Technische Universität München
Zentrum Mathematik
Prof. Dr. P. Gritzmann, Dipl.-Inf. Dipl.-Math. S. Borgwardt

Optimierung 3, SS 2009

Übungsblatt 13

Das Blatt dient zur selbstständigen Wiederholung einiger Vorlesungsinhalte. Ein Lösungsvorschlag wird in der letzten Vorlesungswoche auf die Homepage zur Vorlesung gestellt.

Aufgabe 13.1

a) Beweisen oder widerlegen Sie: Sei $f : \mathbb{R} \to \mathbb{R}$ eine konvexe Funktion. Dann ist $|f|$ definiert durch $|f|(x) := |f(x)|$ für alle $x \in \mathbb{R}$ eine konvexe Funktion.

b) Beweisen oder widerlegen Sie: Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine konvexe, beliebig oft stetig differenzierbare Funktion. Dann ist die Richtungsableitung $f'(x; s)$ eine konvexe Funktion.
Hinweis: Beachten Sie, dass sowohl x als auch s freie Variablen sind.

Lösung zu Aufgabe 13.1

a) Die Behauptung ist falsch. Betrachte z.B. die konvexe Funktion $f(x) := x^2 - 1$. Dann ist $|f|(x) = |x^2 - 1|$, und damit $|f|(-1) = 0$ und $|f|(1) = 0$, aber $|f|(0) = 1$.

b) Die Behauptung ist falsch. Betrachte z.B. $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) := x^4$. f ist offensichtlich konvex und beliebig oft stetig differenzierbar. Wählen wir die Richtung $s = 1$, so ist $f'(x; s) = f'(x)^T s = f'(x) = 4x^3$, was keine konvexe Funktion ist.

Aufgabe 13.2

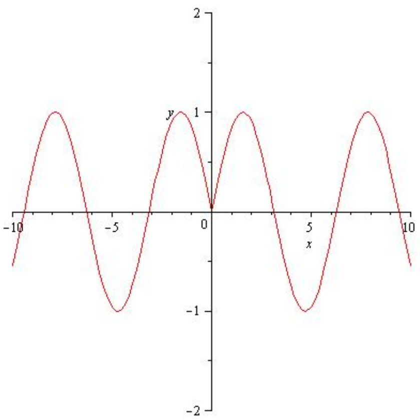
Betrachten Sie die Funktion $f : \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) := \sin |x|$.

a) Skizzieren Sie den Epigraphen von f. Ist f differenzierbar? Begründen Sie Ihre Antwort.

b) Bestimmen Sie $\mathrm{conv}[f]$.

Lösung zu Aufgabe 13.2

a) Der Plot von f sieht in einem Bereich um die 0 wie folgt aus:



Auf $\mathbb{R}\backslash\{0\}$ ist f offensichtlich differenzierbar, allerdings nicht im Nullpunkt, denn $\lim\limits_{x\to 0_+} f'(x) = \lim\limits_{x\to 0_+} \cos x = 1$ und $\lim\limits_{x\to 0_-} f'(x) = \lim\limits_{x\to 0_+} -\cos x = -1$

b) Es gilt $\operatorname{conv}[f](x) = -1$ für alle $x \in \mathbb{R}$, denn f nimmt in allen Punkten der Menge $\{\pm\frac{3}{2}\pi \cdot k : k \in \mathbb{N}\}$ den global minimalen Wert -1 an.

Aufgabe 13.3

Betrachten Sie das folgende nichtlineare Optimierungsproblem P:

$$\begin{array}{rcl} \min & x_1^5 + x_2^2 & \\ x_1 & \geq & 1 \\ x_2 & \geq & 1 \\ x_1 + x_2 & \leq & 3 \end{array}$$

a) Skizzieren Sie den Zulässigkeitsbereich C von P und zeigen Sie die Gültigkeit der Slater-Bedingung.

b) Zeigen Sie, dass P ein konvexes Optimierungsproblem ist.

