Seminararbeit

Dualitätstheorie (Rockafellar- Dualität, Fenchel- Dualität und Lagrange- Dualität)

Basierend auf dem Vortrag “Convex Analysis and Duality with Applications” von Qiji Zhu

Henriette Thier, 418782

5. März 2023

Inhaltsverzeichnis

1 Einleitung **1**

2 Bemerkungen **2**

3 Rockafellar-Dualität **2**
3.1 Voraussetzungen 2
3.2 Rockafellar Dualität 4
3.3 Beispiele 6

4 Fenchel-Dualität **8**
4.1 Fenchel Formulierung 8
4.2 Konvexe, unterhalbstetige Funktionen 8
4.3 Polyhedrische Funktionen 10
4.4 “Decoupling” 14
4.5 Beispiele 15

5 Lagrange-Dualität **19**
5.1 Lagrange Formulierung 19
5.2 Lineare Optimierung 21

6 Zusammenfassung **25**

7 Anhang **26**
7.1 Literatur 26
7.2 Weitere Beweise 26

1 Einleitung

Die vorliegende Arbeit widmet sich systematischen Arten und Weisen, duale Optimierungsprobleme zu finden.

Um es zu motivieren und zu verstehen, was duale Optimierungsprobleme sind, schauen wir uns ein Beispiel an:

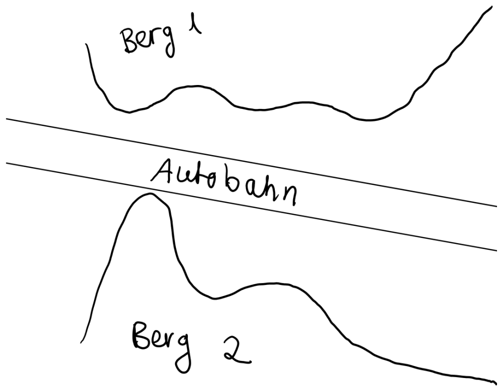

Abbildung 1: Beispiel für ein duales Optimierungsproblem

Unser erstes Optimierungsproblem ist, den minimalen Abstand zweier Berge zueinander herausfinden. Idealisierend gehen wir davon aus, dass die Wände der Berge senkrecht auf dem Boden stehen. Ein weiteres Optimierungsproblem ist, dass man die maximale Breite einer Autobahn herausfinden möchte, welche zwischen den beiden Bergen entlang gehen soll. Die Autobahn soll überall die selbe Breite haben. Wenn man diese beiden Optimierungsprobleme betrachtet, fällt auf, dass

1. Die Breite jeder Autobahn zwischen den Bergen ist kleiner gleich jedem Abstand zwischen den Bergen.

2. Der minimale Abstand zwischen den Bergen ist die maximale Breite der Autobahn.

Wenn man zwei Optimierungsprobleme mit der ersten Eigenschaft hat, sagt man, dass sie schwach dual sind. Falls sie beide Eigenschaften besitzen, dann sind sie (stark) dual zueinander.

Das Kennen des dualen Problems hilft uns auf unterschiedliche Art und Weise:

1. Wir bekommen ein Kriterium, mit welchem wir prüfen können, ob der gegebene Abstand der Berge minimal ist.

2. Wir bekommen ein Kriterium, mit welchem wir prüfen können, ob die gegebene Breite der Autobahn maximal ist.

3. Für das Beantworten des ersten oder zweiten Optimierungsproblems, können wir nun das leichter zu lösende Optimierungsproblem lösen.

Dies führt zu der Frage, ob man für Optimierungsprobleme systematisch duale Optimierungsprobleme formulieren kann.

Im Folgenden schauen wir uns diese Frage an. Dabei widmen wir uns vor allem der **Optimierung von konvexen Funktionen unter linearen Nebenbedingungen**. Diese kommen natürlicherweise in unterschiedlichen Bereichen vor, beispielsweise bei linearen Programmen und im Falle der Entropiefunktion. Außerdem können viele Probleme in der Wirtschaft umformuliert werden zu Optimimierungsproblemen von konvexen Funktionen unter linearen Nebenbedingungen.

Dabei betrachten wir, in welcher Beziehung Dualität und der Lagrange-Multiplikator zueinander stehen.

2 Bemerkungen

Im Folgenden sei der Definitionsbereich von Funktionen ein Banachraum X über den reellen Zahlen. X^* bezeichne den dazugehörigen Dualraum.

Im Folgenden werden wir Langrange-Multiplikatoren benutzen. Für uns reicht es aus, dass l ein Lagrange-Multiplikator von $f(x)$ heißt, falls $l \in -\partial f(x)$. Dabei bezeichne $\partial f(x)$ das Subdifferential von f an der Stelle x, d.h. $\partial f(x) = \{y^* \in X^* : f(y) - f(x) \geq \langle y^*, y - x\rangle \ \forall y \in X\}$.

3 Rockafellar-Dualität

3.1 Voraussetzungen

Definition 3.1.1 (Konvex-Konjugierte/Legendre-Fenchel-Transformation).
Sei $f : X \to \mathbb{R} \cup \{\infty\}$ mit $f(x) \neq \infty$ für ein $x \in X$.

Dann ist die Konvex-Konjugierte der Funktion f:

$$f^* : X^* \to \mathbb{R} \cup \{\infty\}, x^* \mapsto sup_{x \in X}\{\langle x^*, x\rangle - f(x)\}.$$

Bemerkung 3.1.2. Wenn f streng konvex ist, $X \subset \mathbb{R}$ und f nach $\mathbb{R}$ abbildet, nennt man die Konvex-Konjugierte auch Legendre-Transformierte.

Bemerkung 3.1.3. Der Begriff Konvex-Konjugierte kommt daher, dass die Konvex-Konjugierte von f immer konvex ist. Der Beweis ist im Anhang 7.2.

Was also ist die Konvex-Konjugierte? Bildlich nimmt man im $\mathbb{R}^2$ eine Geradenschar $g_i(x) := \langle x^*, x\rangle + y_i$ mit der Steigung x^*. Wir suchen den höchsten y-Abschnitt, welcher erfüllt, dass die Gerade noch unterhalb des Graphen von f liegt.

Also $sup\{y : \langle x^*, x\rangle + y \leq f(x) \ \forall x \in X\} = sup\{y : y \leq -\langle x^*, x\rangle + f(x) \ \forall x \in X\} = inf_{x \in X}\{-\langle x^*, x\rangle + f(x)\} = -sup_{x \in X}\{\langle x^*, x\rangle - f(x)\} = -f^*(x^*)$. Das sagt uns, dass wir f^* genau das Negierte dieses gesuchten y-Abschnitts zuordnen.

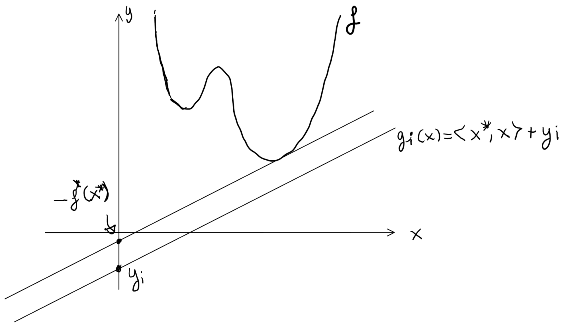

Abbildung 2: Die Konvex-Konjugierte

Die weitere Intuition kommt, wenn man sich folgende Frage stellt: Wie bekommt man die "abgeschlossene, konvexe Hülle" von f in $\mathbb{R}$. Dabei meint die "abgeschlossene, konvexe Hülle einer Funktion", die Funktion, wessen Epigraph die abgeschlossene, konvexe Hülle des Epigraphen von f ist.

Bildlich wollen wir Geraden unter unseren Graphen von f legen und an jedem Punkt den Wert nehmen, der der höchste dieser Geraden ist, um f nahe zu kommen. Wir speichern uns den invertierten y-Achsenabschnitt zu der Steigung y^* der Geraden. Das ist ja genau unser $f^*(y^*)$. Jetzt wollen wir schauen, welche Funktion gerade so über all diesen Geraden liegt, also das Supremum dieser in jedem Punkt ist. Dabei sind die Geraden, welche unter dem Graph von f liegen und diesen berühren, durch $g(x) := \langle y^*, x\rangle - f^*(y^*)$ gegeben. Also ist die "abgeschlossene, konvexe Hülle" von f durch $sup_{y^*\in\mathbb{R}}\{\langle y^*, x\rangle - f^*(y^*)\}$ gegeben. Wenn man sich diesen Term anschaut, ist er nichts anderes als $f^{**}(x)$

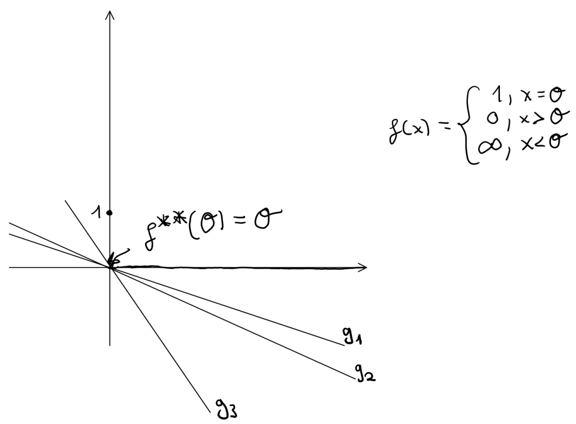

Abbildung 3: Die Bi-Konjugierte

Jetzt wo wir geklärt haben, was die Konvex-Konjugierte ist, kommt unweigerlich die Frage auf, warum wir uns mit dieser beschäftigen. Im ersten Abschnitt zeigen wir mit dieser die Rockafellar Dualität. Zunächst werden wir aber dafür die Fenchel-Young inequality und equality beweisen.

Satz 3.1.4 (Fenchel-Young equality).
Für alle $x \in X$ und alle $x^* \in X^*$ gilt:

$$f(x) + f^*(x^*) = \langle x^*, x\rangle \Leftrightarrow x^* \in \partial f(x).$$

Beweis.

$$\begin{aligned} f(x) + f^*(x^*) = \langle x^*, x\rangle \\ &\Leftrightarrow \langle x^*, x\rangle - f(x) = f^*(x^*) \\ &\Leftrightarrow \langle x^*, x\rangle - f(x) = sup_{y\in X}\{\langle x^*, y\rangle - f(y)\} \\ &\Leftrightarrow \langle x^*, x\rangle - f(x) \geq \langle x^*, y\rangle - f(y) \forall y \in X \\ &\Leftrightarrow f(y) - f(x) \geq \langle x^*, y - x\rangle \forall y \in X \\ &\Leftrightarrow x^* \in \partial f(x) \end{aligned}$$

□

In dem Spezialfall, dass wir eine konvexe, differenzierbare Funktion in $\mathbb{R}$ haben, kannten wir die Fenchel- Young equality schon vorher. Beachte $f(x)+f^*(x^*) = \langle x^*, x\rangle \Leftrightarrow f(x) = \langle x^*, x\rangle - f^*(x^*)$. Also hat die Gerade mit Steigung x^* und y-Achsenabschnitt $-f^*(x^*)$ einen Schnittpunkt mit dem Graph von f an der Stelle $(x, f(x))$. Wir haben aber $f^*(x^*)$ so konstruiert, dass diese Gerade unterhalb des Graphen unserer Funktion liegt. Also haben wir eine Tangente an dem Graphen von f an der Stelle $(x, f(x))$ mit der Steigung x^*. Wenn wir aber eine konvexe, differenzierbare Funktion haben, ist das äquivalent dazu, dass die Ableitung von f an der Stelle x, x^* ist.

