

Problems marked with a (*) are a bit more complex and can be skipped at a first read.
If you don't have a lot of time focus on the Problems/subquestions marked with (♡).

6.1. BONUS PROBLEM. Sei $f \in C^2(\mathbb{R}^n)$ eine konvexe Funktion.

- Zeigen Sie, dass $z \in \mathbb{R}^n$ ein kritischer Punkt von f ist, genau dann wenn z ein globaler Minimierer ist.
- Geben Sie ein Beispiel für solch eine Funktion f in einigen \mathbb{R}^n mit $n > 1$ an, die immer nichtnegativ ist, aber keinen Minimalpunkt hat. Das heißt

$$f(x) > \inf_{\mathbb{R}^n} f \geq 0 \text{ für alle } x \in \mathbb{R}^n.$$

Sie können alle in der Klasse behandelten Sätze verwenden.

6.2. Die Signatur einer 2×2 Matrix. Trotz der Definition ist es nicht notwendig, die Eigenwerte einer Matrix zu berechnen, um ihre Signatur zu finden¹. Beweisen Sie, dass wir für eine 2×2 Matrix M die folgende einfache Regel zur Bestimmung der Signatur in Bezug auf $\det M$ und $\text{Tr}M$ haben:

- Wenn $\det M > 0$, $\text{Tr}M > 0$ ist, dann ist M positiv definit,
- Wenn $\det M > 0$, $\text{Tr}M < 0$ ist, dann ist M negativ definit,
- Wenn $\det M < 0$ ist, dann ist M indefinit,
- Wenn $\det M = 0$ ist, dann ist M degeneriert.

6.3. Isoperimetrische Dreiecke. Unter allen Dreiecken mit einem Umfang von 2 sollen diejenigen mit der größten Fläche gefunden werden. Sie können die Heronsche Formel voraussetzen, die die Fläche eines Dreiecks in Bezug auf die Längen seiner Seiten x, y, z angibt:

$$A = \sqrt{p(p-x)(p-y)(p-z)}, \quad \text{mit } 2p := x + y + z,$$

so dass in unserem Fall $p = 1$.

6.4. Baryzentrum (♡). Seien $y_1, \dots, y_k \in \mathbb{R}^n$ gegeben. Zeigen Sie, dass es genau einen Punkt gibt, für den

$$f(x) = \|x - y_1\|^2 + \dots + \|x - y_k\|^2, \quad x \in \mathbb{R}^n$$

minimal ist, und bestimmen Sie diesen Punkt.

6.5. Lineare Regression I (♡). Sie studieren den Immobilienmarkt in Zürich über ein Jahr, in dem N Häuser verkauft werden. Sie verfolgen die Größe der Häuser x_1, \dots, x_N

¹Fragen Sie ChatGPT nach dem Satz über die Hauptminoren

und die entsprechenden Verkaufspreise y_1, \dots, y_N . Jetzt möchten Sie "die" Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ finden, die

$$\text{Verkaufspreis} = f(\text{Größe des Hauses}),$$

gibt und Sie machen die (nicht unvernünftige) Annahme, dass f affin ist, d.h., $f_{a,b}(x) = ax + b$ für einige Koeffizienten $a, b \in \mathbb{R}$. Unter all diesen Funktionen finden Sie (in Bezug auf die von Ihnen gesammelten Daten) den Wert der Parameter a, b , der den durchschnittlichen quadratischen Fehler minimiert

$$E(a, b) := \sum_{i=1}^N (y_i - f_{a,b}(x_i))^2, \quad a, b \in \mathbb{R}.$$

6.6. Konvexe Funktionen (♥). Entscheide, ob die folgenden Funktionen f_i im konvexen Bereich $U_i \subset \mathbb{R}^n$ konvex sind. Versuche in jedem Fall, das einfachste Argument zu finden, du kannst fast immer lange Berechnungen vermeiden.

