

Die Lagrange - duale Optimierungsaufgabe und ihre geometrische Interpretation

Vortrag von Elisabeth Zschorlich

09.12.2008

Inhaltsverzeichnis

| | | |
|----------|--|-----------|
| 1 | Die Lagrange - duale Optimierungsaufgabe | 2 |
| 1.1 | Einführung | 2 |
| 1.2 | Schwache Dualität | 3 |
| 1.3 | Starke Dualität und Slater - Bedingung | 4 |
| 1.4 | Anwendung: Gemischte Strategien für Matrixspiele | 7 |
| 2 | Geometrische Interpretation der Lagrange - dualen Optimierungsaufgabe | 10 |
| 2.1 | Schwache und starke Dualität anhand von Mengen | 10 |
| 2.2 | Beweis der starken Dualität unter der Slater - Bedingung | 12 |
| 2.3 | Interpretation bei Mehrzieloptimierungsaufgaben | 13 |

1 Die Lagrange - duale Optimierungsaufgabe

1.1 Einführung

Definition 1.1.1

Das Optimierungsproblem mit der Lagrange - dualen Funktion als Zielfunktion

$$\begin{aligned} \max_{(\lambda, \nu) \in \mathbb{R}^m \times \mathbb{R}^p} \quad & g(\lambda, \nu) \\ \text{s.t.} \quad & \lambda \succeq 0, \end{aligned} \tag{1}$$

heißt *Lagrange - duales Optimierungsproblem* zum Optimierungsproblem in Normalform

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq 0 \quad \forall i = 1, \dots, m \\ & h_i(x) = 0 \quad \forall i = 1, \dots, p. \end{aligned} \tag{2}$$

In diesem Zusammenhang nennt man (2) auch *primale Optimierungsaufgabe*.

Definition 1.1.2

Ein Paar $(\lambda, \nu) \in \mathbb{R}^m \times \mathbb{R}^p$ heißt *dual zulässig*, wenn

$$\lambda \succeq 0 \text{ und } g(\lambda, \nu) = \inf\{L(x, \lambda, \nu) : x \in D^1\} > -\infty,$$

d.h. die Lagrange - Funktion ist nach unten beschränkt. Sind $(\lambda^*, \nu^*) \in \mathbf{dom} g^2$ optimal für die Optimierungsaufgabe (1), so nennt man sie *dual optimal* oder *optimale Lagrange - Multiplikatoren*.

Bemerkung 1.1.3

Die Lagrange - duale Optimierungsaufgabe (1) ist eine konvexe Optimierungsaufgabe, da die Zielfunktion konkav und die Restriktionen konvex sind. Diese Eigenschaft hängt nicht davon ab, ob die primale Optimierungsaufgabe (2) konvex ist oder nicht.

Bemerkung 1.1.4

Es ist nicht unüblich, dass die Dimension des Definitionsbereiches $\mathbf{dom} g$ der dualen Funktion kleiner als $m + p$ ist. In vielen solcher Fälle kann man die affine Hülle des Definitionsbereiches $\mathbf{aff}(\mathbf{dom} g)^3$ als Menge von linearen Gleichungsrestriktionen beschreiben. Wir können also eine zum dualen Problem äquivalente Optimierungsaufgabe formulieren, in der diese Gleichungsrestriktionen explizit angegeben werden.

Beispiel 1.1.5

Lagrange - duales Programm eines Linearen Programms in Standardform

Die Lagrange - duale Funktion für das LP in Standardform

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \succeq 0 \end{aligned}$$

¹ $D = \bigcap_{i=0}^m \mathbf{dom} f_i \cap \bigcap_{i=1}^p \mathbf{dom} h_i$

² $\mathbf{dom} g = \{(\lambda, \nu) \in \mathbb{R}^m \times \mathbb{R}^p : g(\lambda, \nu) > -\infty\}$

³ $\mathbf{aff}(\mathbf{dom} g) = \bigcap \{A \subset \mathbb{R}^m \times \mathbb{R}^p : \mathbf{dom} g \subseteq A\}$, wobei A affiner Unterraum, d. h. $A = x_0 + U$ mit $x_0 \in \mathbb{R}^{m+p}$ und $U \in \mathbb{R}^m \times \mathbb{R}^p$ linearer Unterraum

ist gegeben durch

$$g(\lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{sonst.} \end{cases}$$

(vgl.:

$$\begin{aligned} L(x, \lambda, \nu) &= c^T x - \lambda^T x + \nu^T (Ax - b) \\ \Rightarrow g(\lambda, \nu) &= -b^T \nu + \inf_{x \in \mathbb{R}^n} \underbrace{(c - \lambda + A^T \nu)^T x}_{=: q} \end{aligned}$$

Die lineare Funktion q hat bei Minimierung den Grenzwert $-\infty$, es sei denn $q = 0$, d. h. $c - \lambda + A^T \nu = 0$.)

