Richard Huber
richard.huber@uni-graz.at

Projekt 4: Primal-Dual Algorithm

Wir wollen uns im Folgenden mit der proximalen Abbildung beschäftigen, und darauf aufbauend den Primalen-Dualen Algorithmus untersuchen, der es erlaubt eine große Klasse von Optimierungsproblemen zu lösen.

Definition. *Es sei X ein separabler Hilbertraum und $f\colon X \to \mathbb{R} \cup \{\infty\}$ sei eine eigentliche, konvexe und unterhalbstetige Funktion und $\sigma > 0$. Dann ist die proximale Abbildung (Proximal Mapping) definiert als*

$$prox_\sigma^f\colon X \to X, \qquad prox_\sigma^f(u_0) = \operatorname*{argmin}_{u\in X} \frac{\|u-u_0\|_X^2}{2} + \sigma f(u). \tag{1}$$

Aufgabe 1) Zeigen Sie, dass die proximale Abbildung wohldefiniert ist, also dass $\frac{\|u-u_0\|_X^2}{2} + \sigma f(x)$ stets eine eindeutige Lösung besitzt. Folgern Sie weiters, dass es für jedes $u_0 \in X$ genau ein $x \in X$ existiert, sodass $u_0 \in [id + \sigma \partial f](x)$.

Der Hintergrund ist, dass die proximalen Optimierungsprobleme als Hilfsprobleme in manchen Optimierungsverfahren – wie dem Primalen-Dualen Algorithmus – zum Einsatz kommen. Insbesondere lassen sich mit diesem viele nicht-differenzierbare Teile des Zielfunktionals bearbeiten, was mit klassischen Abstiegsverfahren nicht möglich wäre. Um dies zu verstehen, müssen wir allerdings erst Fenchel Dualität betrachten.

Definition. *Es sei X ein Hilbertraum (reflexiver Banachraum würde reichen), und $F\colon X \to \mathbb{R} \cup \{\infty\}$ eigentlich. Wir definieren das konjugierte (oder duale) Funktional*

$$F^*\colon X^* \to \mathbb{R} \cup \{\infty\}, \qquad F^*(\xi) = \sup_{u\in X} \langle \xi, u\rangle - F(u) \tag{2}$$

sowie für $G\colon X^* \to \mathbb{R} \cup \{\infty\}$ eigentlich das biduale Funktional

$$G^*\colon X \to \mathbb{R} \cup \{\infty\}, \qquad G^*(u) = \sup_{\xi\in X^*} \langle \xi, u\rangle - G(\xi). \tag{3}$$

Ein zentrales Resultat der konjugierten Funktion ist das Theorem von Fenchel-Monreau-Rockafellar, welches zeigt, dass für konvexe, unterhalbstetige und eigentlichen Funktionen f die Gleichheit $f = f^{**}$ gilt (siehe Projekt 3).

Aufgabe 2) Gegeben seien $F\colon X \to \mathbb{R} \cup \{\infty\}$ und $G\colon Y \to \mathbb{R} \cup \{\infty\}$, beide eigentlich konvex und unterhalbstetig, sowie $A\colon X \to Y$ stetig für separable Hilberträume X und Y. Wir betrachten das Problem

$$\min_{u\in X} F(u) + G(Au) := P(u). \tag{4}$$

Zeigen Sie, dass dies äquivalent zu

$$\min_{u\in X} \sup_{\xi\in Y^*} \langle A^*\xi, u\rangle - G^*(\xi) + F(u) := L(u,\xi) \tag{5}$$

ist. Folgern Sie unter der Annahme, dass man Supremum und Infimum tauschen darf, dass (5) äquivalent zu

$$\sup_{\xi\in Y^*} -F^*(-A^*\xi) - G^*(\xi) := D(\xi). \tag{6}$$

Man nennt (4) das primale Problem und (6) das duale Problem, während (5) als Sattelpunktsproblem bezeichnet wird. Wie bereits in Aufgabe 7.2 angedeutet, stehen die drei Probleme in engem Zusammenhang.