1. $f_1(x, y) = x^2 + y^2 - 4y$ definiert in $U_1 = \mathbb{R}^2$
2. $f_2(x, y) = x^2 + y^2 - y^4$ definiert in $U_2 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < \frac{1}{10000}\}$
3. $f_3(x, y) = x^2 + y^2 - 4xy$ definiert in $U_3 = \mathbb{R}^2$
4. $f_4(x, y) = x^2 + y^2 - 4xy$ definiert in $U_4 = \{(x, y) \in \mathbb{R}^2 : 0 < 10x < |y|\}$
5. $f_5(x) = \phi(g(x))$, $x \in U_5$ wobei $g \in C^2(U_5)$ irgendeine konvexe Funktion in $U_5 \subset \mathbb{R}^n$ ist und $\phi \in C^2(\mathbb{R})$ irgendeine konvexe und steigende Funktion ist.
6. $f_6(x, y) = (1 + x^2 + y^2)^{1/2}$ definiert in $U_6 = \mathbb{R}^2$
7. $f_7(x, y) = -(1 + x^2 + y^2)^{-1/2}$ definiert in $U_7 = \mathbb{R}^2$
8. $f_8(x) = \sum_{i=1}^n |x_i|^p$ in $U_8 = \mathbb{R}^n$, wobei $p \geq 1$ ein fester Exponent ist.
9. $f_9(x) = \max\{\phi(x), \psi(x)\}$ wobei $\phi, \psi \in C(U_9)$ ein beliebiges Paar von konvexen Funktionen definiert in einem offenen Bereich $U_9 \subset \mathbb{R}^n$ sind.
10. $f_{10}(x) = |x|$ definiert in $U_{10} = \mathbb{R}^n$.
11. $f_{11}(x) = \phi(|x|)$ in $U_{11} = B_1 \subset \mathbb{R}^n$, wobei $\phi \in C(\mathbb{R})$ eine beliebige konvexe Funktion ist.

6.7. Multiple Choice. Unter den folgenden Aussagen über konvexe Funktionen markieren Sie diejenigen (und nur diejenigen), die immer wahr sind.

- (a) If $f \in C^1(U)$ is convex in some open convex set $U \subset \mathbb{R}^n$ and f has a local maximum at $z \in U$, then $\nabla f \equiv 0$ in U .
- (b) If $f \in C^1(U)$ is convex in some open convex set $U \subset \mathbb{R}^n$ and f has a global maximum at $z \in U$, then $\nabla f \equiv 0$ in U .

(c) Assume $f_n \in C^2(\mathbb{R})$ is a sequence of convex functions that converge pointwise to some $f: \mathbb{R} \rightarrow \mathbb{R}$. Is f necessarily convex?

(d) There exists a convex function $f \in C^\infty(\mathbb{R}^2)$ such that

$$f(x) = 1 - 2x_1 + x_2^3 + O(|x|^4) \text{ as } |x| \rightarrow 0.$$

(e) There exists a convex function $f \in C^\infty(\mathbb{R}^2)$ such that

$$f(x) = 1 - 2x_1 + x_2^4 + O(|x|^4) \text{ as } |x| \rightarrow 0.$$

(f) A convex set is not necessarily connected.

6.8. Multiple choice. Die Hesse-Matrix von $f \in C^2(\mathbb{R}^n)$ ist an einem kritischen Punkt x_0 von f positiv semidefinit, d.h.,

$$\langle v, Hf(x_0)v \rangle \geq 0 \text{ für alle } v \in \mathbb{R}^n.$$

Welche der folgenden Aussagen treffen notwendigerweise zu? (Es können mehrere zutreffen).

(a) x_0 ist ein strenges lokales Minimum von f .

(b) x_0 ist ein lokales Minimum von f .

(c) x_0 ist kein lokales Maximum von f .

(d) Keine der obigen Aussagen.

6.9. Minimierung. Die Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ist gegeben durch $f(x, y) = 2x^2 + y^2 - x$. Bestimmen Sie die Extrema von f auf...

(a) ... dem Einheitskreis $\mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$;

(b) ... der abgeschlossenen Einheitskreisscheibe $\mathbb{D} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$.

6.10. Lagrange-Multiplikatoren (♥). Betrachten Sie die Funktion $f(x, y, z) = 3x - y + 2z$ und die Menge

$$M = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1, x + y = 0\}.$$

Bestimmen Sie die Extrema von f auf M und deren Art.

6.11. Kritische Punkte. Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ die Funktion $f(x, y) = (ax^2 + by^2)e^{-x^2 - y^2}$ mit reellen Parametern $a, b \in \mathbb{R}$. Finde alle kritischen Punkte und bestimme ihre Art mit dem Hessischen Test, abhängig von a, b .

Hinweise:

6.2 Verwenden Sie den spektralen Satz und die Eigenschaften:

$$\det(AB) = \det(A) \det(B), \quad \text{Tr}(AB) = \text{Tr}(BA).$$

6.3 Minimieren Sie A^2 anstelle von A . Sie können die Methode der Lagrange-Multiplikatoren verwenden.

6.5 Lassen Sie sich nicht von der Einstellung ablenken, schließlich müssen Sie ein quadratisches Polynom von a, b minimieren....

