Erős dualitás, Karush—Kuhn—Tucker-tétel

Hajnal Péter

2021. tavasz

Erős dualitás: Emlékeztető

Emlékeztető

A primál feladat optimális értékét p^*-gal, a feladat optimális értékét d^*-gal jelöljük ($d^*, p^* \in \mathbb{R} \cup \{-\infty, \infty\}$). Igaz a következő (**Gyenge-dualitás tétel**): $d^* \leq p^*$.

- Az erős dualitás tételről akkor beszélünk, ha $d^* = p^*$-t garantálni tudjuk bizonyos feltételek mellett.
- A „bizonyos” feltételekre sokféle lehetőség van.
- Egész „iparág” alakult ilyenek kifejlesztésére. Mi csak egy lehetőséget ismeretünk.

Slater-tétel

Slater-tétel

Adott a következő optimalizális probléma:

Minimalizáljuk	$c(x)$,-t	
Feltéve, hogy	$f_i(x) \leq 0$	$i = 1, \ldots, k$
	$g_i(x) = 0$	$i = 1, \ldots \ell$

Tegyük fel, hogy

(1) A probléma konvex. Tehát c és f_i konvex függvény, g_i pedig affin függvény. Ez azt jelenti, hogy a $g_i(x) = 0$ ($i = 1, \ldots, \ell$) feltételek a következő alakban írhatók: $Ax - b = 0$, ahol $A \in \mathbb{R}^{\ell \times n}$ és $b \in \mathbb{R}^{\ell}$.

(S) Létezik olyan $s \in \mathcal{D}$, amelyre

   (i) $f_i(s) < 0$ ($i = 1, \ldots, k$) és $g_i(s) = 0$ ($i = 1, \ldots, \ell$). Speciálisan $s \in \mathcal{L}$.

   (ii) Továbbá $s \in \text{int}\, \mathcal{D} = \{x : \ \exists r > 0 \ \ B(x, r) \subset \mathcal{D}\}$, $\mathcal{D}$ belső pontjainak halmaza, ahol $B(x, r)$ az x középpontú r sugarú gömb.

Ekkor erős dualitás van, azaz $d^* = p^*$.

A feltételekről

- (S)-et Slater-feltételnek nevezzük.

- Az (S)-beli feltételt kielégítő s pontokat Slater-pontoknak nevezzük.

- (S)(i) és (S)(ii) gyengíthetők. A tétel állítása igaz marad az alábbi (gyengített) feltételek mellett:

(S) (i)$_0$ Az s Slater-ponttól csak azt kívánjuk meg, hogy $f_i(s) < 0$, ha f_i nem affin, továbbá $f_i(s) \leq 0$, ha affin.

(S) (ii)$_0$ $s \in \text{relint } \mathcal{D} = \mathcal{D}$ belső pontjai $\mathcal{D}$ affin burkában (int $\mathcal{D} \subset$ relint $\mathcal{D}$).

- Az alábbiakban a bizonyítást írjuk le egy fontos feltevés mellett: A (az egyenlőség, affin feltételrendszer mátrixa) teljes sorrangú.

- A feltevés nélkül a bizonyítás lényegi része megmaradna, csupán néhány technikai bonyodalommal lenne hosszadalmasabb.

Bizonyítás: 1. észrevétel

1. Észrevétel

Feltehető, hogy
$$p^* \in \mathbb{R}.$$

- (S)-ből következik, hogy $\mathcal{L} \neq \emptyset$ amiből $p^* < \infty$.

- Továbbá a gyenge dualitásból következik az erős dualitás, ha $p^* = -\infty$. Így feltehető, hogy $p^* > -\infty$. Összegezve .

Bizonyítás: $\mathcal{E}$, $\mathcal{F}$ és észrevételek

- Legyen

$$\begin{aligned}\mathcal{E} = \{(\varphi_1, \varphi_2, \dots, \varphi_k, \gamma_1, \gamma_2, \dots, \gamma_\ell, \tau) \in \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R} :\\ \exists x \in \mathcal{D}, \text{ hogy}\\ \varphi_i \geq f_i(x) \ i = 1, \dots, k\\ \gamma_i = g_i(x) \ i = 1, \dots, \ell\\ \tau \geq c(x)\},\end{aligned}$$

$$\mathcal{F} = \{(0, 0, \dots, 0, \tau) \in \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R} : \ \tau < p^*\}.$$

2. Észrevétel

$\mathcal{E}$ és $\mathcal{F}$ konvex halmazok.

3. Észrevétel

$\mathcal{E}$ zárt a φ_i és τ koordináták növelésére.

Bizonyítás: Lemma

Lemma

$$\mathcal{E} \cap \mathcal{F} = \emptyset.$$

- Indirekt úton fogjuk bizonyítani.
- Tegyük fel, hogy $v \in \mathcal{E} \cap \mathcal{F}$, azaz $v \in \mathcal{E}$ és $v \in \mathcal{F}$.
- $v \in \mathcal{F}$ azt jelenti, hogy $v = (0, \ldots, 0, \tau)$, ahol $\tau < p^*$.
- $v \in \mathcal{E}$ azt jelenti, hogy van olyan $x \in \mathcal{D}$, amelyre $f_i(x) \leq 0$, $g_i(x) = 0$, továbbá $\tau \geq c(x)$.
- Azaz $x \in \mathcal{L}$, továbbá $c(x) \leq \tau < p^*$, ami ellentmondás.