Aufgabe 3) Zeigen Sie, dass wenn es $u^* \in X$, $\xi^* \in X^*$ gibt, sodass $P(u^*) = D(\xi^*)$, so ist u^* Lösung des primalen Problems und ξ^* des dualen Problems, und (u^*, ξ^*) Lösung des Sattelpunktproblems.

Dies motiviert die Idee, das primale- und das duale Problem gleichzeitig zu betrachten und zur Optimierung heranzuziehen. Darauf aufbauend kann der Primale-Duale Algorithmus hergeleitet werden, der gleichzeitig das primale Problem minimiert, das duale Problem maximiert und den Zusammenhang zwischen primalen und dualen Lösungen ausnutzt.

Die Herleitung dieses Algorithmus (basierend auf maximalen monotonen Operatoren) und Konvergenzanalyse ist nicht ganz trivial, daher beschreiben wir im Folgenden nur die resultierende Iteration.

Algorithm 1 Primal-Dual Algorithm

INPUT: Operators A, A^* between Hilbert spaces H_1 and H_2, reals $\sigma, \tau > 0$ such that $\sigma\tau\|A\|^2 < 1$ and proximal mappings $prox_\sigma^F$, $prox_\tau^{G^*}$.

1: Initialise $x^0 = 0 \in H_1$, $\xi^0 = 0 \in H_2$ and $\bar{x} = 0 \in H_1$.
2: **for all** $n = 0, 1, \ldots$ **do**
3: $$\begin{cases} \xi^{n+1} = prox_\sigma^{G^*}(\xi^n + \sigma A\bar{x}) \\ x^{n+1} = prox_\tau^F(x^n - \tau A^* \xi^{n+1}) \\ \bar{x} = 2x^{n+1} - x^n \end{cases}$$

Output: Sequence of pairs $(x^n, \xi^n)_n$ which converge to a solution of (5).

Wir wollen im Folgenden das diskrete Optimierungsproblem für $\Omega = \{1, \ldots, N\} \times \{1, \ldots, N\}$

$$\min_{u \in l^2(\Omega)} \frac{\|k * u - f\|_2^2}{2} + \frac{\|\nabla_N u\|_{l^p}^p}{p} + I_{\cdot \geq 0}(u), \tag{7}$$

für $p \in (1, \infty)$ und $k \in l^1(\Omega)$ betrachten, wobei Faltung unter passenden Erweiterungen mit Null und Einschränkungen zu verstehen ist, analog zu Aufgabe 6.2.

Aufgabe 4) Zeigen Sie, falls $X = X_1 \times X_2$ mit $\langle x, u \rangle_X = \langle x_1, u_1 \rangle_{X_1} + \langle x_2, u_2 \rangle_{X_2}$ für $x = (x_1, x_2)$, $u = (u_1, u_2)$ und $F(x) = F_1(x_1) + F_2(x_2)$, so ist $F^*(\xi) = F_1^*(\xi_1) + F_2^*(\xi_2)$ für $\xi = (\xi_1, \xi_2) \in X^*$ (es gilt $X^* = X_1^* \times X_2^*$). Folgern Sie weiter, dass für solches F die Identität $prox_\sigma^F(x) = \big(prox_\sigma^{F_1}(x_1), prox_\sigma^{F_2}(x_2)\big)^T$ gilt.

Das heißt, für "separierbare" Funktionen sind auch die Konjugierte separierbar und die proximale Abbildung komponentenweise anwendbar.

Aufgabe 5) Bestimmen Sie die relevanten Operationen zur Anwendung des Primal-Dualen Algorithmus auf das Problem (7), im Fall der p-norm in impliziter Form. Implementieren Sie den Algorithmus 1 für die resultierenden Operationen und testen Sie diese für die Daten aus Aufgabe 6.3 und $p \in \{3, 2, 1.5, 1.1\}$. Dabei könnte die Funktion fzero hilfreich sein um den impliziten prox zu bestimmen.