Projekt 3: Fenchel Dualität

Im Folgenden wollen wir uns mit der Fenchel-Dualität von Funktionen beschäftigen, welche ein praktisches Hilfsmittel der Optimierung ist, und in ähnlichem Sinne wie duale Probleme in der linearen Optimierung können diese genutz werden, um Optimierungsprobleme zu vereinfachen.

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

$$F^*\colon X^* \to \mathbb{R} \cup \{\infty\}, \qquad F^*(\xi) = \sup_{u \in X} \langle \xi, u \rangle - F(u) \tag{1}$$

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

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

Wir wollen uns im Folgenden etwas tiefer mit diesen Operatoren beschäftigen, um festzustellen wie sich diese zum Lösen von Problemen nutzen lassen. Eine zentrale Eigenschaft der Konjugation, welche dazu herangezogen werden kann ist, dass $F^{**} = F$ falls F eigentlich, konvex und unterhalbstetig, also dass das zweimalige Konjugieren wieder die Funktion selbst ergibt. Dies im Allgemeinen zu zeigen ist aber nicht ganz trivial, und wir wollen uns langsam heran tasten.

Aufgabe 1) Berechnen Sie die Konjugierten der folgenden (endlichdimensionalen) Funktionen um ein Gefühl für den Konjugationsoperator zu erhalten.

- $f(x) = \frac{1}{2} x^T A x$ für $A \in \mathbb{R}^{n \times n}$ symmetrisch und positiv definit.
- $f(x) = \langle b, x \rangle$ wobei $b \in \mathbb{R}^n$. Beachten Sie insbesondere den Fall $b = 0$.
- gegeben $g\colon \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ und $A \in \mathbb{R}^{n \times m}$ und

$$f(x) = \inf_y g(y) \qquad \text{unter der Nebenbedingung } Ay = x \tag{3}$$

Wir wollen im Folgenden zeigen, dass konjugierte Operatoren gewissen analytische Eigenschaften besitzen, wodurch duale Formulierungen praktisch anwendbar sind.

Aufgabe 2) Es sei $f, g\colon X \to \mathbb{R} \cup \{\infty\}$ eigentlich, mit $f \leq g$ punktweise. Zeigen Sie, dass f^* konvex und unterhalbstetig ist. Zeigen Sie außerdem, dass $g^* \leq f^*$ und folgern Sie unter der Annahme dass f^* eigentlich, dass $f^{**} \leq g^{**}$. Beweisen Sie weiters, dass $f^{**} \leq f$.

Tatsächlich gilt, wenn f eigentlich, konvex und unterhalbstetig ist, dass $f^{**} = f$ (das Fenchel-Moreau-Rockafellar Theorem), wir wollen dies im Folgenden im vereinfachten Fall von differenzierbaren endlichdimensionalen Funktionen betrachten.

Aufgabe 3) Wir betrachten zunächst die Funktion $f(x) = \frac{x^2}{2}$ für $x \in \mathbb{R}$. Bestimmen Sie f^* und skizzieren Sie f sowie für $\xi \in \{-2, -1, 0, 1, 2\}$ die Funktionen $x \mapsto \xi x - f^*(\xi)$ (also was man für f^{**} maximieren muss).

Dies motiviert die Idee, dass die Menge $\{(s, c) \in X^* \times (\mathbb{R} \cup \{\infty\}) \mid \langle s, x \rangle - c \leq f(x)\}$ der affine-linearen Funktionale welche unterhalb von f sind die Menge $\{(\xi, f^*(\xi)) \in X^* \times (\mathbb{R} \cup \{\infty\})\}$ enthält, und tatsächlich $f^*(\xi)$ die größte Konstante ist mit der dies gilt.

Aufgabe 4) Wir nehmen an, dass $f\colon \mathbb{R}^n \to \mathbb{R}$ konvex und differenzierbar auf ganz $\mathbb{R}^n$ sei. Bestimmen Sie für $x \in \mathbb{R}^n$ den Wert $f^*(Df(x))$ und folgern Sie $f^{**} = f$.