6. Solutions

Solution of 6.1:

- (a) Wenn $z \in \mathbb{R}^n$ ein globaler Minimalpunkt ist, dann ist $\nabla f(z) = 0$ nach Satz 11.4. Wenn $\nabla f(z)$, dann gilt nach Satz 11.24

$$f(y) \geq f(z) + \nabla f(z) \cdot (y - z) = f(z) \text{ für alle } y \in \mathbb{R}^n,$$

was bedeutet — nach Definition — dass z ein globaler Minimierer ist.

- (b) $f(x) := e^{x_1+x_2}$ funktioniert. Offensichtlich ist $f(x) > 0$ für alle $x \in \mathbb{R}^2$, und $\lim_{t \rightarrow \infty} f(-t, -t) = \lim_{t \rightarrow \infty} e^{-2t} = 0$, also $\inf_{\mathbb{R}^2} f = 0$. Darüber hinaus ist f konvex, da für alle Einheitsvektoren v gilt

$$\partial_v f(x) = e^{x_1+x_2}(v_1 + v_2),$$

und

$$\partial_{vv} f(x) = e^{x_1+x_2}(v_1 + v_2)^2 \geq 0.$$

Solution of 6.2: Nach dem Spektralsatz gibt es $O \in \mathbb{R}^{2 \times 2}$, so dass

$$OMO^T = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad OO^T = \mathbf{1} \text{ (insbesondere } \det O = \pm 1),$$

mit $\lambda_1, \lambda_2 \in \mathbb{R}$. Somit können wir die (vielleicht bekannte) Tatsache sehen, dass

$$\det M = \det(OMO^T) = \lambda_1 \lambda_2, \quad \text{Tr} M = \text{Tr}(MO^T O) = \text{Tr}(OMO^T) = \lambda_1 + \lambda_2.$$

Es ist jetzt sofort zu überprüfen, dass

- $\lambda_1 \lambda_2 > 0, \lambda_1 + \lambda_2 > 0$ impliziert, dass $\lambda_1 > 0$ und $\lambda_2 > 0$ sind.
- $\lambda_1 \lambda_2 > 0, \lambda_1 + \lambda_2 < 0$ impliziert, dass $\lambda_1 < 0$ und $\lambda_2 < 0$ sind.
- $\lambda_1 \lambda_2 < 0$ impliziert, dass einer der λ_i positiv und der andere negativ ist.
- $\lambda_1 \lambda_2 = 0$ impliziert, dass einer von ihnen null ist.

Solution of 6.3: Die Menge der möglichen (x, y, z) , die Seiten eines Dreiecks darstellen, wird gegeben durch

$$U := \{(x, y, z) \in \mathbb{R}^3 : x > 0, y > 0, z > 0, x + y > z, y + z > x, z + x > y\}.$$

Wir versuchen, die Fläche zu minimieren, was dasselbe ist wie das Minimieren ihres Quadrats, das durch die Heronsche Formel gegeben ist als

$$f(x, y, z) = A^2 = (1-x)(1-y)(1-z), \quad (x, y, z) \in U$$

unter der Bedingung $x + y + z = 2$. Nach dem Weierstraß-Theorem besitzt f einen Maximalpunkt $(x_0, y_0, z_0) \in \bar{U} \cap \{x + y + z = 2\}$.

Fall 1: (x_0, y_0, z_0) liegt auf ∂U . Beachten Sie, dass jeder Punkt in der Abschlussmenge \bar{U} die Bedingung erfüllen muss:

$$U := \{(x, y, z) \in \mathbb{R}^3 : x \geq 0, y \geq 0, z \geq 0, x + y \geq z, y + z \geq x, z + x \geq y\},$$

wie sich aus einer konvergenten Folge von Elementen von U ergibt. Daher muss ein Element auf dem Rand $\partial U = \bar{U} \setminus U$ mindestens eine der folgenden Bedingungen erfüllen (sonst würde es im Inneren liegen)

$$x = 0, \quad y = 0, \quad z = 0, \quad x + y = z, \quad y + z = x, \quad z + x = y.$$

In all diesen Fällen degeneriert unser Dreieck zu einem Segment (eine der Seiten ist null!) und daher muss f an diesen Punkten verschwinden, die Minimalpunkte sein müssen. Überprüfen wir das:

- Wenn (angenommen) $x = 0$, dann $y \geq z$ und $z \geq y$, also $z = y$. Da der Umfang 2 beträgt, finden wir $y = z = 1$, was zu $f(0, 1, 1) = 0$ führt. Die anderen Fälle sind symmetrisch.
- Wenn (angenommen) $x + y = z$, dann $2 - z = z$, also $z = 1$ und somit $f(x, y, 1) = 0$. Die anderen Fälle sind symmetrisch.