Im nächsten Schritt machen wir die Abschätzung, dass f tatsächlich immer über der Geraden $g(x) = \langle x^*, x\rangle - f^*(x^*)$ liegt.

Satz 3.1.5 (Fenchel-Young inequality).
Für alle $x \in X$ und alle $x^* \in X^*$ gilt:

$$f(x) + f^*(x^*) \geq \langle x^*, x\rangle.$$

Beweis.

$$\begin{aligned} f^*(x^*) &= sup_{x\in X}\{\langle x^*, x\rangle - f(x)\} \geq \langle x^*, x\rangle - f(x) \\ &\Rightarrow f^*(x^*) + f(x) \geq \langle x^*, x\rangle. \end{aligned}$$

□

3.2 Rockafellar Dualität

Definition 3.2.1 (Rockafellar Formulierung).
Sei $F : X \times Y \to \mathbb{R} \cup \{\infty\}$ und $v : Y \to \mathbb{R} \cup \{\infty\}, y \mapsto inf_{x\in X}\{F(x, y)\}$ so, dass v wohldefiniert ist und $\exists (x, y) \in X \times Y, F(x, y) < \infty$.
Wir definieren das primale Problem als

$$p = v(0) = inf_{x\in X}\{F(x, 0)\}$$

und das duale Problem als

$$d = v^{**}(0) = sup_{y^*\in Y^*}\{-F^*(0, -y^*)\}.$$

$p - d$ wird als **Dualitätslücke** bezeichnet.

Zunächst müssen wir beweisen, dass $v^{**}(0) = sup_{y^*\in Y^*}\{-F^*(0, -y^*)\}$.

Beweis. Für $\lambda \in Y^*$ gilt:

$$\begin{aligned} v^*(-\lambda) &= sup_{y\in Y}\{\langle -\lambda, y\rangle - v(y)\} \\ &= sup_{y\in Y}\{\langle -\lambda, y\rangle - inf_{x\in X}\{F(x, y)\}\} \\ &= sup_{y\in Y}\{\langle -\lambda, y\rangle + sup_{x\in X}\{-F(x, y)\}\} \\ &= sup_{(x,y)\in X\times Y}\{\langle -\lambda, y\rangle - F(x, y)\} \\ &= sup_{(x,y)\in X\times Y}\{\langle (0, -\lambda), (x, y)\rangle - F(x, y)\} \\ &= F^*(0, -\lambda), \end{aligned} \tag{1}$$

$$\begin{aligned} d &= v^{**}(0) \\ &= sup_{L\in Y^*}\{\langle 0, L\rangle - v^*(L)\} \\ &\overset{L=-\lambda}{=} sup_{\lambda\in Y^*}\{\langle 0, -\lambda\rangle - v^*(-\lambda)\} \\ &= sup_{\lambda\in Y^*}\{-v^*(-\lambda)\} \\ &\overset{(1)}{=} sup_{\lambda\in Y^*}\{-F^*(0, -\lambda)\}. \end{aligned} \tag{2}$$

□

Satz 3.2.2 (Schwache Dualität).

$$v(0) = p \geq d = v^{**}(0).$$

Insbesondere ist d endlich, falls p endlich ist.

Beweis. Wegen der Fenchel-Young inequality ist

$$F(x,0) + F^*(0,-y^*) \geq \langle (x,0),(0,-y^*)\rangle = 0 \quad \forall x \in X, y \in Y^*.$$

Dadurch ist

$$p = v(0) = inf_{x\in X}\{F(x,0)\} \geq sup_{y^*\in Y^*}\{-F^*(0,-y^*)\} = v^{**}(0) = d.$$

□

Damit haben wir gezeigt, dass d kleiner oder gleich p ist. Gibt es Voraussetzungen unter denen d auch gleich p ist? Das beschreibt die (starke) Rockafellar-Dualität.

Satz 3.2.3 (Rockafellar-Dualität). Die folgenden Aussagen sind äquivalent:

1. Das primale Problem besitzt einen Langrange Multiplikator λ, d.h. $-\lambda \in \partial v(0)$.
2. Es gibt keine Dualitätslücke ($p = d$ und p, d sind endlich) und das duale Problem hat den Löser λ.

Beweis. Die Gleichungen

$$-v^*(-\lambda) = -F^*(0,-\lambda), \tag{1}$$

$$sup_{y^*\in Y^*}\{-F^*(0,-y^*)\} = v^{**}(0) \tag{2}$$

haben wir im Beweis zur Rockafellar-Formulierung bewiesen und werden wir hier wieder benutzen.

"⇒" Sei $-\lambda \in \partial v(0)$.
Wegen der Fenchel-Young equality ist

$$v(0) + v^*(-\lambda) = \langle -\lambda, 0\rangle = 0. \tag{3}$$

Daraus folgt mit der schwachen Dualität, dass

$$\begin{aligned} -v^*(-\lambda) &\overset{(1)}{=} -F^*(0,-\lambda) \\ &\leq sup_{y^*\in Y^*}\{-F^*(0,-y^*)\} \overset{(2)}{=} v^{**}(0) = d \\ &\leq p = v(0) \overset{(3)}{=} -v^*(-\lambda) \Rightarrow p = d. \end{aligned} \tag{4}$$

λ ist ein Löser des dualen Problems, da wegen (4)

$$-v^*(-\lambda) \overset{(1)}{=} -F^*(0,-\lambda) \overset{(4)}{=} sup_{y^*\in Y^*}\{-F^*(0,-y^*)\} = d.$$

"$\Leftarrow$ " Sei $p = v(0) = v^{**}(0) = d$ und λ ein Löser des dualen Problems.
Also ist

$$(5) \qquad -v^*(-\lambda) \overset{(1)}{=} -F^*(0,-\lambda) = sup_{y^* \in Y^*}\{-F^*(0,-y^*)\} \overset{(2)}{=} v^{**}(0).$$

$$\begin{aligned} v(0) + v^*(-\lambda) &= v^{**}(0) + v^*(-\lambda) \\ &\overset{(5)}{=} v^{**}(0) - v^{**}(0) \\ &= 0 \\ &= \langle -\lambda, 0 \rangle \\ &\overset{\text{Fenchel-Young Equality, 3.1.4}}{\Rightarrow} -\lambda \in \partial v(0) \\ &\Rightarrow \lambda \text{ ist ein Lagrange- Multiplikator.} \end{aligned}$$

□

Um zu sehen, dass die starke Dualität nicht immer gilt, kann man sich folgendes Beispiel anschauen:

3.3 Beispiele

Beispiel 3.3.1. Sei $M_y = \{(x_1, x_2) \in \mathbb{R}^2 : \sqrt{x_1^2 + x_2^2} - x_1 \leq y\}$ und
$F : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}, (x_1, x_2, y) \mapsto \begin{cases} |x_2 - 1| & \text{falls } (x_1, x_2) \in M_y \\ \infty & \text{sonst} \end{cases}.$
Dann ist nach Definition $v : \mathbb{R} \to \mathbb{R} \cup \{\infty\}, y \mapsto \begin{cases} inf\{|x_2 - 1| : (x_1, x_2) \in M_y\} & \text{falls } M_y \neq \{\} \\ +\infty & \text{sonst.} \end{cases}$
Behauptung:

$$(1) \qquad v(y) = \begin{cases} 0 & y > 0 \\ 1 & y = 0 \\ +\infty & y < 0 \end{cases}$$

Beweis von (1):

1. $y > 0$:
   Wir wissen, dass $|x_2 - 1| \geq 0 \ \forall x_2 \in \mathbb{R} \Rightarrow v(y) \geq 0$.
   Gleichzeitig gilt, dass $|x_2 - 1| = 0 \Leftrightarrow x_2 = 1$.
   Damit ist $v(y) = 0$, wenn wir zeigen können, dass ein $x_1 \in \mathbb{R}$ existiert, mit $(x_1, 1) \in M_y$. Dazu:

$$\begin{aligned} \sqrt{x_1^2 + 1} - x_1 \leq y \Leftrightarrow \sqrt{x_1^2 + 1} &\leq y + x_1 \\ &\Leftrightarrow x_1^2 + 1 \leq (y + x_1)^2 \ \ \& \ \ 0 \leq y + x_1 \\ &\Leftrightarrow x_1^2 + 1 \leq y^2 + 2yx_1 + x_1^2 \ \ \& \ \ -y \leq x_1 \\ &\Leftrightarrow \frac{1 - y^2}{2y} \leq x_1 \ \ \& \ \ -y \leq x_1. \end{aligned}$$

   Setzt man also $x_1 := \max\{\frac{1-y^2}{2y}, -y\} = \frac{1-y^2}{2y}$ und $x_2 := 1$, so ist $(x_1, x_2) \in M_y$ und $|x_2 - 1| = 0$.

2. $y = 0$:

$$\begin{aligned}\sqrt{x_1^2 + x_2^2} - x_1 \le 0 \Leftrightarrow \sqrt{x_1^2 + x_2^2} &\le x_1 \\ \Leftrightarrow x_1^2 + x_2^2 &\le (x_1)^2 \;\;\&\;\; 0 \le x_1 \\ \Leftrightarrow x_2^2 &\le 0 \;\;\&\;\; 0 \le x_1 \\ \Leftrightarrow x_2 &= 0 \;\;\&\;\; 0 \le x_1\end{aligned}$$

$\Rightarrow M_y \neq \{\}$ & $|x_2 - 1| = 1 \ \forall (x_1, x_2) \in M_y$.

3. $y < 0$:

$$\begin{aligned}&\sqrt{x_1^2 + x_2^2} - x_1 \le y \Leftrightarrow \sqrt{x_1^2 + x_2^2} \le y + x_1 \\ &\Rightarrow x_1^2 + x_2^2 \le (y + x_1)^2 \;\;\&\;\; 0 \le y + x_1 \\ &\Rightarrow x_1^2 + x_2^2 \le y^2 + 2yx_1 + x_1^2 \;\;\&\;\; -y \le x_1 \\ &\Rightarrow x_2^2 \le |y|^2 - 2|y|x_1 \;\;\&\;\; -y = |y| \le x_1 \\ &\Rightarrow 0 \le x_2^2 \le |y|^2 - 2|y|^2 = -2y^2 < 0.\end{aligned}$$

Das kann nicht sein. Also ist $M_y = \{\} \Rightarrow v(y) = \infty$.

