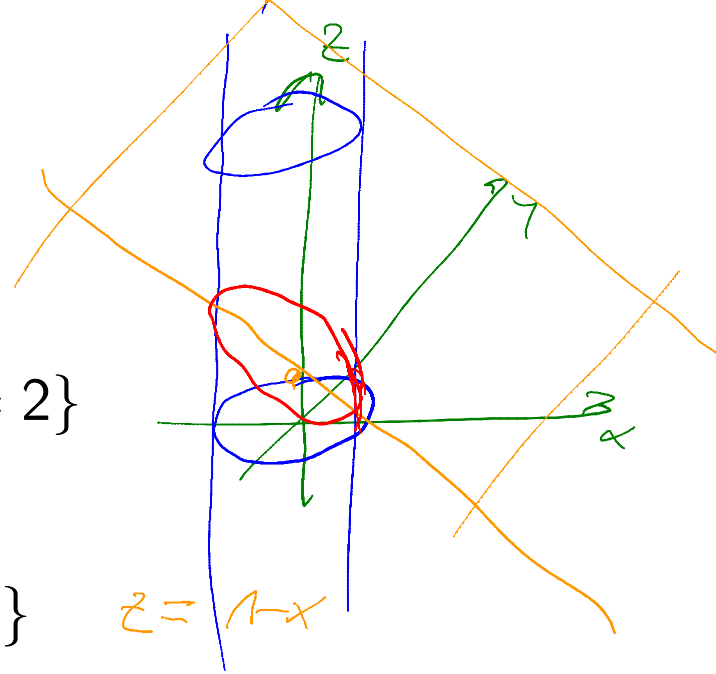
wird das globale Maximum von $f(x, y)$ unter der Nebenbedingung $g(x, y) = 0$ im Punkt $(x, y) = (1, 0)$ angenommen.

Ein weiteres Beispiel.

Man bestimme die lokalen Extremwerte der Funktion

$n = 3$, $m = 2$

$$f(x, y, z) = 2x + 3y + 2z$$



auf dem Durchschnitt des Zylindersmantels

$$M_Z := \{(x, y, z)^T \in \mathbb{R}^3 \mid x^2 + y^2 = 2\}$$

mit der Ebene

$$E := \{(x, y, z)^T \in \mathbb{R}^3 \mid x + z = 1\}$$

Umformulierung: Bestimme die Extremwerte der Funktion $f(x, y, z)$ unter den Nebenbedingungen

$$\begin{aligned} g_1(x, y, z) &:= x^2 + y^2 - 2 = 0 \\ g_2(x, y, z) &:= x + z - 1 = 0 \end{aligned}$$

Fortsetzung des Beispiel.

Die Jacobi–Matrix

$$\mathbf{Jg(x)} = \begin{pmatrix} 2x & 2y & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

2×3

(0,0) erfüllt $g_1(x,y,z)=0$ nicht

hat den Rang 2, d.h. wir können über die Lagrange–Funktion Extremwerte bestimmen:

$$F(x,y,z) = 2x + 3y + 2z + \lambda_1(x^2 + y^2 - 2) + \lambda_2(x + z - 1)$$

Die notwendige Bedingung ergibt das nichtlineare Gleichungssystem

$x, y, z, \lambda_1, \lambda_2$

grad $F = 0$

$$\left\{\begin{aligned} 2 + 2\lambda_1 x + \lambda_2 &= 0 \\ 3 + 2\lambda_1 y &= 0 \\ 2 + \lambda_2 &= 0 \end{aligned}\right.$$

$g_1 = 0$

$$x^2 + y^2 = 2$$

$g_2 = 0$

$$x + z = 1$$

Fortsetzung des Beispiel.

Die notwendige Bedingung ergibt das nichtlineare Gleichungssystem

$$\begin{aligned} 2 + 2\lambda_1 x + \lambda_2 &= 0 \\ 3 + 2\lambda_1 y &= 0 \\ 2 + \lambda_2 &= 0 \\ x^2 + y^2 &= 2 \\ x + z &= 1 \end{aligned}$$

$2\lambda_1 x = 0$

$\lambda_1 \neq 0 \Rightarrow x = 0$

$y = \pm\sqrt{2}$

$z = 1$

Aus der ersten und dritten Gleichung folgt

$$2\lambda_1 x = 0$$

Aus der zweiten Gleichung folgt $\lambda_1 \neq 0$, also $x = 0$.
Damit ergeben sich die möglichen Extremwerte als

$$\boxed{(x, y, z) = (0, \sqrt{2}, 1)} \qquad \boxed{(x, y, z) = (0, -\sqrt{2}, 1)}$$

Komplettierung des Beispiel.

Die möglichen Extremwerte sind also

$$(x, y, z) = (0, \sqrt{2}, 1) \quad (x, y, z) = (0, -\sqrt{2}, 1)$$

und liegen offensichtlich auf der Mantelfläche M_Z des Zylinders Z mit

$$\begin{aligned} Z &= \{(x, y, z)^T \in \mathbb{R}^3 \,|\, x^2 + y^2 \leq 2\} \\ M_Z &= \{(x, y, z)^T \in \mathbb{R}^3 \,|\, x^2 + y^2 = 2\} \end{aligned}$$

Man berechnet nun die zugehörigen Funktionswerte

$$\begin{aligned} f(0, \sqrt{2}, 1) &= 3\sqrt{2} + 2 \\ f(0, -\sqrt{2}, 1) &= -3\sqrt{2} + 2 \end{aligned}$$

Daher liegt im Punkt $(x, y, z) = (0, \sqrt{2}, 1)$ ein Maximum und im Punkt $(x, y, z) = (0, -\sqrt{2}, 1)$ ein Minimum vor.