Wir kommen zu dem Schluss, dass der "Kampf um das Maximum" im Inneren von U ausgetragen wird.

Fall 2: (x_0, y_0, z_0) liegt im Inneren von U . Das Verfahren der Lagrange-Multiplikatoren gilt mit $g(x, y, z) := x + y + z - 1$ und, da ∇g niemals null ist, ergibt sich das System

$$\begin{cases} (1 - y)(1 - z) = \lambda, \\ (1 - z)(1 - x) = \lambda, \\ (1 - x)(1 - y) = \lambda \\ x + y + z = 2, \\ (x, y, z) \in U, \lambda \in \mathbb{R}. \end{cases}$$

Wir lösen dieses System. Zunächst behaupten wir, dass keiner der Werte von x, y, z gleich eins sein kann. Angenommen, $z = 1$, dann ist zwangsläufig $\lambda = 0$ und dann auch $(1 - x)(1 - y) = 0$, also $x = 1$ oder $y = 1$. Da $x + y + z = 2$ ist, finden wir, dass einer von x, y null ist, was im Widerspruch zu $(x, y, z) \in U$ steht.

Nun setzen wir die Differenz der ersten beiden Gleichungen gleich null und erhalten

$$0 = (1 - z)(x - y),$$

was, da $z \neq 1$, zu $x = y$ führt. Wenn wir nun die Differenz zwischen den zweiten beiden Gleichungen bilden, erhalten wir

$$(1 - x)(y - z) = 0,$$

und da $x \neq 1$ ist, erhalten wir $y = z$. Somit muss die einzige Lösung dieses Systems erfüllen

$$x = y = z, x + y + z = 2 \Rightarrow x = y = z = \frac{2}{3},$$

da wir angenommen haben, dass $(x_0, y_0, z_0) \in U$ das System löst, wissen wir, dass es

$$(x_0, y_0, z_0) = \left(\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right), f\left(\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right) = 1/27$$

entspricht, was dem gleichseitigen Dreieck entspricht.

Solution of 6.4: Wir beginnen damit zu beobachten, dass f eine Summe konvexer Funktionen ist, daher ist sie konvex. Außerdem ist f glatt (es ist ein Polynom). Da f konvex ist, ist ein Punkt genau dann ein globaler Minimierer, wenn er ein kritischer Punkt ist. Wir beweisen, dass es einen eindeutigen kritischen Punkt gibt, indem wir gleichzeitig beweisen, dass es einen Minimalpunkt gibt und dass er eindeutig ist.

Wir erhalten für die partiellen Ableitungen von f

$$\partial_{x_j} f(x) = 2 \sum_{i=1}^k (x_j - (y_i)_j).$$

Damit x ein kritischer Punkt ist, muss $\partial_{x_j} f(x) = 0$ für alle j sein, daher $x_j = \frac{1}{k} \sum_{i=1}^k (y_i)_j$. Direkte Berechnung ergibt

$$\partial_{x_i} \partial_{x_j} f(x) = 2k \delta_{ij},$$

somit ist H diagonal mit Einträgen $2k$ für alle $x \in \mathbb{R}^n$. Wir erhalten, dass H positiv definit ist, daher gibt es an dem kritischen Punkt ein Minimum, und der Punkt, der das Minimum definiert, wird durch die Gleichung $x_j = \frac{1}{k} \sum_{i=1}^k (y_i)_j$ eindeutig bestimmt.

Solution of 6.5: Wir sollen die Fehlerfunktion minimieren

$$E : (a, b) \mapsto \sum_{i=1}^N (ax_i + b - y_i)^2,$$

gegeben die festen Parameter $x_1, \dots, x_N, y_1, \dots, y_N$. E ist die Summe der Quadrate von affinen Funktionen, daher ist sie konvex und $C^\infty(\mathbb{R}^2)$. Daher wird jeder kritische Punkt ein globaler Minimierer sein, ein solcher Punkt wird gefunden, indem man löst

$$\begin{cases} 0 = \partial_1 E(a, b) = 2 \sum_{i=1}^n (ax_i + b - y_i)x_i \\ 0 = \partial_2 E(a, b) = \sum_{i=1}^n (ax_i + b - y_i) \end{cases}$$

was äquivalent ist zum linearen System

$$\begin{bmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{bmatrix}.$$

Dessen Lösung ist

$$a = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \quad b = \frac{-\sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i + \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2},$$

wo wir verwendet haben

$$\begin{bmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{bmatrix}^{-1} = \frac{1}{(\sum_{i=1}^n x_i^2) \cdot n - (\sum_{i=1}^n x_i)^2} \begin{bmatrix} n & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix}$$