Damit ist (1) gezeigt. $\square$

$$\begin{aligned}v^*(y^*) &= sup_{y\in\mathbb{R}}\{\langle y^*, y\rangle - v(y)\} \\ &\overset{(1)}{=} sup_{y\ge 0}\{y^* y - v(y)\} \\ &\overset{(1)}{=} max\{y^* \cdot 0 - 1, sup_{y>0}\{y^* y\}\} \\ &= \begin{cases} 0 & y^* \le 0 \\ \infty & y^* > 0 \end{cases}\end{aligned} \tag{2}$$

und

$$\begin{aligned}v^{**}(y) &= sup_{y^*\in\mathbb{R}}\{\langle y, y^*\rangle - v^*(y^*)\} \\ &\overset{(2)}{=} sup_{y^*\le 0}\{y^* y - v(y^*)\} \\ &\overset{(2)}{=} sup_{y\le 0}\{y^* y\} \\ &= \begin{cases} 0 & y \ge 0 \\ \infty & y < 0 \end{cases}.\end{aligned} \tag{3}$$

Damit ist die Dualitätslücke $p - d = v(0) - v^{**}(0) = 1 - 0 = 1$. Dies widerspricht jedoch nicht der Rockafellar-Dualität, da das primale Problem keinen Lagrange-Multiplikator besitzt.
Denn angenommen x^* wäre einer, also $-x^* \in \partial v(0) = \{y^* \in \mathbb{R} : v(y) - v(0) \ge \langle y^*, y - 0\rangle \ \forall y \in \mathbb{R}\}$, dann müsste insbesondere für alle $y > 0$:

$$\begin{aligned}&v(y) - v(0) = 0 - 1 = -1 \ge \langle -x^*, y - 0\rangle = -x^*(y - 0) = -x^* y \\ &\Leftrightarrow \frac{1}{y} \le x^*\end{aligned}$$

sein. Das kann jedoch nicht sein, da $lim_{y\downarrow 0}\frac{1}{y} = \infty$.

4 Fenchel-Dualität

4.1 Fenchel Formulierung

In der Fenchel-Formulierung beschränken wir uns nun darauf, dass F als $f(x) + g(Ax + y)$ geschrieben werden kann.

Definition 4.1.1 (Fenchel Formulierung).
Sei

$$F : X \times Y \to \mathbb{R} \cup \{\infty\}, (x, y) \mapsto f(x) + g(Ax + y),$$

wobei
$A : X \to Y$ ein linearer Operator ist und $f : X \to \mathbb{R} \cup \{\infty\}$, $g : Y \to \mathbb{R} \cup \{\infty\}$ und $\exists (x, y) \in X \times Y, F(x, y) < \infty$.

$$v : Y \to \mathbb{R} \cup \{\infty\}, y \mapsto inf_{x \in X}\{F(x, y)\}$$

sei so, dass v wohldefiniert ist.
Wir definieren das primale Problem als

$$p = v(0) = inf_{x \in X}\{F(x, 0)\} = inf_{x \in X}\{f(x) + g(Ax)\}$$

und das duale Problem als

$$d = v^{**}(0) \overset{\text{Defintion 3.2.1}}{=} sup_{y^* \in Y^*}\{-F^*(0, -y^*)\} = sup_{y^* \in Y^*}\{-f^*(A^* y^*) - g^*(-y^*)\}.$$

Zunächst müssen wir beweisen, dass $sup_{y^* \in Y^*}\{-F^*(0, -y^*)\} = sup_{y^* \in Y^*}\{-f^*(A^* y^*) - g^*(-y^*)\}$.

Beweis.

$$\begin{aligned} F^*(0, -y^*) &= sup_{(x,y) \in X \times Y}\{\langle -y^*, y\rangle - f(x) - g(Ax + y)\} \\ &\overset{u=Ax+y}{=} sup_{(x,u) \in X \times Y}\{\langle -y^*, u - Ax\rangle - f(x) - g(u)\} \\ &= sup_{x \in X}\{\langle y^*, Ax\rangle - f(x)\} + sup_{u \in Y}\{\langle -y^*, u\rangle - g(u)\} \\ &= f^*(A^* y^*) + g^*(-y^*). \end{aligned}$$

□

Die schwache Dualität, dass $p \geq d$ ist, folgt direkt aus der schwachen Rockafellar- Dualität (Satz 3.2.2). Die nächste Frage ist, ob wir die Bedingung in der starken Rockafellar-Dualität, dass p einen Langrange-Multiplikator hat in unserem spezielleren Fall umschreiben können. Dazu werden wir noch spezieller.

4.2 Konvexe, unterhalbstetige Funktionen

Lemma 4.2.1. Falls f und g beide unterhalb stetig und konvex sind, so ist es auch v.

Beweis.

1. Konvex: Sei $\lambda \in (0,1)$ und $(x_1,y_1),(x_2,y_2) \in X \times Y$. Dann ist

$$
\begin{aligned}
&F(\lambda(x_1,y_1)+(1-\lambda)(x_2,y_2)) \\
&= f(\lambda x_1+(1-\lambda)x_2)+g(A(\lambda x_1+(1-\lambda)x_2)+\lambda y_1+(1-\lambda)y_2) \\
&\le \lambda f(x_1)+(1-\lambda)f(x_2)+g(\lambda Ax_1+(1-\lambda)Ax_2+\lambda y_1+(1-\lambda)y_2) \\
&\le \lambda f(x_1)+(1-\lambda)f(x_2)+\lambda g(Ax_1+y_1)+(1-\lambda)g(Ax_2+y_2) \\
&= \lambda F(x_1,y_1)+(1-\lambda)F(x_2,y_2).
\end{aligned}
$$

   Damit folgt leicht, dass auch v konvex ist.

2. Unterhalbstetig: Sei $\lim\limits_{n\to\infty}(x_n,y_n)=(x,y)$ Da A stetig ist, ist

$$\lim_{n\to\infty} Ax_n = Ax.$$

   Da f und g unterhalbstetig sind, ist:

$$\liminf_{n\to\infty} F(x_n,y_n) = \liminf_{n\to\infty} f(x_n)+g(Ax_n+y_n) \ge f(x)+g(Ax+y)=F(x,y). \tag{1}$$

   Sei $\epsilon>0$ und $n:=n(\epsilon)\in\mathbb{N}$ nun so groß, dass $inf_{i\ge n}\{F(x,y_i)\} \ge F(x)-\epsilon$. Nach (1) wissen wir, dass dieses existiert.

$$
\begin{aligned}
inf_{i\ge n}\{v(y_i)\} &= inf_{i\ge n}\{inf_{x\in X}F(x,y_i)\} \\
&= inf_{i\ge n, x\in X}\{F(x,y_i)\} \\
&= inf_{x\in X}\{inf_{i\ge n}F(x,y_i)\} \\
&\ge inf_{x\in X}\{F(x,y)-\epsilon\} \\
&= inf_{x\in X}\{F(x,y)\}-\epsilon \\
&= v(y)-\epsilon.
\end{aligned}
\tag{2}
$$

   Damit gilt aber auch schon, dass

$$\liminf_{n\to\infty} v(y_n) = \lim_{n\to\infty} inf_{i\ge n}\{v(y_i)\} \overset{(2)}{\ge} v(y).$$

   Also ist v unterhalb stetig.

□

Bemerkung 4.2.2. Ein Resultat aus [2] ist:
Falls v unterhalbstetig und konvex ist, ist eine hinreichende Bedingung für die Existenz des Lagrange Multiplikators an der Stelle x, dass

$$x \in ri(dom(v)).$$

ri meint relative interior.
Diese Arten von Bedingungem werdem auch CQ = Constraint Qualification genannt.

Also wissen wir bereits, dass wenn $0 \in ri(dom(v))$ ist, wir eine starke Dualität besitzen.

Die Frage ist nun, ob wir die Domaine von v auch umschreiben können, sodass wir schon mit dem gegebenden f und g prüfen können, ob wir eine starke Dualität haben.

Lemma 4.2.3. $dom(v) = dom(g) - A(dom(f))$.

Beweis. "$\subseteq$"

$$\begin{aligned}
&\text{Sei } y \in dom(v) \Rightarrow \\
&\exists x \in X \text{ mit } f(x) + g(Ax + y) < \infty \\
&\Rightarrow f(x) < \infty \wedge g(Ax + y) < \infty \\
&\Rightarrow x \in dom(f) \wedge Ax + y \in dom(g) \\
&\Rightarrow y \in dom(g) - A(dom(f)).
\end{aligned}$$

"$\supseteq$"

$$\begin{aligned}
&\text{Sei } y \in dom(g) - A(dom(f)) \\
&\Rightarrow y = z - Ax, z \in dom(g) \wedge x \in dom(f) \\
&\Rightarrow x \in dom(f) \wedge y + Ax \in dom(g) \\
&\Rightarrow f(x) + g(Ax + y) < \infty \\
&\Rightarrow v(y) < \infty \\
&\Rightarrow y \in dom(v).
\end{aligned}$$

□

Also haben wir eine Bedingung gefunden für die starke Dualität, für welche wir v nicht explizit ausrechnen müssen.

Satz 4.2.4 (Fenchel-Dualität). Falls f und g unterhalb stetige, konvexe Funktionen sind und die constraint qualification für 0 erfüllt wird:

$$0 \in ri[dom(v)] = ri[dom(g) - A(dom(f))],$$

dann gilt in der Situation von der Definition 4.1.1 starke Dualität, also $p = d$.

Die Frage ist nun, ob wir in manchen Fällen die Überprüfung wegfallen lassen können, dass 0 im relativen Inneren der Domain liegt. Dazu schauen wir uns polyhedrische Funktionen an:

4.3 Polyhedrische Funktionen

Zunächst definieren wir uns, was polyhedrische Mengen und Funktionen sind.

Definition 4.3.1 (Polyhedrische Menge).
Eine polyhedrische Menge P ist eine Menge, welche dargestellt werden kann. als

$$P = \cap_{i \in I}\{x \in X : \langle a_i, x\rangle \leq b_i\},$$

wobei $a_i \in X^*, b_i \in \mathbb{R}$ und I endlich ist.