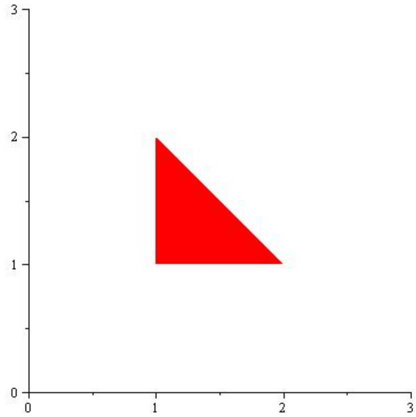
c) Berechnen Sie für alle Randpunkte x von C den Subdifferentialkegel $D_C(x)$.

d) Stellen Sie die KKT-Bedingungen in Sattelpunktform für P auf.

e) Zeigen Sie die Optimalität des Punktes $x^* = (1,1)^T$ für P.

Lösung zu Aufgabe 13.3

a) Der Zulässigkeitsbereich C sieht wie folgt aus:



Der Punkt $(\frac{4}{3}, \frac{4}{3})^T$ liegt im echten Inneren von C, da alle Ungleichungen strikt erfüllt werden. Damit ist die Slater-Bedingung erfüllt.

b) Die Nebenbedingungen sind alle linear, definieren also jeweils einzeln konvexe, abgeschlossene Halbräume. Der Zulässigkeitsbereich ist damit als Schnitt konvexer Mengen konvex. Die Zielfunktion $x_1^5 + x_2^2$ ist als Summe konvexer Funktionen konvex, da x_2^2 konvex ist, und da x_1^5 für $x_1 \geq 0$ konvex ist, und wir unsere Zulässigkeit (unter anderem) auf $x_1 \geq 1$ beschränken.

c) Sei $x \in C$, dann ist $D_C(x) := \mathrm{pos}(\bigcup_{i \in I(x)} \partial\alpha_i(x))$. Wir untersuchen im Folgenden bezüglich des 'Zulässigkeitsbereichsdreiecks' die Punkte auf den drei Kanten und die drei Ecken. Es gilt die Slater-Bedingung, also gilt $D_C(x) = N_C(x)$. Man kann also $D_C(x)$ oder $N_C(x)$ betrachten. Wir berechnen im Folgenden $D_C(x)$. Die Nebenbedingungen unseres Problems nummerieren wir als (1), (2) und (3).

- $I(x) := \{(1)\}$: $\partial\alpha_1(x) = \{\alpha_1'(x)\} = \{(-1, 0)^T\} \Rightarrow D_C(x) = \{(-a, 0)^T : a \in \mathbb{R}_0^+\}$
- $I(x) := \{(2)\}$: $\partial\alpha_2(x) = \{\alpha_2'(x)\} = \{(0, -1)^T\} \Rightarrow D_C(x) = \{(0, -a)^T : a \in \mathbb{R}_0^+\}$
- $I(x) := \{(3)\}$: $\partial\alpha_3(x) = \{\alpha_3'(x)\} = \{(1, 1)^T\} \Rightarrow D_C(x) = \{(a, a)^T : a \in \mathbb{R}_0^+\}$
- $I(x) := \{(1), (2)\}$: $D_C(x) = \{(-a, -b)^T : a, b \in \mathbb{R}_0^+\}$
- $I(x) := \{(1), (3)\}$: $D_C(x) = \{(-a + b, b)^T : a, b \in \mathbb{R}_0^+\}$
- $I(x) := \{(2), (3)\}$: $D_C(x) = \{(b, -a + b)^T : a, b \in \mathbb{R}_0^+\}$

d) Die KKT-Bedingungen in Sattelpunktform für P lauten wie folgt (vgl. 8.3.19, 8.3.21 und 8.3.25):

Ein Vektor $x^* := (\xi_1^*, \xi_2^*)^T$ ist genau dann Minimalpunkt von f auf C, wenn es ein

