Geometrische Interpretation

Stefanie Riedel

10. Mai 2010

1 Starke und schwache Dualität über Wertemengen

Wir betrachten eine einfache geometrische Interpretation dualer Funktionen aus der Menge G:

$$G = \{(\underbrace{(f_1(x), ..., f_m(x)}_{u}, \underbrace{h_1(x), ..., h_p(x)}_{v}, \underbrace{f_0(x)}_{t}) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} | x \in D\} \quad (1.1)$$

Die Menge besitzt einen Wertevorrat auf Grundlage der Nebenbedingungen u,v (u Ungleichungs-NB und v Gleichungs-NB) und der Zielfunktion t.
Der Optimalwert p^* unseres Optimierungsproblems ist hier gegeben durch:

$$p^* = \inf\{t | (u, v, t) \in G, u \leq 0, v = 0\}$$

Wir wollen die duale Funktion bei (λ, ν) ausrechnen und minimieren daher

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

Dies beschreibt daher unsere Lagrange-duale Funktion

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

Für den Fall, dass $g(\lambda, \nu) < \infty$ ist, definiert die Ungleichung $(\lambda, \nu, 1)^T (u, v, t) \geq g(\lambda, \nu)$ eine Stützhyperebene zu unserer Menge G. Man bezeichnet dies manchmal auch als nichtvertikale Stützhyperebene, da die letzte Komponente des Normalenvektors ungleich Null ist.

Wir nehmen an, dass $\lambda \geq 0$, womit klar ist, dass $t \geq (\lambda, \nu, 1)^T (u, v, t)$ für $u \leq 0$ und $v = 0$, da $(\lambda, \nu, 1)^T (u, v, t) = \underbrace{\lambda^T u}_{\leq 0} + \underbrace{\nu^T v}_{=0} + t$ ist.

Daraufhin betrachten wir nun wieder unseren Optimalwert p^*. Dann gilt mit der gerade getroffenen Annahme:

$$\begin{aligned} p^* &= \inf\{t|(u,v,t) \in G, u \leq 0, v = 0\} \\ &\geq \inf\{(\lambda,\nu,1)^T(u,v,t)|(u,v,t) \in G, u \leq 0, v = 0\} \\ &\geq \inf\{(\lambda,\nu,1)^T(u,v,t)|(u,v,t) \in G\} \\ &= g(\lambda,\nu) \end{aligned}$$

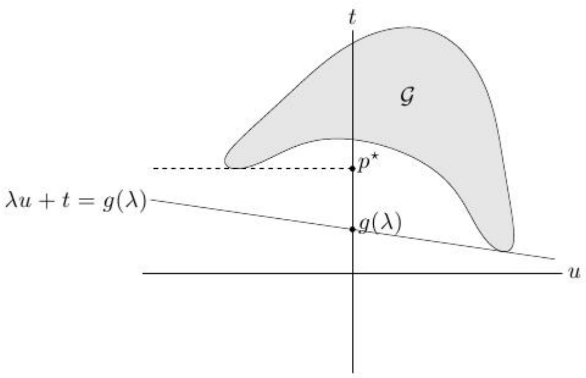
Somit erkennen wir, dass es sich um schwache Dualität handelt, da $p^* \geq g(\lambda,\nu)$ und wir damit eine Dualitätslücke und untere Schranke haben.

Im Nachfolgenden sollen die Beispiele den Sachverhalt graphisch veranschaulichen.

Beispiel 1.1 *Wir betrachten nun ein Problem mit nur einer Ungleichungsnebenbedingung, weshalb unsere Menge G nun wie folgt aussieht:*

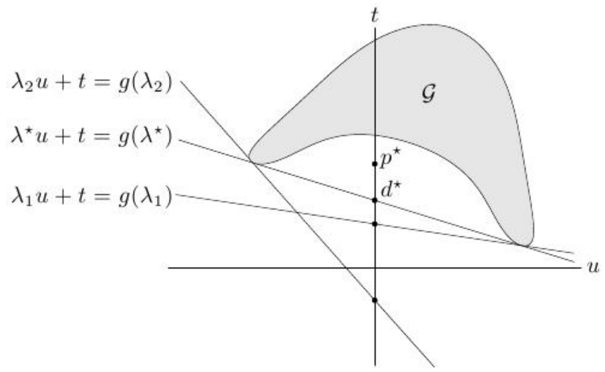
$$G = \{(f_1(x), f_0(x)|x \in D\}.$$

Desweiteren haben wir unseren Optimalwert p^* (beachten, dass $u \leq 0$ ist)



Für irgend ein gegebenes λ können wir die Funktion $(\lambda,1)^T(u,t)$ über unsere Menge G minimieren. Damit erhalten wir eine Stützhyperebene $\lambda u + t = g(\lambda)$ mit negativem Anstieg $-\lambda$.
Der Schnitt der Hyperebene mit der $u = 0$ Achse, gibt uns gerade unser $g(\lambda)$, mit der unteren Grenze $g(\lambda) \leq p^*$.