Durch das Lagrange - duale Problem (1) wird die Lagrange - duale Funktion g unter der Bedingung $\lambda \succeq 0$ maximiert. Dabei ist g genau dann endlich, wenn $A^T \nu - \lambda + c = 0$. Somit können wir eine äquivalente Optimierungsaufgabe formulieren, die die genannten Bedingungen explizit als Gleichungsbedingungen beinhaltet:

$$\begin{aligned} \max_{(\lambda, \nu) \in \mathbb{R}^m \times \mathbb{R}^p} & -b^T \nu \\ \text{s.t.} & A^T \nu - \lambda + c = 0 \\ & \lambda \succeq 0. \end{aligned}$$

Bemerkung 1.1.6

Die Lagrange - duale Optimierungsaufgabe eines dualen Problems ist äquivalent zur primalen Optimierungsaufgabe.

1.2 Schwache Dualität

Satz 1.2.1

Der Optimalwert g^* des Lagrange - dualen Problems stellt die beste untere Schranke für den Optimalwert f_0^* des primalen Problems dar, d. h.

$$g^* \leq f_0^*.$$

Diese Eigenschaft nennt man *schwache Dualität*.

Bemerkung 1.2.2

Die Ungleichung der schwachen Dualität gilt auch für die folgenden Fälle:

- (i) Die primale Aufgabe ist nach unten unbeschränkt ($f_0^* = -\infty$) und das duale Problem ist unzulässig ($g^* = -\infty$).
- (ii) Die primale Aufgabe ist unzulässig ($f_0^* = \infty$) und das duale Problem ist nach oben unbeschränkt ($g^* = \infty$).

- (iii) Die primale Aufgabe ist unzulässig ($f_0^* = \infty$) und das duale Problem ist unzulässig ($g^* = -\infty$).

Definition 1.2.3

Die Differenz zwischen dem Optimalwert des primalen Problems und der größten oberen Schranke durch die Lagrange - duale Funktion $f_0^* - g^*$ heißt *optimale Dualitätslücke*. Sie ist immer nichtnegativ.

Bemerkung 1.2.4

Das Konzept der schwachen Dualität ist sinnvoll, weil es für schwer zu lösende primale Optimierungsaufgaben eine untere Schranke für deren Optimalwert festlegt. Diese Schranke kann durch die Konvexität des dualen Problems in vielen Fällen leichter bestimmt werden als der eigentlich gesuchte Optimalwert.

1.3 Starke Dualität und Slater - Bedingung

Definition 1.3.1

Wenn die Gleichung $g^* = f_0^*$ gilt, d. h. dass die optimale Dualitätslücke gleich Null ist, sprechen wir von *starker Dualität*.

Bemerkung 1.3.2

Starke Dualität ist im Allgemeinen nicht erfüllt. Betrachten wir aber eine konvexe Optimierungsaufgabe

$$\begin{aligned} \min_{x \in \mathbb{R}^n} & f_0(x) \\ \text{s.t.} & f_i(x) \leq 0 \quad \forall i = 1, \dots, m \\ & Ax = b \end{aligned}$$

mit f_0, \dots, f_m konvex, so wird für gewöhnlich (jedoch nicht immer) starke Dualität erreicht. Es gibt Bedingungen, die bei konvexen Optimierungsaufgaben gelten können, aus denen starke Dualität folgt. Es ist oft leichter diese nachzuweisen, anstatt starke Dualität direkt zu zeigen.

Definition 1.3.3

Die *Slater - Bedingung* bedeutet, dass ein $x \in \mathbf{relint}D^4$ existiert, sodass

$$f_i(x) < 0 \quad \forall i = 1, \dots, m \quad \text{und} \quad Ax = b.$$

Solch ein Punkt x heißt *strikt zulässig*.

Definition 1.3.4

Die Slater - Bedingung für konvexe Optimierungsaufgaben mit affin - linearen f_i für die ersten $i = 1, \dots, k$ ($k \leq m$) kann verfeinert werden und lautet wie folgt:

Es existiert ein $x \in \mathbf{relint}D$, sodass

$$f_i(x) \leq 0 \quad \forall i = 1, \dots, k, \quad f_i(x) < 0 \quad \forall i = k + 1, \dots, m \quad \text{und} \quad Ax = b.$$

⁴ $\mathbf{relint}D = \{x \in D : \exists \epsilon > 0 | U_\epsilon(x) \cap \mathbf{aff}(D) \subset D\}$

Für die affin - linearen Ungleichungsbedingungen muss also nicht die strikte Ungleichheit gelten.