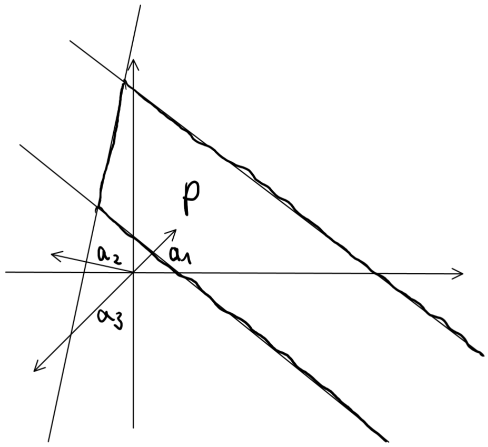

Abbildung 4: Polyhedrische Menge

Definition 4.3.2 (Polyhedrische Funktionen).
Eine polyhedrische Funktion $f : X \to \mathbb{R} \cup \{\infty\}$ ist eine Funktion, welche dargestellt werden kann als

$$f = max_{i \in I} h_i + \iota_P,$$

wobei die h_i affine Abbildungen sind, I endlich und P eine polyhedrische Menge ist.
Hierbei bezeichnet ι_P die Indikatorfunktion von P, welche definiert ist als:

$$\iota_P : X \to \mathbb{R} \cup \{\infty\}, x \mapsto \begin{cases} 0 & x \in P \\ \infty & \text{sonst} \end{cases}$$

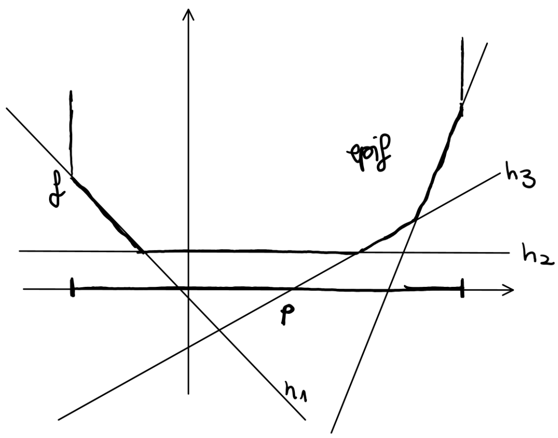

Abbildung 5: Polyhedrische Funktion

Polyhedrische Mengen und Funktionen hängen wie folgt zusammen:

Bemerkung 4.3.3. Ein Resultat aus [2] ist:
f ist eine polyhedrische Funktion genau dann wenn $epi(f)$ eine polyhedrische Menge ist.

Polyhedrische Funktionen haben außerdem die schöne Eigenschaft, dass wir die Existenz eines Lagrange Multiplikators an der Stelle x leichter nachweisen können.

Bemerkung 4.3.4. Ein Resultat aus [2] ist:
Falls v polyhedrisch und konvex ist, ist eine hinreichende Bedingung für die Existenz des Lagrange Multiplikators an der Stelle x, dass

$$x \in dom(v)$$

Um damit eine schöne Formulierung der starken Rockafellar-Dualität zu erhalten, wollen wir nun zeigen, dass, wenn f und g polyhedrisch sind und X, Y endlich dimensional, v ebenfalls polyhedrisch ist.

Das machen wir in zwei Schritten.

Lemma 4.3.5. Falls f und g beide polyhedrisch sind, so ist es auch das in Definition 4.1.1 eingeführte $F(x, y) = f(x) + g(Ax + y)$. Dabei ist $f : X \to \mathbb{R} \cup \{\infty\}$ und $g : Y \to \mathbb{R} \cup \{\infty\}$. $A : X \to Y$ ist eine lineare Abbildung.

Beweis. Falls f und g beide polyhedrisch sind, so ist $f = max_{i \in I} f_i + \iota_P$ und $g = max_{i \in I'} g_i + \iota_{P'}$, mit f_i ist affin $\forall i \in I$ und g_i ist affin $\forall i \in I'$. Außerdem sind I, I' endlich und P, P' polyhedrische Mengen.

1. Wenn f_i affin ist, ist
$$f_i(x) = \langle s_i, x\rangle + a_i,$$
wobei $s_i \in X^*$ und $a_i \in \mathbb{R}$. Also ist auch
$$f_i' : X \times Y \to \mathbb{R}, (x, y) \mapsto \langle (s_i, 0), (x, y)\rangle + a_i = f_i(x)$$
affin.

2. Wenn P eine polyhedrische Menge in X ist, so ist
$$P = \cap_{i \in D}\{x \in X : \langle c_i, x\rangle \leq b_i\},$$
$c_i \in X^*, b_i \in \mathbb{R}\ \forall i \in D$ und D endlich. Also ist
$$\begin{aligned} P_2 &= P \times Y \\ &= \cap_{i \in D}\{(x, y) \in X \times Y : \langle c_i, x\rangle \leq b_i\} \\ &= \cap_{i \in D}\{(x, y) \in X \times Y : \langle (c_i, 0), (x, y)\rangle \leq b_i\} \end{aligned}$$
eine polyhedrische Menge in $X \times Y$.

3. Wenn $g_i : Y \to \mathbb{R}$ affin ist, ist
$$g_i(y) = \langle t_i, y\rangle + d_i,$$
wobei $t_i \in Y^*$ und $d_i \in \mathbb{R}$. Also gilt, wenn $x \in X$, $y \in Y$ und A ein linearer Operator von X nach Y ist:
$$\begin{aligned} g_i(Ax + b) &= \langle t_i, Ax + y\rangle + d_i \\ &= \langle t_i, Ax\rangle + \langle t_i, y\rangle + d_i \\ &= \langle A^* t_i, x\rangle + \langle t_i, y\rangle + d_i \\ &= \langle (A^* t_i, t_i), (x, y)\rangle + d_i. \end{aligned}$$

Also ist auch

$$g_i^{'} : X \times Y \to \mathbb{R}, (x,y) \mapsto \langle (A^* t_i, t_i), (x,y) \rangle + d_i = g_i(Ax+y)$$

affin.

4. Wenn P' eine polyhedrische Menge in Y ist, so ist

$$P' = \cap_{i \in D'} \{ y \in Y : \langle f_i, y \rangle \leq e_i \},$$

wobei $f_i \in Y^*, e_i \in \mathbb{R} \forall i \in D'$ und D' endlich. Dann ist auch

$$\begin{aligned} P_2' &= \{(x,y) \in X \times Y | Ax + y \in P'\} \\ &= \cap_{i \in D'} \{(x,y) \in X \times Y : \langle f_i, Ax+y \rangle \leq e_i\} \\ &= \cap_{i \in D'} \{(x,y) \in X \times Y : \langle f_i, Ax \rangle + \langle f_i, y \rangle \leq e_i\} \\ &= \cap_{i \in D'} \{(x,y) \in X \times Y : \langle A^* f_i, x \rangle + \langle f_i, y \rangle \leq e_i\} \\ &= \cap_{i \in D'} \{(x,y) \in X \times Y : \langle (A^* f_i, f_i), (x,y) \rangle \leq e_i\} \end{aligned}$$

eine polyhedrische Menge in $X \times Y$.

5. Aus den vorherigen Schritten folgt, dass

$$\begin{aligned} F(x,y) &= f(x) + g(Ax+y) \\ &= max_{i \in I} f_i(x) + \iota_P(x) + max_{i \in I'} g_i(Ax+y) + \iota_{P'}(Ax+y) \\ &= max_{i \in I} f_i'(x,y) + \iota_{P_2}(x,y) + max_{i \in I'} g_i'(x,y) + \iota_{P_2'}(x,y) \\ &= max_{(i,i') \in I \times I'} (f_i'(x,y) + g_{i'}'(x,y)) + \iota_{P_2 \cap P_2'}(x,y). \end{aligned}$$

Die Summe von affinen Funktionen ist affin. Da I und I' endlich sind, ist damit auch $I \times I'$ endlich. Außerdem ist der Schnitt polyhedrischer Mengen polyhedrisch. Insgesamt ist F damit auch eine polyhedrische Funktion.

□

Bemerkung 4.3.6. Ein Resultat, welches in [1], Seite 24, genutzt wird, ist:
Projektionen von polyhedrischen Mengen sind in endlich-dimensionalen
$\mathbb{R}$-Vektorräumen ebenfalls polyhedrisch.

Lemma 4.3.7. Falls F polyhedrisch, $X \times Y$ endlich dimensional und das in Definition 4.1.1 eingeführte $v : Y \to \mathbb{R} \cup \{\infty\}, y \mapsto inf_{x' \in X} F(x', y)$ wohldefiniert ist, so ist auch v polyhedrisch.

Beweis. Angenommen F ist polyhedrisch. Dann wissen wir, dass $epi(F)$ eine polyhedrische Menge ist.
Behauptung: $epi(v) = \{(y,z) \in Y \times \mathbb{R} : \exists x \in X, (x,y,z) \in epi(F)\} := P$.

- "$\subseteq$ " Sei $(y,z) \in epi(v)$. Dann ist $z \geq v(y) = inf_{x' \in X} F(x', y)$. F ist polyhedrisch. Dadurch wissen, dass auch $x \mapsto F(x,y)$ polyhedrisch ist. Also ist

$$g : X \to \mathbb{R} \cup \{\infty\}, x \mapsto F(x,y) = max_{i \in I} f_i + \iota_D,$$

mit f_i affin $\forall i \in I$, I endlich und D polyhedrisch (also auch abgeschlossen). Da die f_i affin sind, also stetig, ist jeweils $g(x) = f_i(x)$ auf abgeschlossenen Intervallen J_i. Weil die f_i affin sind, sind sie in der jeweiligen x_k Richtung entweder konstant oder streng monoton mit $\lim\limits_{x_k \to \infty} f_i(x_1, \ldots, x_{dim(X)}) \in \{-\infty, \infty\}$ und $\lim\limits_{x_k \to -\infty} f_i(x_1, \ldots, x_{dim(X)}) \in \{-\infty, \infty\}, k = 1 \ldots dim(x)$, $i \in I$. Dabei ist dies unabhängig davon, welche Werte $x_i,\ i \in \{1 \ldots dim(x)\} \setminus k$ haben.
Daher gibt es drei Fälle:

1. Die f_i nehmen auf J_i ihr Minimum m_i (auch) am Rand an.
2. Die f_i besitzen auf J_i kein Infimum.
3. $J_i = \{\}$.

Weil v wohldefiniert ist, kann der zweite Fall ausgeschlossen werden, da es sonst auch kein $inf_{x' \in X} F(x^{'}, y) \in \mathbb{R} \cup \{\infty\}$ gäbe. Außerdem wissen wir dadurch, dass $D \neq \{\}$, also dass es ein m_i geben muss. Das Minimum der m_i ist das von g. Dieses existiert, da I endlich ist. Also existiert ein $x \in X$ mit $inf_{x' \in X} F(x^{'}, y) = F(x, y) \Rightarrow (x, y, z) \in epi(F) \Rightarrow (y, z) \in P$.

- "$\supseteq$ " Sei $(y, z) \in P \Rightarrow \exists x \in X, (x, y, z) \in epi(F) \Rightarrow z \geq F(x, y) \geq inf_{x' \in X} F(x^{'}, y) = v(y)$.

Mit der Bemerkung 4.3.6, ist also $epi(v)$ eine polyhedrische Menge. Mit der Bemerkung 4.3.3 folgt, dass dann auch v eine polyhedrische Funktion ist.

□

Damit kann man die Bedingung der Rockafellar-Dualität, dass v einen Langrange-Multiplikator an der Stelle 0 besitzt, umschreiben.