Beispiel 1.2 *Dieses Beispiel ist ähnlich dem obrigen, nur dass wir hier 3 Werte von λ betrachten, wobei auch der optimale Wert von λ, nämlich λ^*, dabei ist.*



Aus diesen 3 Werten ergeben sich 3 verschiedene Stützhyperebenen $\lambda_1 u + t = g(\lambda_1)$, $\lambda_2 u + t = g(\lambda_2)$ und $\lambda^* u + t = g(\lambda^*)$
Für λ^* ist der Abstand zwischen p^* und der sich aus der Hyperebene bei $u = 0$ ergebene duale Optimalwert d^* minimal.
Auch hier haben wir keine starke Dualität, da wir eine positive Dualitätslücke $p^* - d^*$ haben.

Epigraphvariation

Wir betrachen hier nun eine andere Variante der geometrischen Interpretation der Dualität von Termen der Menge G.
Wir definieren uns eine neue Menge $A \subseteq \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}$ für die gilt, dass

$$A = G + (\mathbb{R}^m_+ \times \{0\} \times \mathbb{R}_+) \tag{1.2}$$

Anders können wir A auch noch darstellen als

$$A = \{(u, v, t) | x \in D, f_i(x) \leq u_i, i = 1, ..., m, h_i(x) = v_i, i = 1, ..., p, f_0(x) \leq t\}$$

Das heißt, dass A hier eine Art Epigraph von G ist, mit allen Punkten aus G und noch weiteren „schlechten"Werten.

Der Optimalwert von A ist hierbei gerade $p^* = \inf\{t | (0, 0, t) \in A\}$.

Um unsere Funktion im Punkt (λ, ν) mit $\lambda \geq 0$ ausrechen zu können, minimieren wir die affine Funktion $(\lambda, \nu, 1)^T (u, v, t)$ über A.

Für $\lambda \geq 0$ ist $g(\lambda, \nu) = \inf\{(\lambda, \nu, 1)^T(u, v, t)|(u, v, t) \in A\}$
Ist das Infimum endlich, so definiert der Ausdruck

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

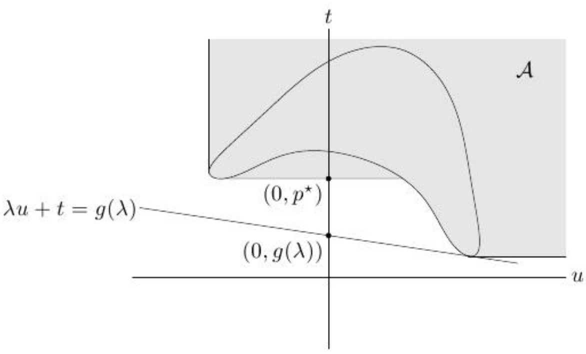
eine nichtvertikale Stützhyperebene zu unserer Menge A

Für $(0, 0, p^*) \in bdA$ (was hier der Rand von A sein soll) haben wir die untere Dualitätsgrenze

$$g(\lambda, \nu) \leq (\lambda, \nu, 1)^T(0, 0, p^*) = p^* \tag{1.3}$$

Starke Dualität gilt in diesen Fall nur genau dann, wenn $p^* = g(\lambda, \nu)$ für ein Paar (λ, ν), d.h.
es existiert eine nichtvertikale Stützhyperebene zur Menge A an ihrem Randpunkt $(0, 0, p^*)$

Beispiel 1.3 *Wir betrachten hier wieder ein Problem mit einer Ungleichungsnebenbedingung im Bezug auf die Menge* $A = \{(u, t)|\exists x \in D, f_0(x) \leq t, f_1(x) \leq u\}$*, welche aus der angezeichneten originalen Menge G und ihrer Verschiebung in Richtung u und t dargestellt ist.*
Ebenfalls wieder gegeben ist der Optimalwert im Punkt $(0, p^*)$*.*



Für irgend ein gegebenes λ *können wir wieder die Funktion* $(\lambda, 1)^T(u, t)$ *minimieren nur diesmal über A. Damit erhalten wir eine Stützhyperebene* $\lambda u + t = g(\lambda)$ *mit negativem Anstieg* $-\lambda$*.*
Der Schnitt der Hyperebene mit der $u = 0$ *Achse, gibt uns gerade unser* $g(\lambda)$*, mit der unteren Grenze* $g(\lambda) \leq p^*$*.*

2 Starke Dualität unter Constraint Qualification

Wir wollen hier zeigen, dass die Slater CQ die starke Dualität für konvexe Probleme garantiert

Beweis: (Starke Dualität unter Slater)

Wir betrachten ein primales Problem mit $f_0, ..., f_m$ konvex und nehmen an, dass Slater gilt.
Es existiert ein $\tilde{x} \in \text{relint}D$ mit $f_i(\tilde{x}) < 0, i = 1, ..., m$ und $A\tilde{x} = b$
Um den Beweis zu vereinfachen, machen wir zwei zusätzliche Annahmen:
1) D hat ein nichtleeres Inneres, sodass $\text{relint}D = \text{int}\, D$
2) und desweitern soll $\text{rang}A = p$ sein.
Das führt uns dazu, dass gerade $p^* < \infty$ (können nur $p^* = -\infty$ oder p^* endlich haben)
Wir haben die vorher verwendete konvexe Menge A und definieren und eine zweite konvexe Menge B für die gilt:

$$B = \{(0, 0, s) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} | s < p^*\}$$