Dieses System kann eindeutig gelöst werden, wenn die Determinante ungleich Null ist, also wenn

$$n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \neq 0,$$

aber dies ist immer eine positive Zahl nach der QM-AM-Ungleichung, es sei denn, die $x_1 = \dots = x_N$ sind alle gleich. In diesem Fall wird das System zu

$$\begin{bmatrix} x_1^2 & x_1 \\ x_1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \frac{\sum_{i=1}^n y_i}{n} \begin{bmatrix} x_1 \\ 1 \end{bmatrix},$$

was durch $a = 0$ und $b = \frac{\sum_{i=1}^n y_i}{n}$ gelöst wird.

Solution of 6.6:

- f_1 ist konvex, da seine Hesse-Matrix $2\mathbf{1}_{2 \times 2}$ ist, was positiv definit ist.

Ähnlich,

$$Hf_2(x, y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 - 12y^2 \end{bmatrix},$$

was genau dann positiv definit ist, wenn $2 - 12y^2 > 0$ ist, und dies ist im gegebenen Bereich U der Fall (aber nicht im gesamten $\mathbb{R}^2 \dots$).

- f_3 und f_4 sind nicht konvex, da ihre Einschränkung auf die Linie $t \mapsto (t, t)$ **nicht** konvex ist

$$f_3(t, t) = f_4(t, t) = -2t^2.$$

- Wir beobachten, dass $|\cdot|$ konvex ist (wie jede andere Norm auch), dies kann mit der Definition überprüft werden:

$$|tx + (1-t)y| \leq t|x| + (1-t)|y|$$

wobei wir die positive Homogenität und die Dreiecksungleichung verwendet haben. Dies beweist, dass f_{10} konvex ist.

- Wir zeigen, dass auch f_5 konvex ist, mit einer Berechnung unter Verwendung der Kettenregel, zunächst

$$\partial_v(\phi(g(x))) = \phi'(g(x))\partial_v g(x),$$

und dann

$$\partial_{vv}(\phi(g(x))) = \underbrace{\phi''(g(x))}_{\geq 0} (\partial_v g(x))^2 + \underbrace{\phi'(g(x))\partial_{vv} g(x)}_{\geq 0} \geq 0 \geq 0.$$

Es folgt sofort, dass f_6 konvex ist, da $t \mapsto (1+t^2)1/2$ konvex ist (Analysis I) und ebenso $x \mapsto |x|$.

Ebenso ist f_8 konvex, da es eine Summe konvexer Funktionen ist. Da $x \mapsto x_i$ konvex ist und ebenso $t \mapsto |t|^p$, ist auch jede $|x_i|^p$ konvex.

- f_7 ist nicht konvex, da ihre Einschränkung auf die Linie $t \mapsto (t, t)$ die Funktion $t \mapsto -(1 + 2t^2) - 1/2$ ist, die nicht konvex ist (Analysis I).
- f_9 ist immer konvex, schreiben Sie einfach beide Konvexitätsungleichungen auf

$$\begin{aligned}\phi(tx + (1-t)y) &\leq t\phi(x) + (1-t)\phi(y) \leq tf_6(x) + (1-t)f_6(x), \\ \psi(tx + (1-t)y) &\leq t\psi(x) + (1-t)\psi(y) \leq tf_6(x) + (1-t)f_6(x),\end{aligned}$$

daher gilt dieselbe Schranke für $\max\{\phi(tx + (1-t)y), \psi(tx + (1-t)y)\}$.

- Es ist nicht notwendigerweise konvex, nehmen Sie $f_{11}(x) := e^{-|x|}$, was $\phi(t) = e^{-t}$ entspricht. f_{11} kann nicht konvex sein, da es seinen absoluten Höchstwert (d. h., 1) im Inneren des Punktes $x = 0$ erreicht und nicht konstant ist.

Solution of 6.7:

- (a) Falsch. Nehmen Sie $f(x) + \max\{x_1, \frac{1}{2}\}^4$, $U = B_1 \subset \mathbb{R}^n$ welches konvex ist, hat ein lokales Maximum bei $x = 0$ (es ist dort lokal konstant), aber sein Gradient verschwindet nicht überall.
- (b) Wahr. Da z ein lokales Maximum ist, gilt $\nabla f(z) = 0$, aber f ist konvex, daher gilt

$$f(x) \geq f(z) + \nabla f(z) \cdot (y - z) = f(z) \text{ für alle } x \in U,$$

aber andererseits gilt $f(x) \leq f(z)$ für alle $x \in U$, weil z ein lokales Maximum ist. Wir folgern, dass f notwendigerweise konstant ist, daher verschwindet ihr Gradient identisch.