Das Theorem von Fenchel-Moreau-Rockafellar im Allgemeinen zu beweisen ist nicht ganz simpel, der obige vereinfachte Fall zeigt aber die grundsätzliche Idee.

Wie wir bereits wissen, ist $f^{**} \leq f$ und f^{**} konvex, tatsächlich gilt dass $\operatorname{conv}(f) = f^{**}$, wobei $\operatorname{conv}(f)$ die größte konvexe Funktion kleiner gleich f bezeichnet.

Aufgabe 5) Es sei $f\colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ eigentlich und für $x \in X$ sei

$$g_f(x) = \inf_{K \in \mathbb{N}} \left\{ \inf_{\substack{x_1,\dots,x_K \in \mathbb{R}^n, \\ \lambda_1,\dots,\lambda_K \geq 0, \sum_{k=1}^K \lambda_k = 1}} \sum_{k=1}^K \lambda_k f(x_k) \qquad \text{unter der Bedingung } \sum_{k=1}^K \lambda_k x_k = x \right\}. \tag{4}$$

Zeigen Sie dass g_f konvex ist, $g_f \leq f$ mit Gleichheit falls f konvex ist, und folgern Sie, dass $\operatorname{conv}(f) = g_f$. Folgern Sie unter Annahme der Anwendbarkeit des Fenchel-Moreau-Rockafellar Theorems auf $\operatorname{conv}(f)$, dass $f^{**} = \operatorname{conv}(f)$.

Nun wollen wir dieses theoretische Wissen auf ein Beispiel aus der Big-Data anwenden.

Aufgabe 6) Wir betrachten die Matrix $A_0 \in \mathbb{R}^{m,n}$ sodass jede Zeile von A_0 den Werten eines Bildes entspricht, wobei wir annehmen, dass alle Bilder das Selbe zeigen, allerdings ist dem aufgrund von Rauschen und ungleicher Skalierung nicht so. Ein Ansatz das echte Bild zu bestimmen, bestünde nun darin

$$\min_{A \in \mathbb{R}^{n,m}} \|A - A_0\|_F + I_{\operatorname{rk}(A) \leq 1} + I_{\|A\|_F \leq \alpha} \tag{5}$$

zu lösen. Das heißt wir suchen eine Rank-1 Approximation. Allerdings ist (5) sehr schwer zu lösen, da nicht konvex. Zeigen Sie, dass $G^{**} = I_{\|A\|_* \leq \alpha}$ für $G = I_{\operatorname{rk}(A) \leq 1} + I_{\|A\|_F \leq \alpha}$ gilt. Dabei bezeichnet $\|A\|_F^2 = \sum_{ij} A_{ij}^2 = \sum_k \sigma_k(A)^2$ die Frobeniusnorm (also Summe der Quadrate bzw. l^2 norm der Singularwerte) welche von $\langle A, B\rangle = \sum_{ij} A_{ij} B_{ij}$ erzeugt wird. Darüber hinaus bezeichnet $\|A\|_* = \sum_k \sigma_k(A)$ die Nuklearnorm, welche dual zu $\|A\|_\infty = \max_k \sigma_k(A)$ ist. Außerdem könnte für Sie hilfreich sein, dass $\operatorname{rk}(A) \leq 1$ genau dann wenn $A = ab^T$ für Zeilenvektoren a, b und es gilt $\|ab^T\|_F = \|a\|_2 \|b\|_2$.

Daher ist es ein natürlicher Schritt, statt (5) das Optimierungsproblem

$$\min_{A \in \mathbb{R}^{n,m}} \|A - A_0\|_F + I_{\|A\|_* \leq \alpha} \tag{6}$$

zu betrachten, da es in gewissen Sinne das ähnlichste konvexe Optimierungsproblem ist.