Die beiden Mengen schneiden sich gegenseitig nicht, d.h. $A \cap B = \emptyset$
Nach dem Trennungssatz (vgl. §2.5.1) existieren $(\tilde{\lambda}, \tilde{\nu}, \mu) \neq 0$ und α sodass

$$(u, v, t) \in A \Rightarrow \tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \geq \alpha \tag{2.1}$$

$$(u, v, t) \in B \Rightarrow \tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \leq \alpha \tag{2.2}$$

Aus den beiden Ungleichungen schließen wir, dass $\tilde{\lambda} \geq 0$ und $\mu \geq 0$ sowie $\mu t \leq \alpha$ und $\mu p^* \leq \alpha$ sein müssen.
Zusammen schließen wir darauf, dass gerade für ein $x \in D$

$$\sum_{i=1}^{m} \tilde{\lambda}_i f_i(x) + \tilde{\nu}^T (Ax - b) + \mu f_0(x) \geq \alpha \geq \mu p^* \tag{2.3}$$

An dieser Stelle müssen wir eine Fallunterscheidung für unser $\mu \geq 0$ machen.
Fall 1.
Mit $\mu > 0$ können wir die Ungleichung (2.3) durch μ teilen um $L(x, \frac{\tilde{\lambda}}{\mu}, \frac{\tilde{\nu}}{\mu}) \geq p^*$ zu bekommen $\forall x \in D$
Durch das Minimieren über x, folgt nun, dass $g(\lambda, \nu) \geq p^*$ und definieren dazu $\lambda = \frac{\tilde{\lambda}}{\mu}$ und $\nu = \frac{\tilde{\nu}}{\mu}$
Für schwache Dualität ist $g(\lambda, \nu) \leq p^*$, sodass wir daraus mit $g(\lambda, \nu) \geq p^*$ dann starke Dualität mit $g(\lambda, \nu) = p^*$ erhalten.
Somit ist zunächst das duale Optimum für $\mu > 0$ erreicht.

Fall 2.
Wir betrachten nun noch den Fall, dass $\mu = 0$.
In (2.3) schließen wir $\forall x \in D$ auf

$$\sum_{i=1}^{m} \tilde{\lambda}_i f_i(x) + \tilde{\nu}^T (Ax - b) \geq 0 \tag{2.4}$$

Wir wenden dies nun auf den Punkt $\tilde{x}$ an, welcher Slater erfüllt und haben damit:

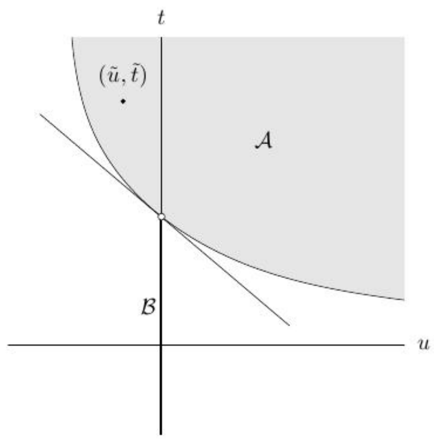
$$\sum_{i=1}^{m} \tilde{\lambda}_i f_i(\tilde{x}) \geq 0$$

Da $f_i(\tilde{x}) < 0$ und $\tilde{\lambda}_i \geq 0$ folgern wir daraus, dass $\tilde{\lambda} = 0$.
Weiterhin ist $(\tilde{\lambda}, \tilde{\nu}, \mu) \neq 0$ und $\tilde{\lambda} = 0, \mu = 0$ woraus $\tilde{\nu} \neq 0$ folgt.
Damit impliziert (2.4), dass $\forall x \in D$ gerade gelten muss, dass $\tilde{\nu}^T(Ax - b) \geq 0$

Aber: mit $\tilde{x}$ folgt $\tilde{\nu}^T(A\tilde{x} - b) = 0$ und $\tilde{x} \in \operatorname{int} D$ existieren Punkte in D mit $\tilde{\nu}^T(Ax - b) < 0$ (außer für $A^T\tilde{\nu} = 0$
Die widerspricht aber unserer Annahme, dass $\operatorname{rang} A = p$

Das heißt, dass also unser $\mu > 0$ sein muss.
□

Beispiel 2.1 *(Illustration des Beweises starker Dualität)*
Wir haben die zwei konvexen Mengen A und B, welche sich nicht schneiden. B ist in diesem Fall die dicke schwarze Linie, welche den Punkt $(0, p^*)$ nicht enthält.
Die beiden Mengen werden durch eine Hyperebene getrennt.



Slaters Bedingung garatiert uns, dass irgend eine trennende Hyperebene nicht-vertikal sein muss, wenn die Hyperebene links des Punktes $(\tilde{u}, \tilde{t}) = (f_1(\tilde{x}), f_0(\tilde{x}))$ verläuft, in welchem $\tilde{x}$ strikt zulässig ist.