Beachte:

Für $k = m$ und $\text{dom} f_0$ offen entspricht die verfeinerte Slater - Bedingung der Zulässigkeit.

Satz (von Slater) 1.3.5

Gilt die Slater - Bedingung oder ihre Verfeinerung bei einem konvexen Optimierungsproblem, so gilt starke Dualität.

Bemerkung 1.3.6

Der Satz von Slater besagt auch, dass der duale Optimalwert angenommen wird, also $g^* > -\infty$. Es existieren somit ein dual zulässiges Paar (λ^*, ν^*) und ein primal zulässiger Punkt x^* mit $g(\lambda^*, \nu^*) = g^* = f_0^* = f_0(x^*)$ und (x^*, λ^*, ν^*) ist KKT - Punkt.

Beispiele 1.3.7

- (i) *Lösung von linearen Gleichungssystemen durch die Methode der kleinsten Fehlerquadrate*

Wir betrachten das Problem

$$\begin{array}{ll} \min_{x \in \mathbb{R}^n} & x^T x \\ \text{s.t.} & Ax = b. \end{array}$$

Das zugehörige duale Problem lautet

$$\max_{\nu \in \mathbb{R}^p} g(\nu) = -\frac{1}{4} \nu^T A A^T \nu - b^T \nu$$

(vgl.: Vortrag von Stefanie Fuchs am 2.12.2008, Bsp. II.5, 1. Anstrich), welches ein unrestringiertes, konkaves und quadratisches Maximierungsproblem darstellt. Hier bedeutet die Slater - Bedingung gerade, dass das primale Problem zulässig ist, d. h. b ist im Bild von $A \in \mathbb{R}^{p \times n}$ und somit $f_0^* < \infty$. Es gilt mit Slater also $f_0^* = g^*$.

Beachte:

In diesem Beispiel gilt starke Dualität auch, wenn b nicht im Bild von A und somit die primale Optimierungsaufgabe unzulässig ist ($f_0^* = \infty$), weil dann ein $z \in \mathbb{R}^p$ existiert mit $A^T z = 0$, aber $b^T z \neq 0$ ⁵. Daraus kann man folgern, dass die duale Funktion entlang $\{tz | t \in \mathbb{R}\}$ nach oben unbeschränkt ist und somit auch $g^* = \infty$.

Der Fall $f_0^* = -\infty$, also unbeschränkt, wird durch die schwache Dualität gewährleistet, d. h. $g^* = -\infty$.

- (ii) *Das Lagrange - duale Programm eines quadratischen Programms mit quadratischen Nebenbedingungen (QCQP)*

Wir betrachten das QCQP

$$\begin{array}{ll} \min_{x \in \mathbb{R}^n} & \frac{1}{2} x^T P_0 x + q_0^T x + r_0 \\ \text{s.t.} & \frac{1}{2} x^T P_i x + q_i^T x + r_i \leq 0, \quad \forall i = 1, \dots, m \end{array}$$

⁵Es gilt: $\text{Kern}(A^T) = \text{Bild}(A)^\perp$.

mit $P_0 \in S_{++}^n$ und $P_i \in S_+^n, \forall i = 1, \dots, m$. Die Lagrange - Funktion ist

$$L(x, \lambda) = \frac{1}{2}x^T P(\lambda)x + q(\lambda)^T x + r(\lambda),$$

wobei

$$P(\lambda) = P_0 + \sum_{i=1}^m \lambda_i P_i, \quad q(\lambda) = q_0 + \sum_{i=1}^m \lambda_i q_i, \quad r(\lambda) = r_0 + \sum_{i=1}^m \lambda_i r_i.$$

Einen Ausdruck für die Lagrange - duale Funktion $g(\lambda)$ mit allgemeinem λ zu finden ist möglich, aber schwierig. Wir nehmen daher den für das duale Optimierungsproblem relevanten Fall $\lambda \succeq 0$ an, wodurch die positive Definitheit und Symmetrie von $P(\lambda)$ erhalten bleibt.

Wir erhalten also das duale Optimierungsproblem

$$\begin{aligned} \max_{\lambda \in \mathbb{R}^m} \quad & g(\lambda) = -\frac{1}{2}q(\lambda)^T P(\lambda)^{-1}q(\lambda) + r(\lambda) \\ \text{s.t.} \quad & \lambda \succeq 0. \end{aligned}$$

(vgl.:

$$\begin{aligned} g(\lambda) &= \inf_{x \in \mathbb{R}^n} L(x, \lambda) \quad \text{mit } L(x, \lambda) \text{ quadratische Funktion und } P(\lambda) \in S_{++}^n \\ L_x(x, \lambda) &= P(\lambda)x + q(\lambda) \stackrel{!}{=} \mathbf{0} \\ \Rightarrow x^* &= -P(\lambda)^{-1}q(\lambda) \\ \Rightarrow g(\lambda) &= L(x^*, \lambda) \end{aligned}$$

Die Slater - Bedingung besagt hier, dass starke Dualität gilt, wenn ein x existiert, welches strikt zulässig für die quadratischen Ungleichungsrestriktionen ist.