Bemerkung 4.3.8. Falls X, Y endlich dimensional sind, f und g polyhedrische, konvexe Funktionen sind und

$$0 \in dom(v) = dom(g) - A(dom(f)),$$

dann ist $p = d$.

4.4 "Decoupling"

Im nächsten Schritt sehen wir uns an, wie man die Löser des primalen und dualen Problems in Abhängigkeit bringen kann.

Satz 4.4.1 ("Decoupling"/Entkuppeln). Falls für p, d der Definition 4.1.1, $p = d$ ist und x ein Löser des primalen Problems ist und y ein Löser des dualen Problems ist, dann gilt:

$$f(x) + f^*(A^*y) - \langle y, Ax \rangle = 0$$

und

$$g(Ax) + g^*(-y) + \langle y, Ax \rangle = 0.$$

Beweis. Sei $p = d$, x ein Löser des primalen Problems und y ein Löser des dualen Problems. Dann gilt:

$$\begin{aligned} p &= v(0) \\ &= inf_{x' \in X}\{f(x') + g(Ax')\} \\ &\overset{\text{x Löser von p}}{=} f(x) + g(Ax) \\ &\overset{p=d}{=} d \\ &= sup_{y^* \in Y^*}\{-f^*(A^*y^*) - g^*(-y^*)\} \\ &\overset{\text{y Löser von d}}{=} -f^*(A^*y) - g^*(-y). \end{aligned}$$

Also:

$$\begin{aligned} &f(x) + g(Ax) + f^*(A^*y) + g^*(-y) = 0 \\ \overset{\text{Künstliche Null}}{\Rightarrow} & [f(x) + f^*(A^*y) - \langle y, Ax\rangle] + [g(Ax) + g^*(-y) + \langle y, Ax\rangle] = 0. \end{aligned}$$

Die Fenchel- Young inequality, Satz 3.1.5, besagt, dass

$$f(x) + f^*(A^*y) \geq \langle A^*y, x\rangle = \langle y, Ax\rangle$$

und

$$g(Ax) + g^*(-y) \geq \langle -y, Ax\rangle = -\langle y, Ax\rangle.$$

Damit sind aber die Terme in den eckigen Klammern beide nicht negativ. Daher müssen schon beide gleich 0 sein. □

4.5 Beispiele

Beispiel 4.5.1 (Optimierung der Entropie). Wir beschäftigen uns mit folgendem allgemeinen Fall:

$$min_{x\in X}\{f(x) : Ax = b \in \mathbb{R}^N\},$$

wobei X ein $\mathbb{R}$-Banachraum ist. Außerdem soll f unterhalbstetig und konvex sein, $A : X \to \mathbb{R}^N$ eine lineare Abbildung und $b \in A(ri(dom(f)))$.

Dieser allgemeine Fall tritt beispielsweise auf, wenn man die Entropie maximieren möchte und lineare Bedingungen an sie hat. Dabei ist f über einem Signalraum definiert (ein Raum, welcher eine Folge von Signalen mit einer Norm darstellt, also $\mathbb{R}^n, l_p, L_p$) und stellt die negative Entropie dar (die Entropie ist konkav, damit ist das Negative der Entropie konvex). Das Maximum der Entropie ist das Minimum von f.

Auf Grund der Allgemeinheit, kann es in vielen verschiedenen Anwendungen angewandt werden. Zum Beispiel bei der Konstruktion von CT-scans von Körpern. Dort gibt es viele 2D-Bilder und das 3D Bild muss auf Grund dessen erstellt werden. $X \in \mathbb{R}^N$ ist hierbei eine Diskretisierung des Bildes.

Um unser Problem zu lösen, setzen wir das f der Definition 4.1.1 auf die Funktion f, welche wir minimieren wollen und g, wie folgt:

$$g(y) = \iota_b(y) = \begin{cases} 0 & y = b \\ \infty & y \neq b \end{cases}.$$

Damit sind die Funktionen der Definition 4.1.1 wie folgt gegeben:

$$F : X \times \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}, (x, y) \mapsto f(x) + g(Ax + y) = \begin{cases} f(x) & Ax + y = b \\ \infty & \text{sonst} \end{cases},$$

$$v : \mathbb{R}^N \to \mathbb{R}\cup\{\infty\}, y \mapsto inf_{x\in X}\{f(x)+g(Ax+y)\} = \begin{cases} inf_{x\in X}\{f(x) : Ax = b - y\} & \exists x \in X \text{ mit } Ax = b - y \\ \infty & \text{sonst} \end{cases}$$

und das primale Problem

$$p = v(0) = \begin{cases} inf_{x\in X}\{f(x) : Ax = b\} & \exists x \in X \text{ mit } Ax = b \\ \infty & \text{sonst} \end{cases}.$$

Da nach unseren Voraussetzungen $b \in A(ri(dom(f)))$, wissen wir, dass es insbesondere ein $x \in X$ existiert, mit $Ax = b$. Damit ist p unsere allgemeine Problemstellung.

Da f und g konvex und unterhalbstetig und $0 \in ri[dom(g) - A(dom(f))]$, da

$$0 \in ri[dom(g) - A(dom(f))] = ri[b - A(dom(f))] \Leftrightarrow b \in A(ri(dom(f)))$$

können wir Satz 4.2.4 anwenden. Durch ihn wissen wir, dass es eine starke Dualität zwischen dem primalen und dualen Problem der Definition 4.1.1 gibt.

Wir berechnen nun das duale Problem:

Zunächst berechnen wir g^*:

$$g^*(y^*) = sup_{y \in X}\{\langle y^*, y\rangle - g(y)\} = \langle y^*, b\rangle,$$

da $g(y) = \infty$ ist, wenn $y \neq b$.

Unser duales Problem ist nach Definition 4.1.1:

$$\text{(NR 1)} \qquad \begin{aligned} d &= sup_{y^* \in \mathbb{R}^N}\{-f^*(A^*y^*) - g^*(-y^*)\} \\ &= sup_{y^* \in \mathbb{R}^N}\{\langle y^*, b\rangle - f^*(A^*y^*)\} \\ &\overset{\text{Definition 3.1.1}}{=} (f^* \circ A^*)^*(b). \end{aligned}$$

Hier fließt mehrmals ein, dass $\mathbb{R}^N$ zu seinem Dualraum isomorph ist, und somit als dieser angesehen werden kann. Falls $N < dimX$, ist das duale Problem normalerweise deutlich einfacher zu lösen als das primale Problem.

Beispiel 4.5.2 (Erster Spezialfall). Wir setzen nun im Beispiel 4.5.1 $X = \mathbb{R}^M$ und

$$f(x) := \frac{1}{2}x^\intercal \Sigma x,$$

wobei $\Sigma \in \mathbb{R}^{MxM}$ symmetrisch und positiv definit, $A\Sigma^{-1}A^\intercal \in \mathbb{R}^{NxN}$ invertierbar und $b \in Range(A) \subset \mathbb{R}^N$ sein soll.

Dann sind unsere Bedingungen, dass f unterhalbstetig und konvex ist und

$$b \in A[ri(dom(f))] = A(\mathbb{R}^M) = Range(A)$$

erfüllt.

Wir haben normalerweise Probleme eine quadratische Funktion mit linearen Nebenbedingungen zu optimieren, in diesem Fall kann man es jedoch explizit machen, da

$$\text{(1)} \qquad f^*(y) = \frac{1}{2}y^\intercal \Sigma^{-1} y,\ y \in X^* = \mathbb{R}^M.$$

Beweis von (1):

$$\begin{aligned} f^*(y) &= sup_{x \in \mathbb{R}^M}\{\langle y, x\rangle - f(x)\} \\ &= sup_{x \in \mathbb{R}^M}\{\langle y, x\rangle - \frac{1}{2}x^\intercal \Sigma x\} \\ &= sup_{x \in \mathbb{R}^M}\{\langle y, x\rangle - \frac{1}{2}\langle x, \Sigma x\rangle\}. \end{aligned}$$

Da f eine quadratische Funktion ist und damit konvex, ist $-f$ konkav. $x \mapsto \langle y, x\rangle$ ist linear also auch konvav. Damit ist auch die Summe konkav. Falls die Ableitung von $x \mapsto \langle y, x\rangle - f(x)$ also eine Nullstelle besitzt, ist das hinreichend um zu wissen, dass dort ein Maximum ist.

$$\begin{aligned} 0 &= \frac{\partial}{\partial x}(\langle y, x\rangle - \frac{1}{2}\langle x, \Sigma x\rangle) \\ &= y^\intercal - x^\intercal \frac{1}{2}(\Sigma + \Sigma^\intercal) \\ &\overset{\Sigma=\Sigma^\intercal}{=} y^\intercal - x^\intercal \Sigma^\intercal. \end{aligned}$$

Also ist am Maximum

$$y = \Sigma x \text{ und damit } x = \Sigma^{-1} y. \tag{NR 2}$$

Im nächsten Schritt setzten wir dies ein und erhalten:

$$\begin{aligned} f^*(y) &= \langle y, \Sigma^{-1} y \rangle - \frac{1}{2} \langle \Sigma^{-1} y, \Sigma \Sigma^{-1} y \rangle \\ &= \frac{1}{2} y^\intercal \Sigma^{-1} y. \end{aligned}$$

□

Es folgt

$$f^* \circ A^\intercal(y) = \frac{1}{2} y^\intercal A \Sigma^{-1} A^\intercal y, \; y \in \mathbb{R}^N.$$

Wenn Σ symmetrisch ist, ist es auch Σ^{-1} und damit auch $A\Sigma^{-1}A^\intercal$. Da nach Annahme $A\Sigma^{-1}A^\intercal$ invertierbar ist, können wir den Beweis von (1) erneuert mit $f^* \circ A^\intercal$ durchführen um

$$d = (f^* \circ A^\intercal)^*(b) = sup_{y \in \mathbb{R}^N} \{ \langle b, y \rangle - f^* \circ A^\intercal(y) \},$$

(vgl. (NR 1)), herauszubekommen und den Löser dieses Maximierungsproblems.

