Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

Fachbereich Informatik
Dr. Marco Hülsmann

Nichtlineare Optimierung
Übungsblatt 6, WS 2019/20
– Musterlösungen –

Aufgabe 1 (Duale Lagrange-Funktion)

Beweisen Sie Satz 3.10, also zeigen Sie, daß für alle $\mu \leq 0$ und $\lambda \in \mathbb{R}^m$ gilt:

$$\tilde{L}(\lambda, \mu) \leq p^*,$$

wobei p^* der optimale Wert des primalen Optimierungsproblems und $\tilde{L}$ die duale Lagrange-Funktion ist.

Musterlösung:

Ist die zulässige Menge G leer, so ist $p^* = \infty$, und die Ungleichungs ist trivialerweise erfüllt. Sei ansonsten $x \in G$ ein zulässiger Punkt, es gelte also

$$\forall_{i=1,\ldots,m}\ g_i(x) = 0\ \wedge\ \forall_{j=1,\ldots,\ell}\ h_j(x) \leq 0$$

Wegen $\mu \leq 0$ gilt

$$-\sum_{i=1}^{m} \lambda_i(x) g_i(x) - \sum_{j=1}^{\ell} \mu_j h_j(x) \leq 0\ \Rightarrow\ \mathcal{L}(x; \lambda, \mu) \leq f(x)$$

Dies gilt auch, wenn man auf beiden Seiten das Infimum anwendet und insbesondere dann auch für die optimale Lösung $x^* \in G$, also:

$$\tilde{\mathcal{L}}(x; \lambda, \mu) \leq p^*$$

Aufgabe 2 (Dualität)

In der Informationstheorie ist die Entropie der mittlere Informationsgehalt einer Nachricht, welche dementsprechend zu maximieren ist. Sei $x = (x_1, ..., x_n)$ ein Vektor aus Wahrscheinlichkeiten des Auftretens eines Wortes. Die Entropie von x ist dann gegeben durch

$$H(x) := -\sum_{i=1}^{n} x_i \log(x_i)$$

Die Basis des Logarithmus ist irrelevant. Wir wählen hier $\log = \ln$. Die Entropie wird maximal bei Gleichverteilung und minimal im Falle einer Dirac-Verteilung, wo genau ein Wort mit Wahrscheinlichkeit 1 auftritt und alle anderen mit Wahrscheinlichkeit 0 auftreten. Betrachten Sie das folgende nichtlineare Optimierungsproblem mit linearen Nebenbedingungen:

$$\begin{aligned} H(x) = \sum_{i=1}^{n} x_i \log(x_i) & \qquad \text{min!} \\ Ax & \leq b \\ \langle e, x \rangle & = 1 \end{aligned}$$

Dabei sei A eine Matrix, b ein Vektor und e der Vektor, der aus lauter Einsen besteht.

(i) Zeigen Sie, daß die zugehörige Lagrange-Funktion konvex ist.

(ii) Formulieren Sie das zugehörige duale Optimierungsproblem!

Musterlösung:

(i) Für die Lagrange-Funktion gilt

$$\mathcal{L}(x; \lambda, \mu) = \sum_{i=1}^{n} x_i \log(x_i) - \lambda \left(\sum_{i=1}^{n} x_i - 1 \right) - \langle \mu, Ax - b \rangle$$

mit $\lambda \in \mathbb{R}$ und $\mu \in \mathbb{R}^n$. Die zweiten partiellen Ableitungen des rechten Teils von $\mathcal{L}$ sind alle 0, es kommt somit nur auf die Zielfunktion $f(x) = \sum_{i=1}^{n} x_i \log(x_i)$. Für $x_i > 0$ gilt

$$\frac{\partial f}{\partial x_i} = \log(x_i) + 1, \; \frac{\partial^2 f}{\partial x_i^2} = \frac{1}{x_i} > 0$$

und somit ist f auf $\mathbb{R}^n_{>0}$ und daher $\mathcal{L}$ auf $\mathbb{R}^n_{>0} \times \mathbb{R} \times \mathbb{R}^n$ sogar streng konvex. Für $x = 0$ und $y \geq 0$ gilt wegen der Monotonie des Logartihmus für $\vartheta \in [0, 1]$ und $0 \cdot \log(0) = 0$ (l'Hospital!) für $\varphi(x) := x \log(x)$:

$$\begin{aligned} \varphi(\vartheta x + (1 - \vartheta)y) &= \varphi((1 - \vartheta)y) = (1 - \vartheta)y \log((1 - \vartheta)y) \\ &\leq (1 - \vartheta)y \log(y) = (1 - \vartheta)\varphi(y) \end{aligned}$$

Somit ist f auch auf $\mathbb{R}^n_{\geq 0}$ und daher $\mathcal{L}$ auf $\mathbb{R}^n_{\geq 0} \times \mathbb{R} \times \mathbb{R}^n$ konvex.

(ii) Minimiere $\mathcal{L}$:

$$\begin{aligned} & \frac{\partial L}{\partial x_i} = \log(x_1) + 1 - \lambda - \langle \mu, a_i \rangle = 0 \\ \Leftrightarrow \quad & x_i = \exp\left(\langle \mu, a_i \rangle + \lambda - 1\right) \end{aligned}$$

Sei $\mathcal{L}_1(x; \lambda, \mu) := \sum_{i=1}^{n} x_i \log(x_i) - \lambda \sum_{i=1}^{n} x_i - \langle \mu, Ax \rangle$. Dann lautet das duale Optimierungsproblem:

$$\begin{aligned} \lambda + \langle \mu, b \rangle + \mathcal{L}_1\left(\exp\left(\langle \mu, a_i \rangle + \lambda - 1\right)\right) & \quad \max! \\ \mu \leq & \quad 0 \end{aligned}$$