(iii) *Ein nichtkonvexes, quadratisches Programm mit starker Dualität*

In seltenen Fällen gilt starke Dualität auch bei nichtkonvexen Optimierungsproblemen. Als wichtiges Beispiel dafür betrachten wir nun das Minimierungsproblem mit nichtkonvexer, quadratischer Zielfunktion über der Einheitskugel

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & x^T A x + 2b^T x \\ \text{s.t.} \quad & x^T x \leq 1, \end{aligned}$$

wobei $A \in S^n, A \not\succeq 0$ und $b \in \mathbb{R}^n$. Dieses Programm wird manchmal *Trust - Region - Problem* genannt. Eine Approximation zweiter Ordnung einer Funktion wird dabei über der Einheitskugel minimiert.

Die Lagrange - Funktion ist

$$L(x, \lambda) = x^T A x + 2b^T x + \lambda(x^T x - 1) = x^T (A + \lambda I)x + 2b^T x - \lambda,$$

womit sich die Lagrange - duale Funktion ergibt durch

$$g(\lambda) = \begin{cases} -b^T (A + \lambda I)^{\dagger} b - \lambda & A + \lambda I \succeq 0, b \in \text{im}(A + \lambda I) \\ -\infty & \text{sonst,} \end{cases}$$

wobei $(A + \lambda I)^\dagger$ die Pseudoinverse⁶ von $A + \lambda I$ ist.

(vgl.: Für den Fall $A + \lambda I \succeq 0$ ist $L(x, \lambda)$ quadratisch und konvex.

$$\begin{aligned} \Rightarrow L_x(x, \lambda) &= 2(A + \lambda I)x + 2b \stackrel{!}{=} 0 \\ (b \in \text{im}(A + \lambda I)) \Rightarrow x^* &= -(A + \lambda I)^\dagger b \\ \Rightarrow L(x^*, \lambda) &= b^T \underbrace{((A + \lambda I)^\dagger)^T (A + \lambda I)}_{\text{symmetrisch}} (A + \lambda I)^\dagger b - 2b^T (A + \lambda I)^\dagger b - \lambda \\ &= -b^T (A + \lambda I)^\dagger b - \lambda \end{aligned}$$

Das Lagrange - duale Problem ist demnach

$$\begin{aligned} \max_{\lambda \in \mathbb{R}} \quad & -b^T (A + \lambda I)^\dagger b - \lambda \\ \text{s.t.} \quad & A + \lambda I \succeq 0 \\ & b \in \text{im}(A + \lambda I). \end{aligned}$$

Bei diesem Beispiel ist die optimale Dualitätslücke immer Null. Hier gilt nämlich ein allgemeineres Resultat (siehe Anhang §B.1 im Buch „Convex Optimization“, Stephen P. Boyd):

Bei jedem Optimierungsproblem mit quadratischer Zielfunktion und einer quadratischen Ungleichungsbedingung, für das die Slater - Bedingung gilt, gilt starke Dualität.

1.4 Anwendung: Gemischte Strategien für Matrixspiele

Anwendung findet die starke Dualität beispielsweise in der Spieltheorie. Hier kann man aus der starken Dualität ein wesentliches Resultat für Nullsummenspiele⁷ ableiten.

Wir betrachten ein Spiel mit zwei Spielern:

- Spieler 1 wählt den Spielzug $k \in \{1, \dots, n\}$.
- Spieler 2 wählt den Spielzug $l \in \{1, \dots, m\}$.
- Spieler 1 zahlt dann G_{kl} an Spieler 2, d. h. $G \in \mathbb{R}^{n \times m}$ ist die Auszahlungsmatrix des Spiels.

\Rightarrow Spieler 1 will seine Zahlung minimieren.

\Rightarrow Spieler 2 will seine Zahlung maximieren.

Beide Spieler spielen mit *gemischten Strategien*, was bedeutet, dass jeder Spieler seinen Spielzug zufällig und unabhängig vom Spielzug des anderen Spielers wählt. Die Strategien $u \in \mathbb{R}^n$

⁶Für $Y \in \mathbb{R}^{m \times n}$ ist $X \in \mathbb{R}^{n \times m}$ die (Moore - Penrose -) Pseudoinverse, wenn die Bedingungen (i) $YXY = Y$, (ii) $XYX = X$, (iii) $(YX)^T = YX$ und (iv) $(XY)^T = XY$ gelten.