Wie in (NR 2) erhält man:

$$y' = (A \Sigma^{-1} A^\intercal)^{-1} b.$$

Falls es einen Löser des primalen Problem x' gibt, wissen wir durch das Decoupeling-Theorem, Satz 4.4, dass für x' und y' gilt:

$$f(x') + f^*(A^\intercal y') = \langle A^\intercal y', x' \rangle.$$

Mit der Fenchel-Young Equality, Satz 3.1.4, ist

$$A^\intercal y' \in \partial f(x') = \{ \nabla f(x') \} = \{ \Sigma x' \},$$

da f konvex und differenzierbar ist. Also folgt

$$x' = \Sigma^{-1} A^\intercal y' = \Sigma^{-1} A^\intercal (A \Sigma^{-1} A^\intercal)^{-1} b.$$

Beispiel 4.5.3 (**Zweiter Spezialfall: Boltzmann-Shannon Entropie**). Wir setzten hier im Beispiel 4.5.1 $X = \mathbb{R}^M$ und

$$f = f' + \langle c, x \rangle$$

mit

$$f'(x) := \sum_{n=1}^{M} p(x_n),$$

wobei

$$p(t) := \begin{cases} t \cdot \ln(t) - t & t > 0 \\ 0 & t = 0 \\ \infty & t < 0 \end{cases}.$$

f' stellt die Bolzmann- Shannon Entropie dar. Dabei sei $c \in \mathbb{R}^M$. Damit wir die Voraussetzungen von dem Beispiel 4.5.1 erfüllen, benötigen wir

$$b \in A[ri(dom(f))] \Leftrightarrow \exists x \in (\mathbb{R}_+^*)^M, Ax = b \tag{1}$$

(2) f ist unterhalb-stetig und konvex.

Dabei ist

$$\mathbb{R}^*_+ := \{x \in \mathbb{R} : x \geq 0\}$$

und

$$\mathbb{R}_+ := \{x \in \mathbb{R} : x > 0\}.$$

Das Erste nehmen wir an, das Zweite zeigen wir:
Sei $x \in (\mathbb{R}^*_+)^M$.

$$\frac{d}{dx_n}(f(x)) = ln(x_n) + c_n.$$

Also ist

$$\frac{d}{dx_n dx_m} f(x) = \begin{cases} \frac{1}{x_n} & n = m \\ 0 & \text{sonst} \end{cases}.$$

Daher ist die Hessematrix von f $\forall x \in (\mathbb{R}^*_+)^M$ positiv definit. Also ist f auf $(\mathbb{R}^*_+)^M$ streng konvex.

Wir zeigen nun, dass f auf $dom(f) = (\mathbb{R}_+)^M$ stetig ist. Auf $(\mathbb{R}_+)^M$ ist f stetig als Verknüpfung(Summe, Produkt) stetiger Funktionen. Außerdem ist

$$lim_{x_n \downarrow 0} p(x_n) + c_n \cdot x_n = 0 = p(0) + c_n \cdot 0,\ n = 1 \dots M.$$

Daraus folgt die Behauptung. Insgesamt ist dann f auf ganz $dom(f) = \mathbb{R}^M_+$ konvex und unterhalbstetig. $\square$

Falls es einen Löser des primalen Problems

$$min_{x \in \mathbb{R}^N} \{f(x) + \langle c, x \rangle : Ax = b\}$$

gibt, wollen wir diesen nun in Zusammenhang zu dem Löser des dualen Problems aus Definition 4.1.1 stellen.

Angenommen dieser Löser x' läge am Rand von $\mathbb{R}^M_+$. Dann wäre er o.B.d.A. in der 1 bis i-ten Komponente 0. Gleichzeitig wissen wir, dass ein weiteres $x^* \in ri(\mathbb{R}^M_+)$ existiert mit $Ax^* = b$, nach (1). Also ist auch für

$$x_\lambda := (1 - \lambda)x' + \lambda x^*,$$

$$Ax_\lambda = A((1 - \lambda)x' + \lambda x^*) = (1 - \lambda)A(x') + \lambda A(x^*) = b,\ \lambda \in (0, 1).$$

Dabei liegt x_λ in $ri(\mathbb{R}^M_+)$.

Der 1 bis i-te Eintrag von $v := x^* - x'$ ist positiv, da $x^* \in (\mathbb{R}^*_+)^M$. Damit ist

$$lim_{\lambda \downarrow 0} D_v f(x_\lambda) = lim_{\lambda \downarrow 0} ln(x_{\lambda_1}) \cdot v_1 + c_1 \cdot v_1 + ... + ln(x_{\lambda_1}) \cdot v_i + c_i \cdot v_i + .. + ln(x_{\lambda_M}) \cdot v_M + c_M \cdot v_M = -\infty.$$

Daher gibt es mit dem Mittelwertsatz ein $\lambda > 0$ und $f(x_\lambda) < f(x')$. Das ist ein Widerspruch zur Annahme. Also wissen wir, dass falls wir einen Löser x' des primalen Problems besitzen, dieser in $ri(\mathbb{R}^M_+)$ liegt.

Das duale Problem besitzt aufgrund der starken Dualität einen Löser ϕ.

Mit dem Decoupeling-Theorem, Satz 4.4, wissen wir, dass:

$$f(x') + f^*(A^\intercal \phi) = \langle \phi, Ax' \rangle = \langle A^\intercal \phi, x' \rangle.$$

Daher ist mit der Fenchel-Young equality, Satz 3.1.4,

$$A^\intercal \phi \in \partial f(x') = \{\nabla f(x')\} = \{(ln(x'_1) + c_1, .., ln(x'_n) + c_n)^\intercal\},$$

da f auf $(\mathbb{R}^*_+)^M$ konvex und differenzierbar ist und $x' \in (\mathbb{R}^*_+)^M$. Daraus folgt, dass

$$x'_n = exp((A^\intercal \phi - c)_n),\ n = 1 \dots M.$$

5 Lagrange-Dualität

5.1 Lagrange Formulierung

Mit der Lagrange Formulierung schauen wir uns nun Optimierungsprobleme an, bei denen eine unterhalbstetige, konvexe Funktion unter linearen und konvexen Nebenbedingungen optimiert wird.

Definition 5.1.1 (Lagrange Formulierung). Im Folgenden seien X, Y and Z endlichdimensionale Banach-Räume and sei $\leq_K$ eine partielle Ordnung in Y induziert von einem abgeschlossenen, konvexen Kegel $K \subset Y$. Sei

$$f : X \to \mathbb{R} \cup \{\infty\}$$

unterhalbstetig und konvex,

$$g : X \to Y$$

unterhalbstetig und konvex und

$$h : X \to Z$$

affin.
Wir betrachten das Optimierungsproblem

$$\text{(NR)} \qquad inf_{x \in C}\{f(x) : g(x) \leq_K 0, h(x) = 0\},$$

wobei $C \subset X$ abgeschlossen und konvex ist.

Bemerkung 5.1.2. Ein Kegel ist hier eine Teilmenge des Vektorraums, welches unter Multiplikation mit positiven Skalaren abgeschlossen ist. Der Polarkegel von K ist definiert, als

$$K^+ := \{x^* \in X^* : \langle x^*, x\rangle \geq 0, \forall x \in K\}.$$

Die Ordnung, welche durch K induziert ist, ist $x \leq_K y \Leftrightarrow y - x \in K$.

Definition 5.1.3 (Langragian).
Die Langragian ist definiert als

$$L_{\lambda,x} : Y \times Z \to \mathbb{R} \cup \{\infty\}, (y, z) \mapsto f(x) + \langle\lambda, (g(x) - y, h(x) - z)\rangle,$$

wobei $\lambda \in K^+ \times Z^*$ und $x \in C$.

Um die Rockafellar Dualität nutzen zu können, schreiben wir unser Optimierungsproblem um:

Lemma 5.1.4. $sup_{\lambda \in K^+ \times Z^*} L_{\lambda,x}(y, z) = \begin{cases} f(x) & \text{falls } g(x) \leq_K y, h(x) = z \\ \infty & \text{sonst .} \end{cases}$

Beweis. 1. Sei $x \in C$ so, dass $h(x) = z$ und $g(x) \leq_K y$, dann ist für $\lambda = (\lambda_1, \lambda_2), \lambda_1 \in K^+$ und $\lambda_2 \in Z^*$:

$$\langle\lambda, (g(x) - y, h(x) - z)\rangle = \langle\lambda_1, g(x) - y\rangle + \langle\lambda_2, 0\rangle = -\langle\lambda_1, y - g(x)\rangle \leq 0,$$

da $\lambda_1 \in K^+$ und $y - g(x) \in K$. Also gilt

$$sup_{\lambda \in K^+ \times Z^*} L_{\lambda,x}(y,z) \leq f(x).$$

Da $0 \in K^+$ und $L_{(0,\lambda_2),x}(y,z) = f(x)$ ist damit

$$sup_{\lambda \in K^+ \times Z^*} L_{\lambda,x}(y,z) = f(x).$$

2. Sei $h(x) \neq z$, dann gibt es eine Basis von Z mit dem Element $h(x) - z$ (da Z ein endlichdimensionaler Banachraum ist). Setze $\lambda_2^n(h(x) - z) = n,\ n \in \mathbb{N}$ und $\lambda_2^n(b)$ =0 für alle anderen Elemente der Basis. Sei $\lambda_n := (0, \lambda_2^n)$. Dann ist

$$lim_{n\to\infty} L_{\lambda_n,x}(y,z) = lim_{n\to\infty} f(x) + n = \infty.$$

3. Sei $g(x) \not\leq_K y$, dann ist $y - g(x) \notin K$. Da K konvex ist und $K \neq Y$, da $y - g(x) \notin K$, kann eine Hyperebene zwischen K und $y - g(x)$ gelegt werden (wobei Elemente von K auf der Hyperebene liegen können). Setze λ_1^1 so, dass $\lambda_1^1(e) = 0$ für alle Elemente e der Hyperebene ist und $\lambda_1^1(e) < 0$ für alle Elemente e auf der Seite von $y - g(x)$ der Hyperebene. Dann ist insbesondere $\lambda_1^1(k) \geq 0$ für alle $k \in K$. Also $\lambda_1^1 \in K^+$. Außerdem ist $\lambda_1^1(g(x) - y) > 0$. Setze $\lambda_1^n = n \cdot \lambda_1^1 \in K^+ \forall n \in \mathbb{N}$. Dann ist für $\lambda_n := (\lambda_1^n, 0)$:

$$lim_{n\to\infty} L_{\lambda_n,x}(y,z) = lim_{n\to\infty} f(x) + n \cdot \lambda_1^1(g(x) - y) = \infty.$$

□

Bemerkung 5.1.5. Unser Optimierungsproblem in (NR) der Definition 5.1.1 kann also umgeschrieben werden zu:

$$p = v(0,0) = inf_{x\in C} sup_{\lambda \in K^+ \times Z^*} L_{\lambda,x}(0,0).$$

Wenn wir

$$F : X \times Y \times Z \to \mathbb{R} \cup \{\infty\}, (x,y,z) \mapsto \begin{cases} sup_{\lambda \in K^+ \times Z^*} L_{\lambda,x}(y,z) & x \in C \\ \infty & \text{sonst} \end{cases}$$

setzen, erhalten wir so unser primales Problem von der Rockafellar-Dualität, Definition 3.2.1.

Unser d in der Rockafellar Formulierung, Definition 4.1.1, können wir ebenfalls unschreiben zu einer von der Lagragian abhängigen Form.