- (c) Wahr, wählen Sie beliebige $x, y \in U$ und $t \in [0, 1]$, nach Definition wissen wir

$$f_n(tx + (1-t)y) \leq tf_n(x) + (1-t)f_n(y) \text{ für alle } n \in \mathbb{N},$$

wenn wir x, y, t festhalten und $n \rightarrow \infty$ gehen lassen, finden wir die Ungleichung für die Konvexität von f an diesen Punkten. Da x, y, t beliebig waren, folgern wir, dass f konvex sein muss.

- (d) Falsch. Eine konvexe Funktion liegt über ihrer Tangentenebene an einem Punkt, daher gilt in diesem Fall

$$f(x) \geq 1 - 2x_1 \text{ für alle } x \in \mathbb{R}^n,$$

(wir finden es aus der Taylorentwicklung bei $x = 0$). Aber dann folgern wir, dass für ein kleines $\rho > 0$ und ein großes $M > 0$ gelten muss

$$1 - 2x_1 + x_2^3 + M|x|^4 \geq f(x) \geq 1 - 2x_1 \text{ für alle } |x| < \rho,$$

was umgeformt wird zu

$$x_2^3 \geq -M|x|^4 \text{ für alle } |x| < \rho,$$

was unmöglich ist, wenn wir $x = (0, -r, 0, \dots, 0)$ nehmen und $r \downarrow 0$ lassen.

- (e) Wahr, $f = 1 - 2x_1 + x_2^4$ ist selbst konvex.
- (f) Falsch, nach Definition ist es wegstreckenweise verbunden durch gerade Segmente, also ist es verbunden.

Solution of 6.8:

- (a) Falsch, $f \equiv 0$ ist ein Gegenbeispiel.
- (b) Falsch, $f = x_1^3$ bei $x_0 = 0$ ist ein Gegenbeispiel.
- (c) Falsch, $f = -x_1^4$ bei $x_0 = 0$ ist ein Gegenbeispiel.
- (d) Wahr, denn alle anderen Aussagen sind tatsächlich falsch.

Solution of 6.9:

- (a) Wir verwenden Lagrange-Multiplikatoren. Zunächst beobachten wir, dass $\mathbb{S}^1 = g^{-1}(0)$, wobei $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben ist durch $g(x, y) = x^2 + y^2 - 1$. Daher ist eine Lagrange-Funktion $L: \mathbb{R}^3 \rightarrow \mathbb{R}$ für unser Problem gegeben durch

$$L(x, y, \lambda) = f(x, y) - \lambda g(x, y)$$

Wir gehen vor wie im Beispiel 11.7. Gemäß 11.6 müssen an einem lokalen Extremum $p = (x, y) \in g^{-1}(0)$ die folgenden Gleichungen erfüllt sein:

$$0 = \partial_x L(x, y, \lambda) = 2x(2 - \lambda) - 1, \tag{1}$$

$$0 = \partial_y L(x, y, \lambda) = y(2 - 2\lambda), \tag{2}$$

$$0 = \partial_\lambda L(x, y, \lambda) = -(x^2 + y^2 - 1). \tag{3}$$

Wir machen eine Fallunterscheidung basierend auf Gleichung (2):

- (i) Wenn $y = 0$: Dann ist $x^2 = 1$ gemäß 3, d.h., $x = \pm 1$. Daher gibt es lokale Extrema bei $(1, 0)$ und $(-1, 0)$ mit Funktionswerten $f(1, 0) = 1$ und $f(-1, 0) = 2$.
- (ii) Wenn $\lambda = 1$: Dann ist $x = \frac{1}{2}$ gemäß 1. Einsetzen in 3 ergibt $y = \pm \frac{\sqrt{3}}{2}$. Daher gibt es zusätzliche lokale Extrema bei $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ und $(\frac{1}{2}, -\frac{\sqrt{3}}{2})$ mit Funktionswerten $f(\frac{1}{2}, \frac{\sqrt{3}}{2}) = f(-\frac{1}{2}, -\frac{\sqrt{3}}{2}) = \frac{3}{4}$.

Durch Vergleich der Funktionswerte an allen lokalen Extrema sehen wir, dass die Funktion f auf \mathbb{S}^1 ein globales Maximum bei $(-1, 0)$ und zwei globale Minima bei $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ und $(\frac{1}{2}, -\frac{\sqrt{3}}{2})$ hat.