⁷Ein Tripel $G = (S, T; H)$ heißt Matrixspiel oder 2 - Personen - Nullsummenspiel, wenn S und T zwei endliche, nicht leere Mengen sind und H eine Abbildung $H : S \times T \rightarrow \mathbb{R}^2$, die jedem $(s, t) \in S \times T$ einen Vektor $H(s, t) = (H_1(s), H_2(t))$ zuordnet, wobei $H_1(s) = -H_2(t)$ ist.

und $v \in \mathbb{R}^m$ der Spieler sind somit durch die Wahrscheinlichkeitsverteilungen ihrer Spielzüge gegeben:

$$P(k = i) = u_i, \quad \forall i = 1, \dots, n, \quad P(l = i) = v_i, \quad \forall i = 1, \dots, m.$$

Die erwartete Auszahlung von Spieler 1 an Spieler 2 ist demnach

$$\sum_{k=1}^n \sum_{l=1}^m u_k G_{kl} v_l = u^T G v.$$

\Rightarrow Spieler 1 wählt u so, dass es Lösung von $\min u^T G v$ ist.

\Rightarrow Spieler 2 wählt v so, dass es Lösung von $\max u^T G v$ ist.

Wir betrachten das Spiel nun aus Sicht der einzelnen Spieler:

| <i>Betrachtungen aus der Sicht von</i> | |
|---|---|
| <i>Spieler 1</i> | <i>Spieler 2</i> |
| <p>Annahme: Die Strategie u von Spieler 1 ist Spieler 2 bekannt, was sich anhand seines Maximierungsvorhabens in der erwarteten Auszahlung niederschlägt:</p> $\sup\{u^T G v \mid v \succeq 0, \mathbf{1}^T v = 1\} = \max_{i=1, \dots, m} (G^T u)_i.$ <p>\Rightarrow offensichtlicher Vorteil von Spieler 2</p> <p>Ziel von Spieler 1: Minimierung der gegebenen Worst - Case - Auszahlung durch die Wahl seiner Strategie u</p> <p>Gesucht ist demnach die Lösung u^* von</p> $\begin{aligned} \min_{u \in \mathbb{R}^n} \quad & \max_{i=1, \dots, m} (G^T u)_i \\ \text{s. t.} \quad & u \succeq 0 \\ & \mathbf{1}^T u = 1, \end{aligned} \quad (3)$ <p>was eine stückweise lineare, konvexe Optimierungsaufgabe darstellt. u^* ist somit die kleinste erwartete Auszahlung, die Spieler 1 unter obiger Annahme an Spieler 2 leisten muss.</p> | <p>Annahme: Die Strategie v von Spieler 2 ist Spieler 1 bekannt, was sich anhand seines Minimierungsvorhabens in der erwarteten Auszahlung niederschlägt:</p> $\inf\{u^T G v \mid u \succeq 0, \mathbf{1}^T u = 1\} = \min_{i=1, \dots, n} (G v)_i.$ <p>\Rightarrow offensichtlicher Vorteil von Spieler 1.</p> <p>Ziel von Spieler 2: Maximierung der gegebenen Auszahlung an ihn durch die Wahl seiner Strategie v</p> <p>Gesucht ist demnach die Lösung v^* von</p> $\begin{aligned} \max_{v \in \mathbb{R}^m} \quad & \min_{i=1, \dots, n} (G v)_i \\ \text{s. t.} \quad & v \succeq 0 \\ & \mathbf{1}^T v = 1, \end{aligned} \quad (4)$ <p>was ein konvexes Problem mit stückweise linearer, konkaver Zielfunktion ist. v^* ist somit die größte erwartete Auszahlung, die Spieler 2 unter obiger Annahme bekommt.</p> |

Tabelle 1: Analyse des Spiels aus Sicht der Spieler

Intuitiv ist klar, dass $u^* \geq v^*$ gilt. Die Differenz $u^* - v^*$ können wir als den Vorteil interpretieren, den ein Spieler hat, wenn er die Strategie seines Gegners kennt. Wir beweisen diese

Aussagen, indem wir zeigen, dass die Optimierungsaufgaben (3) und (4) ein primal - duales Paar sind:

Zuerst formulieren wir (3) als ein LP:

$$\begin{aligned} \min_{(t,u) \in \mathbb{R} \times \mathbb{R}^n} \quad & t \\ \text{s.t.} \quad & u \succeq 0 \\ & \mathbf{1}^T u = 1 \\ & G^T u \preceq t \mathbf{1}. \end{aligned}$$