$y^* := (\eta_1^*, \eta_2^*, \eta_3^*)^T \in [0,\infty[^3$ gibt, so dass

$$\begin{aligned} -\xi_1 + 1 &\leq 0 & (1)\\ -\xi_2 + 1 &\leq 0 & (2)\\ \xi_1 + \xi_2 - 3 &\leq 0 & (3)\\ 5\xi_1^4 - \eta_1^* + \eta_3^* &= 0 & (4)\\ 2\xi_2 - \eta_2^* + \eta_3^* &= 0 & (5)\\ \eta_1^*(-\xi_1+1) + \eta_2^*(-\xi_2+1) + \eta_3^*(\xi_1+\xi_2-3) &= 0 & (6) \end{aligned}$$

e) Für den Beweis genügt es in unserem Fall (konvexes Problem, Regularitätsbedingung) zu zeigen, dass $x^* = (1,1)^T$ die aufgestellten KKT-Bedingungen erfüllt.

- (1): $-1+1 \leq 0$ ok
- (2): $-1+1 \leq 0$ ok
- (3): $1+1-3 \leq 0$ ok
- $(4): 5 - \eta_1^* + \eta_3^* = 0$, also $\eta_1^* = 5 - \eta_3^*$
- $(5): 2 - \eta_2^* + \eta_3^* = 0$, also $\eta_2^* = 2 - \eta_3^*$
- (6) liefert $\eta_3^*(1+1-3) = 0$, also $\eta_3^* = 0$ und $\eta_1^* = 5$, $\eta_2^* = 2$ (und damit $\eta_1^*, \eta_2^*, \eta_3^* \geq 0$)

Also ist $(x^*, y^*)^T = (1,1,5,2,0)^T$ ein Sattelpunkt der Lagrangefunktion, und damit x^* minimal für P.

Aufgabe 13.4

Betrachten Sie das Problem $\min\limits_{x\in\mathbb{R}^2} f(x)$ für die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x) := \xi_1^2 + 2\xi_1\xi_2 + \xi_2^2$ für $x = (\xi_1, \xi_2)^T$. Sei ferner $x_0 := (1,1)^T$.

a) Führen Sie den ersten Schritt des Gradientenverfahrens mit Startpunkt x_0 und exakter Schrittweitenbestimmung durch.

b) Führen Sie den ersten Schritt des Gradientenverfahrens mit Startpunkt x_0 und Schrittweitenbestimmung nach der Armijo-Regel mit $\rho = \beta = \frac{1}{2}$ und $\gamma = 1$ durch.

c) Bestimmen Sie für den ersten Schritt eines Gradientenverfahrens mit Startpunkt x_0 alle Schrittweiten, die den Wolfe-Powell-Bedingungen (in Abhängigkeit von $0 < \rho < \kappa < 1$) genügen.

Lösung zu Aufgabe 13.4

a) Es gilt $f'(x) = (2\xi_1 + 2\xi_2, 2\xi_1 + 2\xi_2)^T$, also wählen wir als Suchrichtung $s := (s_1, s_2)^T = -\frac{f'(x)}{\|f'(x)\|} = \frac{(-4,-4)^T}{\|(-4,-4)^T\|} = (-\frac{1}{2}\sqrt{2}, -\frac{1}{2}\sqrt{2})^T$. Nun müssen wir die exakte Schrittweite bestimmen.

Es soll nun also insbesondere $\frac{d}{d\lambda}f(x_0+\lambda s)=0$ sein:

$$\begin{aligned}
0 &= \frac{d}{d\lambda}f(x_0+\lambda s) = \\
&= \frac{d}{d\lambda}((\xi_1+\lambda s_1)^2+2(\xi_1+\lambda s_1)(\xi_2+\lambda s_2)+(\xi_2+\lambda s_2)^2) = \\
&= 2(\xi_1+\lambda s_1)s_1+2(\xi_1+\lambda s_1)s_2+2(\xi_2+\lambda s_2)s_1+2(\xi_2+\lambda s_2)s_2 = \\
&= 4\cdot(2(1-\frac{1}{2}\sqrt{2}\lambda)(-\frac{1}{2}\sqrt{2})) = \\
&= -\frac{1}{2}\sqrt{2}\cdot 8(1-\frac{1}{2}\sqrt{2}\lambda) = 4\sqrt{2}(\frac{1}{2}\sqrt{2}\lambda-1)
\end{aligned}$$

Damit erhalten wir $\lambda=\sqrt{2}$ und $x_1=x_0+\lambda s=(0,0)^T$.

b) Erneut wählen wir $s=(-\frac{1}{2}\sqrt{2},-\frac{1}{2}\sqrt{2})^T$. Diesmal bestimmen wir die Schrittweite nach der Armijo-Regel, wir suchen also den Armijo-Exponenten i^* und setzen dann $\lambda=\beta^{i^*}\gamma$.