- (b) Da wir den Rand $\mathbb{S}^1 = \partial\mathbb{D}$ bereits in Teil (a) untersucht haben, müssen wir nur das Innere auf Extrema überprüfen. Dazu berechnen wir

$$Df(x, y) = (4x - 1, 2y).$$

An einem lokalen Extremum $p = (x, y)$ von f müssen wir haben

$$0 = Df(x, y) = (4x - 1, 2y),$$

d.h., $p = (x, y) = (\frac{1}{4}, 0)$. Offensichtlich liegt p im Inneren der Einheitskreisscheibe $\mathbb{D}^\circ = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$. Wir berechnen die Hesse-Matrix

$$H_f(x, y) = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix},$$

die positiv definit ist, was auf ein lokales Minimum bei p hinweist.

Die Funktionswert von f an p ist $f(p) = f(\frac{1}{4}, 0) = -\frac{1}{8}$. Vergleichen dieses Funktionswert mit denjenigen aus Teil (a), sehen wir, dass f auf der abgeschlossenen Einheitskreisscheibe \mathbb{D} ein globales Minimum an $(\frac{1}{4}, 0)$ und ein globales Maximum an $(-1, 0)$ hat.

Solution of 6.10: Die Lagrange-Funktion, die zu f und M gehört, ist gegeben durch

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

An einem lokalen Extremum $(x, y, z) \in M$ müssen die folgenden Gleichungen erfüllt sein:

$$\begin{aligned} \partial_x L(x, y, z, \lambda_1, \lambda_2) &= 3 - 2\lambda_1 x - \lambda_2 = 0 \\ \partial_y L(x, y, z, \lambda_1, \lambda_2) &= -1 - 2\lambda_1 y - \lambda_2 = 0 \\ \partial_z L(x, y, z, \lambda_1, \lambda_2) &= 2 - 2\lambda_1 z = 0 \\ \partial_{\lambda_1} L(x, y, z, \lambda_1, \lambda_2) &= -(x^2 + y^2 + z^2 - 1) = 0 \\ \partial_{\lambda_2} L(x, y, z, \lambda_1, \lambda_2) &= -(x + y) = 0 \end{aligned}$$

Aus $-x - y = 0$ folgt $y = -x$, daher $z^2 = 1 - 2x^2 \implies z = \pm\sqrt{1 - 2x^2}$. Aus $2 - 2\lambda_1 z = 0$ folgt $\lambda_1 = \frac{1}{z} = \frac{1}{\pm\sqrt{1 - 2x^2}}$. Außerdem haben wir

$$3 - 2\lambda_1 x = -1 + 2\lambda_1 x \implies 1 = \lambda_1 x.$$

Also, $z = x = -y$ und daher $x^2 = 1 - 2x^2 \iff 3x^2 = 1 \iff x = \pm\frac{1}{\sqrt{3}}$. Durch Einsetzen in f erhalten wir

$$f\left(\frac{1}{\sqrt{3}}(1, -1, 1)\right) = \frac{6}{\sqrt{3}} = 2\sqrt{3}, \quad f\left(-\frac{1}{\sqrt{3}}(1, -1, 1)\right) = -\frac{6}{\sqrt{3}} = -2\sqrt{3}.$$

Somit haben wir alle Extrema von f auf M gefunden, und f erreicht ein Maximum bei $\frac{1}{\sqrt{3}}(1, -1, 1)$ und ein Minimum bei $-\frac{1}{\sqrt{3}}(1, -1, 1)$.

Solution of 6.11: Zunächst, wenn $a = b = 0$, ist f die Nullfunktion. Alle Punkte sind kritisch und gleichzeitig globale Maxima und Minima.

Für kritische Punkte $(x, y) \in \mathbb{R}^2$ berechnen wir den Gradienten

$$\nabla f = \begin{bmatrix} 2x(a - ax^2 - by^2)e^{-x^2 - y^2} \\ 2y(b - ax^2 - by^2)e^{-x^2 - y^2} \end{bmatrix} = 0.$$

Um die kritischen Punkte zu klassifizieren, berechnen wir die Hesse-Matrix $H(x, y)$:

$$\begin{bmatrix} 2(a - 5ax^2 - by^2 + 2ax^4 + 2bx^2y^2) & 4xy(ax^2 - a + by^2 - b) \\ 4xy(ax^2 - a + by^2 - b) & 2(b - 5by^2 - ax^2 + 2by^4 + 2ax^2y^2) \end{bmatrix} e^{-x^2-y^2}.$$

- Wenn $a = 0, b \neq 0$, dann

$$\nabla f = \begin{bmatrix} -2xby^2e^{-x^2-y^2} \\ 2y(b - by^2)e^{-x^2-y^2} \end{bmatrix} = 0.$$

Aus der zweiten Koordinate folgern wir $y = 0$ oder $y = -1$ oder $y = 1$. Wenn $y = 0$, kann x beliebig sein, solange die Bedingung der ersten Koordinate erfüllt ist, sonst $x = 0$.