Wir bezeichnen den Lagrange - Multiplikator für $G^T u \preceq t \mathbf{1}$ mit $\lambda \in \mathbb{R}^m$, für $u \succeq 0$ mit $\mu \in \mathbb{R}^n$ und für $\mathbf{1}^T u = 1$ mit $\nu \in \mathbb{R}$. Somit ergibt sich die Lagrange - Funktion mit

$$L(t, u, \lambda, \mu, \nu) = t + \lambda^T (G^T u - t \mathbf{1}) - \mu^T u + \nu (1 - \mathbf{1}^T u) = \nu + (1 - \mathbf{1}^T \lambda) t + (G \lambda - \nu \mathbf{1} - \mu)^T u,$$

sodass sich die Lagrange - duale Funktion wie folgt ergibt:

$$g(\lambda, \mu, \nu) = \begin{cases} \nu & \mathbf{1}^T \lambda = 1, \quad G \lambda - \nu \mathbf{1} = \mu \\ -\infty & \text{sonst.} \end{cases}$$

Die duale Optimierungsaufgabe lautet demnach

$$\begin{aligned} \max_{(\lambda, \mu, \nu) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}} \quad & \nu \\ \text{s.t.} \quad & \lambda \succeq 0 \\ & \mathbf{1}^T \lambda = 1 \\ & \mu \succeq 0 \\ & G \lambda - \nu \mathbf{1} = \mu, \end{aligned}$$

bzw. wenn wir μ eliminieren:

$$\begin{aligned} \max_{(\lambda, \nu) \in \mathbb{R}^m \times \mathbb{R}} \quad & \nu \\ \text{s.t.} \quad & \lambda \succeq 0 \\ & \mathbf{1}^T \lambda = 1 \\ & G \lambda \succeq \nu \mathbf{1}. \end{aligned}$$

Wenn wir λ in v umbenennen, sieht man schnell, dass dieses Problem äquivalent zu (4) ist. Durch die schwache Dualität ergibt sich $v^* \leq u^*$. Wenn die beiden LPs zulässig sind, gilt auch die verfeinerte Slater - Bedingung (siehe Def. 1.3.4, Beachte) und somit sogar starke Dualität. Das bedeutet, dass der Vorteil eines Spielers, der des Gegners Strategie kennt, ausgedrückt durch $u^* - v^*$, Null entspricht.

⇒ Resultat für Matrixspiele:

In einem Matrixspiel mit gemischten Strategien existiert kein Vorteil für einen Spieler, der die Strategie seines Gegenspielers kennt.

2 Geometrische Interpretation der Lagrange - dualen Optimierungsaufgabe

2.1 Schwache und starke Dualität anhand von Mengen

Wir können die Lagrange - Funktion geometrisch interpretieren, indem wir sie als Wertemenge der Restriktionen und der Zielfunktion auffassen:

$$\mathcal{L} = \{(f_1(x), \dots, f_m(x), h_1(x), \dots, h_p(x), f_0(x)) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} | x \in D\}.$$

Den Optimalwert f_0^* von (2) kann man im Sinne von \mathcal{L} ausdrücken als

$$f_0^* = \inf\{t | (u, v, t) \in \mathcal{L}, u \preceq 0, v = 0\}.$$

Um zur Lagrange - dualen Funktion zu gelangen, drücken wir die Lagrange - Funktion als affine Funktion

$$(\lambda, \nu, 1)(u, v, t)^T = \sum_{i=1}^m \lambda_i u_i + \sum_{i=1}^p \nu_i v_i + t, (u, v, t) \in \mathcal{L},$$

aus. Wir erhalten also als duale Funktion

$$g(\lambda, \nu) = \inf\{(\lambda, \nu, 1)(u, v, t)^T | (u, v, t) \in \mathcal{L}\}.$$

Ist dieses Infimum endlich, dann kann man die Lagrange - duale Funktion anhand der Ungleichung

$$(\lambda, \nu, 1)(u, v, t)^T \geq g(\lambda, \nu)$$

als stützende Hyperebene an \mathcal{L} verstehen. Diese wird als *nichtsenkrecht* bezeichnet, weil die letzte Komponente des Normalenvektors nicht null ist.

