Technische Universität München
Zentrum Mathematik
Lehrstuhl für Mathematische Optimierung, M1
Prof. Dr. Michael Ulbrich
Dennis Klöckner, M.Sc. Wintersemester 2013/14

First Order Primal-Dual Optimization Methods (Primal-duale Optimierungsmethoden erster Ordnung)

Übungsblatt 6

Aufgabe 6.1 (ADMM bei Matrix Vervollständigung):
Wir betrachten für $M \in \mathbb{R}^{m\times n}$ das Problem

$$\min_{X\in\mathbb{R}^{m\times n}} \|X\|_* \quad \text{u.d.N.} \quad X_{ij} = M_{ij}\ \forall (i,j) \in \Omega, \tag{1}$$

wobei $\Omega \subset \{(i,j)\colon i \in \{1,\ldots,m\}, j \in \{1,\ldots,n\}\}$ eine Indexmenge ist und $\|\cdot\|_*$ die nukleare Norm beschreibt, die definiert ist als die Summe der Singulärwerte.

a) Formulieren Sie ein äquivalentes unrestringiertes Problem

$$\min_X \|X\|_* + \delta_{\mathcal{U}}(X)$$

für eine geeignete Menge $\mathcal{U}$. Entkoppeln Sie beide Ausdrücke durch Einführen einer neuen Variablen Y, sodass $X - Y = 0$ und sich die Zielfunktion als $\|X\|_* + \delta_{\mathcal{U}}(Y)$ schreiben lässt, und stellen Sie die augmentierte Lagrangefunktion

$$L^a(\gamma; X, Y, \Lambda)$$

auf, wobei $\Lambda \in \mathbb{R}^{m\times n}$ der entsprechende Lagrangemultiplikator ist und $\gamma > 0$ gilt. Überlegen Sie sich hierbei, wie man diese mit dem Frobeniusskalarprodukt $\langle A, B\rangle := \sum_{i,j} A_{ij}B_{ij}$ und der Frobeniusnorm $\|A\|_F := \sqrt{\langle A, A\rangle_F} = \left(\sum_{i,j} A_{ij}^2\right)^{1/2}$ für $A, B \in \mathbb{R}^{m\times n}$ schreiben kann.

b) Begründen Sie, dass für Iterierte X^k, Λ^k die Lösung des Y-Teilproblems gegeben ist durch

$$\arg\min_Y L^a(\gamma; X^k, Y, \Lambda^k) = P_{\mathcal{U}}\left(X^k + \frac{1}{\gamma}\Lambda^k\right),$$

mit

$$\left[P_{\mathcal{U}}\left(X^k + \frac{1}{\gamma}\Lambda^k\right)\right]_{ij} = \begin{cases} M_{ij} & \text{, falls } (i,j) \in \Omega, \\ \left(X^k + \frac{1}{\gamma}\Lambda^k\right)_{ij} & \text{, sonst.} \end{cases}$$

c) Zeigen Sie, dass für gegebene Iterierte Y^{k+1}, Λ^k die Lösung des X-Teilproblems gegeben ist durch

$$\arg\min_X L^a(\gamma; X, Y^{k+1}, \Lambda^k) = \operatorname{prox}_{\frac{1}{\gamma}\|\cdot\|_*}(Y^{k+1} - \frac{1}{\gamma}\Lambda^k).$$

d) Schreiben Sie den gesamten Algorithmus (inklusive Multiplikatorupdates) auf und begründen Sie kurz, dass die Voraussetzungen von Konvergenzsatz 3.4 erfüllt sind.

   Hinweis: Sie dürfen hierbei verwenden, dass

$$\operatorname{prox}_{\tau\|\cdot\|_*}(Y) = \arg\min_X \tau\|X\|_* + \frac{1}{2}\|X - Y\|_F^2 = \mathcal{D}_\tau(Y)$$

   gilt, wobei $\mathcal{D}_\tau$ der sogenannte Singulärwert Shrinkage Operator ist, definiert als

$$D_\tau(X) := U\mathrm{Diag}\left(\max\{\sigma_i - \tau, 0\}\right) V^*,$$

   und $X = U\mathrm{Diag}(\sigma_i)V^*$ die Singulärwertzerlegung von X mit den Singulärwerten σ_i ist. Außerdem definiert $\mathrm{Diag}(v)$ eine Diagonalmatrix, die den Vektor v auf der Diagonalen enthält.

Aufgabe 6.2 (Dualität):
Wir betrachten

$$\min_{x\in C} f(x) \quad \text{u.d.N.} \quad Ax = b$$

für eine Funktion $f\colon C \to \mathbb{R}$, eine nichtleere Menge $C \subset \mathbb{R}^n$ und $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$ und definieren $d\colon \mathbb{R}^m \to \mathbb{R}$ als

$$d(\lambda) := \inf_{x\in C} L(x, \lambda) = \inf_{x\in C} \Big(f(x) + \lambda^T(Ax - b)\Big)$$

mit der Lagrangefunktion $L(x, \lambda)$ bezüglich der Gleichungsnebenbedingung $Ax = b$. Es bezeichne weiter $f^* := \inf_{x\in X} f(x)$ und $d^* := \sup_{\lambda\in\mathbb{R}^m} d(\lambda)$, wobei $X := \{x \in C : Ax = b\}$.

a) Wir betrachten erneut das Beispiel aus Aufgabe 2.2 b) mit $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(x) := x_1$, $A := (0, 1)$, $b := 0$ und $C := \{x \in \mathbb{R}^2 : x_2 \geq x_1^2\}$. Somit gilt $f^* = 0$ und $\bar{x} = 0$ ist ein zulässiger Punkt mit $f^* = f(\bar{x}) = 0$.

   Zeigen Sie, dass $d^* = 0$ gilt, aber kein $\bar{\lambda} \in \mathbb{R}$ existiert mit $d(\bar{\lambda}) = 0$.

b) Nun untersuchen wir die ganzzahlige Optimierungsaufgabe mit $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(x) := -x_1$, $A := (1, 1)$, $b := 3$ und $C := \{(0, 0), (2, 1), (1, 2), (4, 0), (0, 4)\}$. Zeigen Sie, dass es einen zulässigen Punkt $\bar{x} \in \mathbb{R}^2$ gibt mit $f^* = f(\bar{x})$, und außerdem $\bar{\lambda} \in \mathbb{R}$ existiert mit $d^* = d(\bar{\lambda})$. Folgern Sie, dass hier $d(\bar{\lambda}) = d^* < f^* = f(\bar{x})$ erfüllt ist.

Aufgabe 6.3 (Moreau-Zerlegung):
Es sei $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ konvex, eigentlich und unterhalbstetig, $x \in \mathbb{R}^n$ und $\tau > 0$. Zeigen Sie, dass dann

$$x = \operatorname{prox}_{\tau f}(x) + \tau \operatorname{prox}_{f^*/\tau}(x/\tau)$$

gilt, wobei $\operatorname{prox}_{\tau f}$ wie in Aufgabe 1.3 definiert ist.

Hinweis: Zeigen Sie zunächst, dass $y = \operatorname{prox}_{\tau f}(x)$ äquivalent ist zu $0 \in \partial(\tau f)(y) + (y - x)$. Nutzen Sie weiter die Charakterisierung der Subdifferentiale von f und f^* aus Satz 2.16 um zu zeigen, dass dies äquivalent ist zu $\frac{x-y}{\tau} = \operatorname{prox}_{f^*/\tau}(x/\tau)$ und folgern Sie daraus die Behauptung.