Also sind die kritischen Punkte $(x, 0)$ mit $x \in \mathbb{R}$ beliebig, sowie $(0, -1)$ und $(0, 1)$. Für $a = 0$:

$$H(x, 0) = \begin{bmatrix} 0 & 0 \\ 0 & 2b \end{bmatrix} e^{-x^2}.$$

Die Matrix ist singular, wir beobachten dies auch, da die kritischen Punkte auf der gesamten Linie $\mathbb{R} \times \{0\}$ liegen.

Mit $H(0, -1) = H(0, 1) = \begin{bmatrix} -2b & 0 \\ 0 & -4b \end{bmatrix} e^{-1}$ sind $(0, -1)$ und $(0, 1)$ sowohl Minima (falls $b < 0$) als auch Maxima (falls $b > 0$).

- Der Fall $a \neq 0, b = 0$ folgt symmetrisch zu (degenerierten) kritischen Punkten $(0, y)$ mit $y \in \mathbb{R}$ beliebig, sowie Minima $(-1, 0)$ und $(1, 0)$ falls $a < 0$ und sonst Maxima.
- Der Fall $a \neq 0, b \neq 0$: Wenn $x = 0$, dann ebenso $y = 0$ oder $y = 1$ oder $y = -1$. Analog für $y = 0$. Wir erhalten kritische Punkte $(0, 0), (0, 1), (0, -1), (1, 0), (-1, 0)$. Wenn $x \neq 0$ und $y \neq 0$, dann $a - ax^2 - by^2 = 0 = b - ax^2 - by^2$. Subtrahieren dieser Gleichungen ergibt $a = b$. Wir behandeln diese zusätzlichen Punkte separat (siehe unten). Wir berechnen

$$H(0, 0) = \begin{bmatrix} 2a & 0 \\ 0 & 2b \end{bmatrix}, H(0, \pm 1) = \begin{bmatrix} 2(a - b) & 0 \\ 0 & -4b \end{bmatrix} e^{-1}, H(\pm 1, 0) = \begin{bmatrix} -4a & 0 \\ 0 & 2(b - a) \end{bmatrix} e^{-1}$$

- Wenn $a > b > 0$, haben wir ein Minimum bei $(0, 0)$, zwei Sattelpunkte bei $(0, \pm 1)$ und zwei Maxima bei $(\pm 1, 0)$.
- Wenn $b > a > 0$, haben wir ein Minimum bei $(0, 0)$, zwei Maxima bei $(0, \pm 1)$ und zwei Sattelpunkte bei $(\pm 1, 0)$.
- Wenn $a > 0 > b$, haben wir einen Sattelpunkt bei $(0, 0)$, zwei Minima bei $(0, \pm 1)$ und zwei Maxima bei $(\pm 1, 0)$.
- Wenn $b > 0 > a$, haben wir einen Sattelpunkt bei $(0, 0)$, zwei Maxima bei $(0, \pm 1)$ und zwei Minima bei $(\pm 1, 0)$.
- Wenn $0 > a > b$, haben wir ein Maximum bei $(0, 0)$, zwei Minima bei $(0, \pm 1)$ und zwei Sattelpunkte bei $(\pm 1, 0)$.

- Wenn $0 > b > a$, haben wir ein Maximum bei $(0, 0)$, zwei Sattelpunkte bei $(0, \pm 1)$ und zwei Minima bei $(\pm 1, 0)$.
- Im Fall $a = b$ finden wir zusätzliche kritische Punkte: Da $a \neq 0$ ist, ist $1 - x^2 - y^2 = 0$. Somit liegt (x, y) auf dem Kreis mit Radius 1.

Für Punkte mit $x^2 + y^2 = 1$ und $a = b$ ist die Hesse-Matrix

$$H(x, y) = \begin{bmatrix} -4ax^2 & -4axy \\ -4axy & -4ay^2 \end{bmatrix} e^{-1}.$$

Sie hat Determinante 0, ist also singulär, wie erwartet, da der gesamte Kreis $x^2 + y^2 = 1$ eine kritische Menge ist.