Setzen wir nun die Ungleichungsbeschränkung $\lambda \succeq 0$ der dualen Optimierungsaufgabe voraus, erhalten wir schwache Dualität:

$$\begin{aligned} f_0^* &= \inf\{t | (u, v, t) \in \mathcal{L}, u \preceq 0, v = 0\} \geq \inf\{(\lambda, \nu, 1)(u, v, t)^T | (u, v, t) \in \mathcal{L}, u \preceq 0, v = 0\} \\ &\geq \inf\{(\lambda, \nu, 1)(u, v, t)^T | (u, v, t) \in \mathcal{L}\} = g(\lambda, \nu). \end{aligned}$$

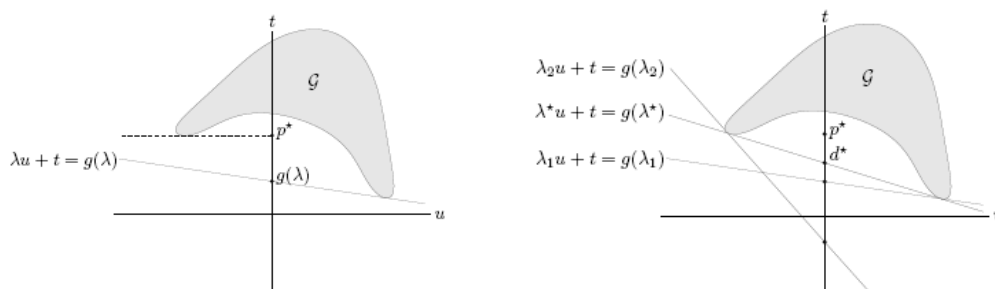


Abbildung 1: Illustration für ein Problem mit einer Ungleichungsrestriktion

Epigraph - Variante

Man kann die Dualität auch in leicht abgewandelter Form interpretieren, was illustrieren soll, warum starke Dualität für die meisten konvexen Optimierungsaufgaben gilt. Wir betrachten hierfür die Minkowski - Summe $\mathcal{A} \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}$ mit

$$\begin{aligned} \mathcal{A} &= \mathcal{L} + (\mathbb{R}_+^m \times \{0\} \times \mathbb{R}_+) \\ &= \{(u, v, t) | \exists x \in D : f_i(x) \leq u_i, \forall i = 1, \dots, m, h_i(x) = v_i, \forall i = 1, \dots, p, f_0(x) \leq t\}. \end{aligned}$$

\mathcal{A} ist also eine Art Epigraph von \mathcal{L} , d. h. \mathcal{A} enthält alle Elemente von \mathcal{L} und alle Punkte, die einen größeren Zielfunktionswert oder Ungleichungsfunktionswert haben.

Wir können den Optimalwert f_0^* von (2) mit \mathcal{A} wie folgt ausdrücken:

$$f_0^* = \inf\{t | (0, 0, t) \in \mathcal{A}\}.$$

Um zur Lagrange - dualen Funktion zu gelangen, minimieren wir die affine Funktion $(\lambda, \nu, 1)(u, v, t)^T$ über \mathcal{A} . Wir erhalten also für $\lambda \succeq 0$

$$g(\lambda, \nu) = \inf\{(\lambda, \nu, 1)(u, v, t)^T | (u, v, t) \in \mathcal{A}\}.$$

Ist dieses Infimum endlich, dann kann man die Lagrange - duale Funktion anhand der Ungleichung

$$(\lambda, \nu, 1)(u, v, t)^T \geq g(\lambda, \nu)$$

als nichtsenkrechte, stützende Hyperebene an \mathcal{A} verstehen.

Die schwache Dualitätsungleichung

$$f_0^* = (\lambda, \nu, 1)(0, 0, f_0^*)^T \geq g(\lambda, \nu)$$

erhalten wir in diesem Fall, da $(0, 0, f_0^*) \in \partial\mathcal{A}$. Starke Dualität würde bei der Epigraph - Interpretation bedeuten, dass eine nichtsenkrechte, stützende Hyperebene an \mathcal{A} im Randpunkt $(0, 0, f_0^*)$ existiert.

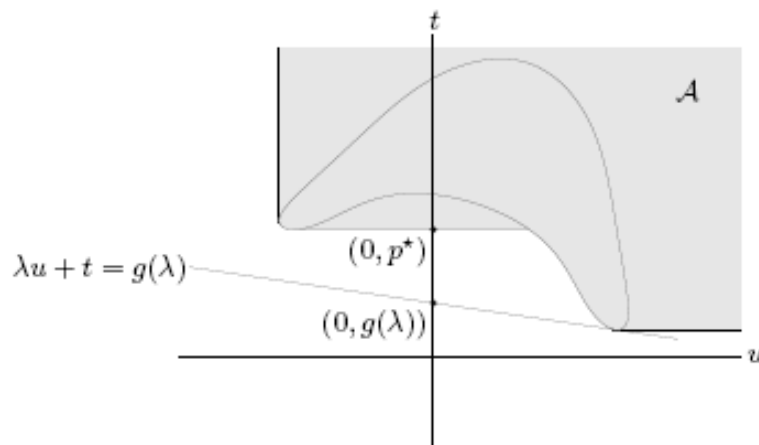


Abbildung 2: Illustration der Epigraph - Interpretation

2.2 Beweis der starken Dualität unter der Slater - Bedingung

Wir beweisen hier, dass aus der Slater - Bedingung die starke Dualität folgt. Wir betrachten also