Lemma 5.1.6.

$$d = v^{**}(0,0) \overset{(1)}{=} sup_{\sigma \in Y^* \times Z^*}\{-v^*(-\sigma)\} \overset{(2)}{=} sup_{\lambda \in K^+ \times Z^*} inf_{x \in C} L_{\lambda,x}(0,0).$$

Beweis. (1) kommt daher, dass

$$\begin{aligned} v^{**}(0,0) &\overset{\text{Definition 3.1.1}}{=} sup_{\sigma \in Y^* \times Z^*}\{\langle (0,0), -\sigma\rangle - v^*(-\sigma)\} \\ &= sup_{\sigma \in Y^* \times Z^*}\{-v^*(-\sigma)\}. \end{aligned}$$

Da $Y^* \times Z^*$ ein $\mathbb{R}$-Vektorraum ist, ist es egal, ob wir $-\sigma$ oder σ in der Gleichung einsetzen.

Zu zeigen ist noch (2). Für $\lambda \in Y^* \times Z^*$ gilt:

$$\begin{aligned} v^*(-\lambda) &= sup_{(y,z)\in Y\times Z}\{\langle -\lambda, (y,z)\rangle - v(y,z))\} \\ &= sup_{(y,z)\in Y\times Z}[\langle -\lambda, (y,z)\rangle - inf_{x\in C}\{f(x)|g(x) \leq_K y, h(x) = z\}] \\ &= sup_{(x,y,z)\in C\times Y\times Z}[\langle -\lambda, (y,z)\rangle - f(x)|g(x) \leq_K y, h(x) = z\}]. \end{aligned}$$

Da das Supremum über alle $z \in Z$ genommen wird und gefordert wird, dass $h(x) = z$ ist können wir das z ersetzen durch $h(x)$. Sei außerdem $\zeta = y - g(x) \in K$, dann können wir die Bedingung $g(x) \leq_K y$ umschreiben zu $\zeta \in K$ und den Term umformen zu:

$$\begin{aligned} v^*(-\lambda) &= sup_{(x,\zeta)\in C\times K}[\langle -\lambda, (g(x), h(x))\rangle + \langle -\lambda, (\zeta, 0)\rangle - f(x)] \\ &= -inf_{(x,\zeta)\in C\times K}[L_{\lambda,x}(0,0) + \langle \lambda, (\zeta, 0)\rangle] \\ &= \begin{cases} -inf_{x\in C} L_{\lambda,x}(0,0) & \text{falls } \lambda \in K^+ \times Z^* \\ \infty & \text{sonst} \end{cases}. \end{aligned}$$

Das Letzte folgt daher, dass wenn $\lambda \in K^+ \times Z^*$, wegen $\zeta \in K$ immer $\langle \lambda, (\zeta, 0)\rangle \geq 0$. Außerdem ist $0 \in K$.

Ist anderseits $\lambda \notin K^+ \times Z^*$, dann gibt es ein $\zeta' \in K$ mit $\langle \lambda, (\zeta', 0)\rangle < 0$. Setze $\zeta_n = n \cdot \zeta'$, dann ist $lim_{n\to\infty} L_{\lambda,x}(0,0) + \langle \lambda, (\zeta_n, 0)\rangle = -\infty$.

□

Bemerkung 5.1.7. Die schwache Rockafellar-Dualität, Satz 3.2.2, gibt uns also, dass

$$inf_{x\in C} sup_{\lambda\in K^+\times Z^*} L_{\lambda,x}(0,0) \geq sup_{\lambda\in K^+\times Z^*} inf_{x\in C} L_{\lambda,x}(0,0).$$

Diese Ungleichung gilt bereits aus der Definition von dem Supremum und Infimum. Die Rockafellar Dualität, Satz 3.2.3, gibt uns jedoch neue Erkenntnisse über lineare Optimierung.

5.2 Lineare Optimierung

Mit der neuen Beschränkung auf lineare Optimierungsprobleme können wir die Fenchel und Lagrange-Dualität nun anwenden, um eine starke Dualität zu erhalten mit der minimalen Einschränkung, dass das duale oder primale Optimierungsproblem lösbar ist.

Zunächst definieren wir, was ein lineares Optimierungsproblem ist.

Definition 5.2.1 (Standardversion des linearen Optimierungsproblems). Sei $c \in \mathbb{R}^N, b \in \mathbb{R}^M, A \in \mathbb{R}^{M\times N}$. Wir betrachten die von K_i induzierte Ordnung $\leq := \leq_{K^i}$ auf $\mathbb{R}^i$ mit $K^i := \mathbb{R}^i_+$. Also $x \leq y \Leftrightarrow x_j \leq y_j,\ j = 1 \ldots i$.

Die Standardversion des linearen Optimierungsproblems ist gegeben durch:

(NR 1) $$\tilde{p} = max_{x\geq 0}\{\langle c, x\rangle : Ax \leq b\}.$$

Satz 5.2.2 (Duales Problem zur Standardversion des linearen Optimierungsproblems). Sei ein lineares Optimierungsproblem $\tilde{p}$ in der Standardversion gegeben, so ist das zu $\tilde{p}$ duale Optimierungsproblem:

(NR 2) $$\tilde{d} = min_{\lambda\geq 0}\{\langle b, \lambda\rangle : A^\mathsf{T}\lambda \geq c\}.$$

und es gilt $\tilde{p} \leq \tilde{d}$.

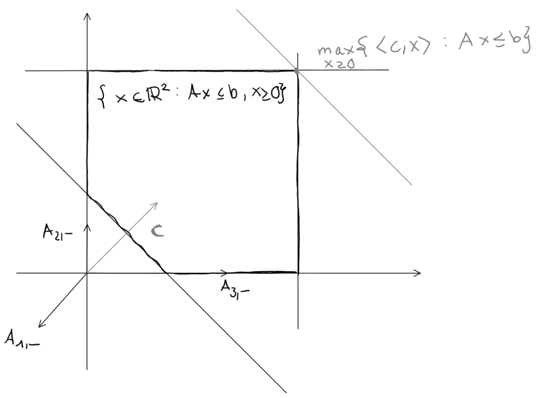

Abbildung 6: Standardversion Lineares Optimierungsproblem

Beweis.

$$\begin{aligned} p :=& -\tilde{p} \\ &= -max_{x\geq 0}\{\langle c, x\rangle : Ax \leq b\} \\ &= min_{x\geq 0}\{\langle -c, x\rangle : Ax \leq b\}. \end{aligned}$$

Wir schreiben uns dazu unsere Langragian:

$$L_{\lambda,x}(y) = \langle -c, x\rangle + \langle \lambda, (Ax - b) - y\rangle. \tag{NR 3}$$

Durch die Bemerkung 5.1.5 wissen wir, dass wir p schreiben können als:

$$p = inf_{x\geq 0} sup_{\lambda \geq 0} L_{\lambda,x}(0).$$

(Wir können den Dualraum von $\mathbb{R}^M$ als $\mathbb{R}^M$ auffassen und dann ist der Polarkegel von $\mathbb{R}^M_+$ wieder $\mathbb{R}^M_+$.)

Das zu p duale Problem ist nach Lemma 5.1.6:

$$d = sup_{\lambda\geq 0} inf_{x\geq 0} L_{\lambda,x}(0).$$

$$\begin{aligned} inf_{x\geq 0} L_{\lambda,x}(0) &= inf_{x\geq 0}\{\langle -c, x\rangle + \langle \lambda, Ax - b\rangle\} \\ &= inf_{x\geq 0}\{\langle -c, x\rangle + \langle \lambda, Ax\rangle - \langle \lambda, b\rangle\} \\ &= inf_{x\geq 0}\{\langle -c, x\rangle + \langle A^\intercal \lambda, x\rangle - \langle \lambda, b\rangle\} \\ &= inf_{x\geq 0}\{\langle -c + A^\intercal \lambda, x\rangle\} - \langle \lambda, b\rangle \\ &\overset{(*)}{=} \begin{cases} -\langle \lambda, b\rangle & \text{falls } A^\intercal\lambda \geq c \\ -\infty & \text{sonst} \end{cases}. \end{aligned} \tag{1}$$

(*) kommt daher, dass wenn

$$-c + A^*\lambda \geq 0 \Rightarrow \langle -c + A^\intercal\lambda, x\rangle \geq 0 \ \forall x \geq 0 \text{ und } \langle -c + A^\intercal\lambda, 0\rangle = 0.$$

Ist anderseits $-c + A^\intercal\lambda \not\geq 0$, dann existiert ein $i \in \{1 \dots N\}$ mit $(-c + A^\intercal\lambda)_i < 0$. Setze $x^n_j = 0$ für $j \neq i$ und $x^n_i = -\frac{n}{(-c+A^*\lambda)_i} > 0$, dann ist $\langle -c + A^\intercal\lambda, x^n\rangle = -n$. Da man $n \in \mathbb{N}$ beliebig wählen kann, folgt die Behauptung.

Also ist unser zu p duales Problem

$$\begin{aligned} d &= sup_{\lambda\geq 0} inf_{x\geq 0} L_{\lambda,x}(0) \\ &\overset{(1)}{=} -sup_{\lambda\geq 0}\{-\langle \lambda, b\rangle : A^\intercal\lambda \geq c\} \\ &= -inf_{\lambda\geq 0}\{\langle \lambda, b\rangle : A^\intercal\lambda \geq c\} \\ &\overset{(**)}{=} -min_{\lambda\geq 0}\{\langle \lambda, b\rangle : A^\intercal\lambda \geq c\} \\ &\overset{\text{(NR 2), Satz5.2.2}}{:=} -\tilde{d}. \end{aligned}$$

Bei $(**)$ geht ein, dass $h : \lambda \to \langle \lambda, b\rangle$ linear ist. $M := \{A^\intercal \lambda \geq c\}$ ist abgeschlossen. h ist in der jeweiligen λ_k Richtung entweder konstant oder streng monoton mit $\lim\limits_{\lambda_k \to \infty} h(\lambda_1, \ldots, \lambda_M) \in \{-\infty, \infty\}$ und $\lim\limits_{\lambda_k \to -\infty} h(\lambda_1, \ldots, \lambda_M) \in \{-\infty, \infty\}$, $k = 1 \ldots M$. Dabei ist dies unabhängig davon, welche Werte $\lambda_i,\ i \in \{1 \ldots M)\} \setminus k$ besitzen.
Dadurch gibt es drei Fälle:

1. h nimmt auf M sein Minimum (auch) am Rand an.
2. h besitzt auf M kein Infimum.
3. $M = \{\}$

Da im 2. und 3. Fall, $inf_{\lambda \geq 0}\{\langle \lambda, b\rangle : A^\intercal \lambda \geq c\} = min_{\lambda \geq 0}\{\langle \lambda, b\rangle : A^\intercal \lambda \geq c\} = -\infty$ ist, und im 1. Fall das Infimum von h auf M sein Minimum ist, folgt $(**)$.

Offenbar folgt aus $p \geq d$, Satz 3.2.2, dass $\tilde{p} \leq \tilde{d}$.

□

Nun können wir die Fenchel-Dualität nutzen um zu zeigen, dass unsere beiden Optimierungsprobleme stark dual sind zueinander, falls eines lösbar ist.