Dazu beginnen wir unsere Tests bei $i=0$:

$$\begin{aligned}
& f(x_0+\beta^i\gamma s)\leq f(x_0)+\beta^i\gamma\rho(f'(x_0))^T s \\
\Leftrightarrow\ & f((1,1)^T+\frac{1}{2}^0\cdot 1\cdot(-\frac{1}{2}\sqrt{2},-\frac{1}{2}\sqrt{2})^T)\leq f((1,1)^T)+\frac{1}{2}^0\cdot 1\cdot\frac{1}{2}f'((1,1)^T)^T(-\frac{1}{2}\sqrt{2},-\frac{1}{2}\sqrt{2})^T \\
\Leftrightarrow\ & f((1-\frac{1}{2}\sqrt{2},1-\frac{1}{2}\sqrt{2})^T)\leq f((1,1)^T)+\frac{1}{2}f'((1,1)^T)^T(-\frac{1}{2}\sqrt{2},-\frac{1}{2}\sqrt{2})^T \\
\Leftrightarrow\ & 4(1-\frac{1}{2}\sqrt{2})^2\leq 4+\frac{1}{2}(4,4)(-\frac{1}{2}\sqrt{2},-\frac{1}{2}\sqrt{2})^T \\
\Leftrightarrow\ & 4(1-\sqrt{2}+\frac{1}{2})\leq 4-2\sqrt{2} \\
\Leftrightarrow\ & 2-2\sqrt{2}\leq 0
\end{aligned}$$

Diese Aussage ist wahr, also ist bereits $i^*=0$ und wir erhalten die Schrittweite $\lambda=\beta^{i^*}\gamma=\frac{1}{2}^0\cdot 1=1$. Es ergibt sich

$$x_1=x_0+\lambda s=(1,1)^T+1\cdot(-\frac{1}{2}\sqrt{2},-\frac{1}{2}\sqrt{2})^T=(1-\frac{1}{2}\sqrt{2},1-\frac{1}{2}\sqrt{2})^T$$

c) Wir benutzen erneut $s:=(-\frac{1}{2}\sqrt{2},-\frac{1}{2}\sqrt{2})^T$. Wir kontrollieren nacheinander die Schrittweiten, die die zwei Wolfe-Powell Bedingungen erfüllen.

WP1:

$$\begin{aligned}
& f(x_0+\lambda s) \leq f(x_0)+\lambda\rho(f'(x_0))^T s \\
\Leftrightarrow\quad & f((1,1)^T+\lambda(-\frac{1}{2}\sqrt{2},-\frac{1}{2}\sqrt{2})^T) \leq f((1,1)^T)+\lambda\rho f'((1,1)^T)^T(-\frac{1}{2}\sqrt{2},-\frac{1}{2}\sqrt{2})^T \\
\Leftrightarrow\quad & 4(1-\lambda\frac{1}{2}\sqrt{2})^2 \leq 4-\lambda\rho 4\sqrt{2} \\
\Leftrightarrow\quad & 1-\lambda\sqrt{2}+\lambda^2\frac{1}{2} \leq 1-\lambda\rho\sqrt{2} \\
\Leftrightarrow\quad & \lambda\rho\sqrt{2}-\lambda\sqrt{2}+\lambda^2\frac{1}{2} \leq 0 \\
\Leftrightarrow\quad & \lambda(\rho\sqrt{2}-\sqrt{2}+\lambda\frac{1}{2}) \leq 0 \\
\Leftrightarrow\quad & \lambda=0 \text{ oder } \lambda \leq (1-\rho)\cdot 2\sqrt{2}
\end{aligned}$$