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq 0 \quad \forall i = 1, \dots, m \\ & Ax = b \end{aligned} \tag{5}$$

mit f_0, \dots, f_m konvex, und nehmen an, dass die Slater - Bedingung gilt:

$$\exists x \in \mathbf{relint}D : f_i(x) < 0 \quad \forall i = 1, \dots, m \text{ und } Ax = b.$$

Um den Beweis zu vereinfachen, treffen wir noch zwei zusätzliche Annahmen: D hat ein nicht-leeres Inneres, d. h. $\mathbf{relint}D = \mathbf{int}D$, und $\mathbf{rank}(A) = p$. Außerdem können wir davon ausgehen, dass f_0^* endlich ist, weil (5) durch die Slater - Bedingung zulässig ist und im nach unten unbeschränkten Fall die Behauptung durch die schwache Dualität gewährleistet wird ($f_0^* = -\infty \geq -\infty = g^*$).

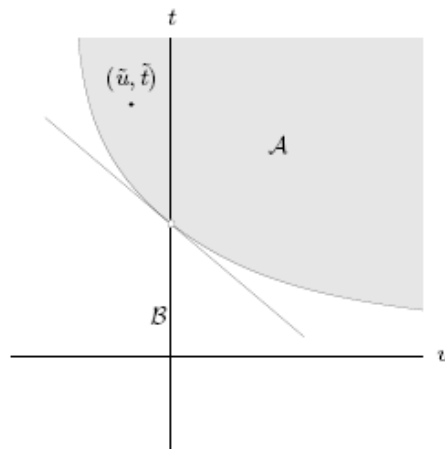


Abbildung 3: **Illustration des Beweises:** Die Menge \mathcal{A} ist schattiert dargestellt und die Menge \mathcal{B} ist die dicke, vertikale Linie, die den Punkt $(0, f_0^*)$ nicht einschließt. Die beiden Mengen sind konvex und disjunkt, sodass sie durch eine Hyperebene getrennt werden können. Die Slater - Bedingung garantiert, dass jede Hyperebene nichtsenkrecht ist, weil sie links vom Punkt $(\tilde{u}, \tilde{t}) = (f_1(\tilde{x}), f_0(\tilde{x}))$, wobei \tilde{x} strikt zulässig, liegen muss.

2.3 Interpretation bei Mehrzieloptimierungsaufgaben

Es gibt eine natürliche Beziehung zwischen der Lagrange - Dualität einer Optimierungsaufgabe ohne Gleichungsrestriktionen

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq 0, \quad \forall i = 1, \dots, m. \end{aligned} \quad (6)$$

und der Skalarisierungsmethode der Mehrzieloptimierungsaufgabe

$$\min(\text{w.r.t. } \mathbb{R}_+^{m+1}) F(x) = (f_1(x), \dots, f_m(x), f_0(x)) \quad (7)$$

(vgl.: Vortrag von Sandra Hartl am 25.11.2008, Abschnitt 4.7.4 und 4.7.5).

Bei der Skalarisierung wählen wir einen positiven Vektor $\tilde{\lambda}$ und minimieren die Funktion $\tilde{\lambda}^T F(x)$, wobei jede Lösung dessen ein Pareto - optimaler Punkt von (7) ist. Weil wir $\tilde{\lambda}$ mit positiven Konstanten skalieren können, wählen wir o. B. d. A. $\tilde{\lambda} = (\lambda, 1)$. Somit minimieren wir nun die Funktion

$$\tilde{\lambda}^T F(x) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x).$$

Diese ist gerade die Lagrange - Funktion von (6).

Beim Beweis der Aussage, dass jeder Pareto - optimale Punkt eines konvexen Mehrzieloptimierungsproblems das Skalarprodukt $\tilde{\lambda}^T F(x)$ minimiert, wird wie in Abschnitt 2.2 die Menge

$$\mathcal{A} = \{t \in \mathbb{R}^{m+1} \mid \exists x \in D : f_i(x) \leq t_i, \forall i = 0, \dots, m\}$$

benutzt. Man hat ebenfalls eine stützende Hyperebene an diese Menge in einem beliebigen Pareto - optimalen Punkt konstruiert.

Der Gewichtungsvektor $\tilde{\lambda}$, wessen erste m Komponenten den Lagrange - Multiplikator der Lagrange - Funktion von (6) darstellen, gibt im Übrigen die relativen Gewichtungen zwischen den Zielfunktionen an. Definieren wir $\tilde{\lambda}$ wie oben, so geben die ersten m Komponenten die Kosten der zugehörigen f_1, \dots, f_m relativ zur Zielfunktion f_0 an.