Satz 5.2.3. Falls das primale Problem oder das duale Problem lösbar ist, dann existiert keine Dualitätslücke. Falls der gemeinsame optimale Wert endlich ist, haben beide Probleme optimale Löser.

Beweis. 1. Wir schauen uns, um die Rockafellar- Dualität anzuwenden, zunächst $-\tilde{p}$ an:

Wenn wir F wie folgt setzen, bekommen wir als $p = inf_{x \in X}\{F(x, 0)\} = -\tilde{p}$

$$F : \mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R} \cup \{\infty\}, (x, y) \mapsto \begin{cases} \langle -c, x\rangle & Ax \leq b + y \ \ \&\ \ x \geq 0 \\ \infty & \text{sonst} \end{cases}$$

Wir können es umschreiben zu

$$F(x, y) = \langle -c, x\rangle + \iota_P$$

mit

$$\begin{aligned} P &= \{(x, y) \in \mathbb{R}^N \times \mathbb{R}^M : Ax - y \leq b \ \ \&\ \ x \geq 0\} \\ &= \cap_{i \in \{1..M\}}\{x \in \mathbb{R}^N : \langle (A_{i,-}, -e_i^M), (x, y)\rangle \leq b_i\} \cap \cap_{i \in \{1..N\}}\{x \in \mathbb{R}^N : \langle -e_i^N, x\rangle \leq 0\}, \end{aligned}$$

wobei e_i^M, der i-te Einheitsvektor in $\mathbb{R}^M$ und e_i^N, der i-te Einheitsvektor in $\mathbb{R}^N$ sei. Dabei fällt uns auf, dass P eine polyhedrische Menge (Definition 4.3.1) und $(x, y) \mapsto \langle -c, x\rangle$ affin ist. Damit ist F eine polyhedrische Funktion (Definition 4.3.2).

Sei $v : Y \to \mathbb{R} \cup \{\infty\}, y \mapsto inf_{x \in X}\{F(x, y)\}$, wie in der Definition zur Rockafellar-Dualität, 3.2.1.

Falls das primale Problem lösbar ist, so ist $\tilde{p} \in \mathbb{R} \Rightarrow p = -\tilde{p} \in \mathbb{R}$, also

(a) $\exists (x, y) \in \mathbb{R}^N \times \mathbb{R}^M, F(x, y) < \infty$
(b) $v(0) = inf_{x \in X}\{\langle -c, x\rangle : Ax \leq b, x \geq 0\} = p \leq \infty \Rightarrow 0 \in domv$

Da F polyhedrisch ist, sagen uns Bemerkungen 4.3.7 und 4.3.4, dass dann $p = d$ und damit $\tilde{p} = \tilde{d}$.

2. Das duale Problem von

$$
\begin{aligned}
d &:= -\tilde{d} \\
&\overset{\text{NR 2, Satz5.2.2}}{=} -min_{\lambda\geq 0}\{\langle b,\lambda\rangle : A^\intercal\lambda \geq c\} \\
&= max_{\lambda\geq 0}\{\langle -b,\lambda\rangle : -A^\intercal\lambda \leq -c\}
\end{aligned}
$$

ist nach Satz 5.2.2:

$$
\begin{aligned}
p &:= min_{x\geq 0}\{\langle -c,x\rangle : (-A^\intercal)^\intercal x \geq -b\} \\
&= -max_{x\geq 0}\{\langle c,x\rangle : Ax \leq b\} \\
&\overset{\text{NR 1, Satz 5.2.1}}{=} -\tilde{p}
\end{aligned}
$$

Falls das duale Problem von $\tilde{d}$ lösbar ist, ist es ebenfalls d. Dann ist nach 1. $d = p \Leftrightarrow \tilde{p} = \tilde{d}$.

3. $h_1 : \mathbb{R}^N \to \mathbb{R}, x \mapsto \langle c,x\rangle, h_2 : \mathbb{R}^M \to \mathbb{R}, \lambda \mapsto \langle b,\lambda\rangle$ sind linear und $D_1 := \{x \in \mathbb{R}^N : Ax \leq b\} \cap \{x \in \mathbb{R}^N : x \geq 0\}$ und $D_2 := \{\lambda \in \mathbb{R}^M : A^\intercal\lambda \geq c\} \cap \{\lambda \in \mathbb{R}^M : \lambda \geq 0\}$ abgeschlossen.

   Da h_1 linear ist, ist h_1 in der jeweiligen x_k Richtung entweder konstant oder streng monoton mit $\lim\limits_{x_k\to\infty} h_1(x_1,\ldots,x_N) \in \{-\infty,\infty\}$ und $\lim\limits_{x_k\to-\infty} h_1(x_1,\ldots,x_N) \in \{-\infty,\infty\}$, $k = 1\ldots N$. Dabei ist dies unabhängig davon, welche Werte $x_i,\ i \in \{1\ldots N\} \setminus k$ besitzen.

   Dadurch gibt es für h_1 drei Fälle:

   (a) h_1 nimmt auf D_1 sein Maximum am Rand an.
   (b) h_1 besitzt auf D_1 kein Supremum.
   (c) $D_1 = \{\}$.

   Also existiert ein optimaler Löser von $\tilde{p}$, falls der gemeinsame optimale Wert endlich ist, also die Menge D_1 nicht leer sind und es ein h_1 ein Supremum besitzt. Genauso können wir schlussfolgern, dass wenn der gemeinsame optimale Wert endlich ist, $\tilde{d}$ einen optimalen Löser besitzt.

□

Man kann das Lineare Optimierungsproblem leicht modifizieren, indem man andere Einschränkungen an x stellt. Dies ändert die duale Funktion, wie hier in den Tabellen dargestellt.

Primale Einschränkung	Duale Variable
$Ax \leq b$	$\lambda \geq 0$
$Ax = b$	λ ist frei wählbar
$Ax \geq b$	$\lambda \leq 0$
Primale Variable	Duale Einschränkung
$x \geq 0$	$A^\intercal\lambda \geq c$
x ist frei wählbar	$A^\intercal\lambda = c$
$x \leq 0$	$A^\intercal\lambda \leq c$

Das sieht man, wenn man die Lagrange-Dualität nutzt und den Beweis zu dem Satz 5.2.2 an der Stelle (NR 3) modifiziert und anstelle von $x \geq 0$, $x \leq 0$ oder x frei wählbar, fordert.

Für die modifizierten primalen und dualen Probleme gilt ebenfalls Satz 5.2.3, weil F polyhedrisch ist.

6 Zusammenfassung

Zusammenfassend haben wir unterschiedliche Arten und Weisen kennengelernt, duale Optimierungsprobleme zu finden. Dabei haben wir uns immer weiter eingeschränkt. Dafür haben wir kompaktere Formen des dualen Problems und einfachere Bedingungen für eine starke Dualität gewonnen.

So ist die Rockafellar-Dualität anwendbar auf alle primalen Optimierungsprobleme (p) von Funktionen, welche von Banachräumen nach $\mathbb{R} \cup \{\infty\}$ gehen. Ein weiteres schwach duales Optimierungsproblem (d) haben wir für alle primalen Optimierungsprobleme gefunden. Die Bedingung damit diese auch stark dual zueinander sind (dass der Lagrange-Multiplikator von v an der Stelle 0 existiert) ist jedoch unhandlich zu überprüfen. Gleichzeitig haben wir festgestellt, dass die Bedingung notwendig ist, falls man die Optimierungsprobleme nicht weiter beschränkt.

In der Fenchel-Dualität haben wir nun uns zunächst weiter beschränkt auf Optimierungsprobleme über Funktionen F, welche sich als $(x, y) \mapsto f(x) + g(Ax + y)$ schreiben lassen. Das schwach duale Optimierungsproblem konnte man in Abhängigkeit von f und g umformulieren. Um handlichere Bedingungen zu bekommen für eine starke Dualität, haben wir f und g zunächst auf unterhalbstetige, konvexe Funktionen beschränkt und dann auf polyhedrische Funktionen. Außerdem haben wir die Löser von (p) zu den Lösern in (d) in Zusammenhang gestellt. In den Beispielen haben wir die Fenchel-Dualität angewandt, um das Optimierungsproblem: Die Entropie unter Nebenbedingungen zu maximieren, näher kennenzulernen.

Bei der Lagrange- Dualität haben wir uns nun Optimierungsprobleme angeschaut, wo eine unterhalbstetige, konvexe Funktion unter linearen und konvexen Nebenbedingungen optimiert wird. Dabei haben wir einen Zusammenhang zwischen dem primalen/schwach dualen Optimierungsproblem und der Lagragian-Funktion kennengelernt. Als weitere Klasse von Optimierungsproblemen haben wir uns die Klasse der linearen Optimierungsprogramme angeschaut. Dort konnten wir die Lagrange- Dualität und die Fenchel-Dualität nutzen, um ein dazu stark duales lineares Optimierungsproblem zu finden. Dabei wissen wir, dass diese stark dual sind, falls eines der beiden lösbar ist.

Weiterführend kann man sich weitere Anwendungen näher anschauen, wo wir die dualen Optimierungsprobleme nutzen können. Insbesondere kann man in der Anwendung die Bedeutung des dualen Problems näher betrachten, um mehr über die Art des primalen Problems zu erfahren.

Weitere Klassen (zum Beispiel diskrete oder nicht-konvexe) Optimierungsprobleme kann man weiter untersuchen und dort nach expliziten dualen Optimierungsproblemen suchen.

7 Anhang

7.1 Literatur

[1] R. Tyrrell Rockafellar. *Conjugate Duality and Optimization.* `https://sites.math.washington.edu/~rtr/papers/rtr054-ConjugateDuality.pdf`. 1974.

[2] Qiji Zhu. “Convex Analysis and Duality with Applications. 3. Duality”. Universitätsvorlesung. 2022.

7.2 Weitere Beweise

Der Beweis der Bemerkung 3.1.3:

Beweis. Sei $\lambda \in [0,1]$ und $x^*, y^* \in X^*$.

$$\begin{aligned}
f^*(\lambda x^* + (1-\lambda)y^*) &= sup_{x\in X}\{\langle \lambda x^* + (1-\lambda)y^*, x\rangle - f(x)\} \\
&= sup_{x\in X}\{\lambda\langle x^*, x\rangle + (1-\lambda)\langle y^*, x\rangle - \lambda f(x) - (1-\lambda)f(x)\} \\
&\leq sup_{x\in X}\{\lambda\langle x^*, x\rangle - \lambda f(x)\} + sup_{x\in X}\{(1-\lambda)\langle y^*, x\rangle - (1-\lambda)f(x)\} \\
&= \lambda sup_{x\in X}\{\langle x^*, x\rangle - f(x)\} + (1-\lambda)sup_{x\in X}\{\langle y^*, x\rangle - f(x)\} \\
&= \lambda f^*(x^*) + (1-\lambda)f^*(y^*).
\end{aligned}$$

□