WP2:

$$\begin{aligned}
& (f'(x_0+\lambda s))^T s \geq \kappa(f'(x_0))^T s \\
\Leftrightarrow\quad & f'((1-\lambda\frac{1}{2}\sqrt{2},1-\lambda\frac{1}{2}\sqrt{2})^T)(-\frac{1}{2}\sqrt{2},-\frac{1}{2}\sqrt{2})^T \geq \kappa f'((1,1)^T)^T(-\frac{1}{2}\sqrt{2},-\frac{1}{2}\sqrt{2})^T \\
\Leftrightarrow\quad & -4\sqrt{2}(1-\lambda\frac{1}{2}\sqrt{2}) \geq \kappa\cdot(-4\sqrt{2}) \\
\Leftrightarrow\quad & 1-\lambda\frac{1}{2}\sqrt{2} \leq \kappa \\
\Leftrightarrow\quad & \sqrt{2}(1-\kappa) \leq \lambda
\end{aligned}$$

Also erfüllen alle $\lambda \in [\sqrt{2}(1-\kappa), 2\sqrt{2}(1-\rho)]$ die beiden Wolfe-Powell Bedingungen.

Aufgabe 13.5
Sie arbeiten für eine große Konsumgüterfirma, die künftig auch im Ausland Fuß fassen möchte. Dazu wurden bereits Verträge mit m Einzelhandelsunternehmen abgeschlossen, für die Standorte $y_1,\dots,y_m$ und die gewünschte Liefermengen $b_1,\dots,b_m$ bekannt sind. Die Marketingexperten Ihres Unternehmens haben bereits eine strategische Planung ausgearbeitet, die n Produktionsstätten mit Produktionskapazitäten $a_1,\dots,a_n$ im Ausland vorsieht. Von diesen Standorten aus müssen die Einzelhändler beliefert werden, die Lieferkosten sind dabei proportional zur (euklidischen) Distanz und zur Liefermenge (es gelte also $Kosten = Distanz \cdot Menge$). Mit der Auswahl der Standorte $x_1,\dots,x_n$ und der Planung der Liefermengen m_{ij} von Standort i zum Einzelhändler j haben sich die Planer bisher aber nicht befasst.

Formulieren Sie das Problem, die Standorte und Liefermengen zu bestimmen, als nichtlineares Optimierungsproblem.

Lösung zu Aufgabe 13.5
Wir wollen natürlich die Gesamtkosten der Standortwahl und des Lieferplanes minimieren. Dazu definieren wir zunächst die euklidische Distanz vom Produktionsstandort x_i zum Einzelhändler y_j als $d_{ij} := dist(x_i, y_j) = \|x_i - y_j\|_2$. Mit m_{ij} bezeichnen wir die Liefermenge von

x_i nach y_j. Damit lautet die Zielfunktion

$$\min \sum_{i=1}^{n} \sum_{j=1}^{m} d_{ij} \cdot m_{ij}.$$

Die Nebenbedingungen beinhalten neben obigen Definitionen noch eine Art „‘Flussbedingung“’, die Produktionskapazitäten und Liefermengen widerspiegelt:

$$\begin{aligned}
\min \sum_{i=1}^{n} \sum_{j=1}^{m} d_{ij} \cdot m_{ij} & \\
d_{ij} = \|x_i - y_j\|_2 & \qquad \forall (i,j) \in [n] \times [m] \\
\sum_{i=1}^{n} m_{ij} = b_j & \qquad \forall j \in [m] \\
\sum_{j=1}^{m} m_{ij} \leq a_i & \qquad \forall i \in [m] \\
m_{ij} \geq 0 & \qquad \forall (i,j) \in [n] \times [m] \\
x_i \in \mathbb{R}^2 & \qquad \forall i \in [n]
\end{aligned}$$

Sind die Standorte x_i festgelegt, so sind auch die d_{ij} nicht mehr variabel und das Problem wird linear. Genau genommen haben wir das bekannte Transportproblem vorliegen, das wir mit Algorithmen aus der Netzwerkfluss-Optimierung leicht lösen können.