Bizonyítás: Szeparációs tétel alkalmazása

Konvex halmazok szeperációs tétele ≈ Farkas lemma

K, L konvex és $K \cap L = \emptyset$ akkor létezik H hipersík, amely elválasztja a két halmazt.

Azaz úgy osztja fel a teret $H^{\leq}$ és $H^{\geq}$ zárt félsíkokra, hogy $H^{\leq} \supset K$ és $H^{\geq} \supset L$.

- A tételből és a 2. Lemmából adódik, hogy létezik $n = (\lambda_1, \lambda_2, \dots, \lambda_k, \mu_1, \dots, \mu_\ell, \nu)$ ($n \in \mathbb{R}^{k+\ell+1} = R^k \times \mathbb{R}^\ell \times \mathbb{R}$) nem-nulla vektor és α valós szám, hogy a $H_{n,\alpha} = \{x \in \mathbb{R}^{k+\ell+1}, n^\mathsf{T} x = \alpha\}$ hipersík két félterére:

$$H^{\geq}_{n,\alpha} = \{x \in \mathbb{R}^{k+\ell+1} : n^\mathsf{T} x \geq \alpha\} \supseteq \mathcal{E},$$

$$H^{\leq}_{n,\alpha} = \{x \in \mathbb{R}^{k+\ell+1} : n^\mathsf{T} x \leq \alpha\} \supseteq \mathcal{F}.$$

Bizonyítás: További észrevételek

• A 3. Észrevételből tudjuk, hogy az első k és az utolsó koordináta növelésével $\mathcal{E}$-ben és így $H^{\geq}$-ben is maradunk.

4. Észrevétel

$\lambda \succeq 0$ és $\nu \geq 0$.

• $(0, 0, p^* - \epsilon) \in \mathcal{F}$, amiből $(0, 0, p^* - \epsilon) \in H^{\leq}$, azaz $\nu(p^* - \epsilon) \leq \alpha$. Mivel $\epsilon > 0$ tetszőleges, határátmenettek kapjuk, hogy

5. Észrevétel

$\nu p^* \leq \alpha$.

• $x \in \mathcal{D}$ esetén $(f(x), g(x), c(x)) \in \mathcal{E}$, speciálisan $\in H^{\geq}$.

6. Észrevétel

Minden $x \in \mathcal{D}$ esetén:
$$\sum_{i=1}^{k} \lambda_i f_i(x) + \sum_{i=1}^{\ell} \mu_i g_i(x) + \nu c(x) \geq \alpha.$$

Bizonyítás: 1. eset: $\nu \neq 0$ ($\nu > 0$)

- Ekkor minden $x \in \mathcal{D}$-re

$$L\left(\frac{\lambda_i}{\nu}, \frac{\mu_i}{\nu}, x\right) = \sum_{i=1}^{k} \frac{\lambda_i}{\nu} f_i(x) + \frac{\mu_i}{\nu} g_i(x) + c(x) \geq \frac{\alpha}{\nu}.$$

- Adódik

$$\widetilde{c}\left(\frac{\lambda_i}{\nu}, \frac{\mu_i}{\nu}\right) \geq \frac{\alpha}{\nu} \geq p^*.$$

$\left(\frac{\lambda_i}{\nu}\right)_{i=1}^{k}$ a duális optimalizálási feladat lehetséges megoldásai.

- Ebből és az előző egyenlőtlenségből adódik, hogy

$$d^* \geq \widetilde{c}\left(\frac{\lambda_i}{\nu}, \frac{\mu_i}{\nu}\right) \geq p^*.$$

- A gyenge dualitással összevetve, kapjuk az erős dualitást.

Bizonyítás: 2. eset: $\nu = 0$

- Ekkor $\sum_{i=1}^{k} \lambda_i f_i(x) + \sum_{i=1}^{\ell} \mu_i g_i(x) \geq \alpha \geq \nu p^* = 0$ minden $x \in \mathcal{D}$ esetén.
- Írjuk fel az egyenlőtlenséget $x = s$ pontra, ahol s egy Slater-pont.

$$\sum_{i=1}^{k} \lambda_i f_i(s) + \sum_{i=1}^{\ell} \mu_i g_i(s) \geq 0, \quad \text{ahol} \quad \lambda_i \geq 0, f_i(s) < 0 \text{ és } g_i(s) = 0.$$

- Ekkor minden i-re λ_i-nek nullának kell, hogy legyen.
- Kiinduló egyenlőtlenségünk újra egyszerűsödik:

$$\sum_{i=1}^{\ell} \mu_i g_i(x) \geq 0,$$

ami átírva $\mu^{\mathsf{T}}(Ax - b) \geq 0$ minden $x \in \mathcal{D}$-re.

Bizonyítás: 2.eset (folytatás)

- Legyen $x = s + \delta$, ahol $\delta \in \mathbb{R}^{k+\ell+1}$ és $|\delta| < r_0$, ahol r_0 olyan kicsi, hogy $B(s, r_0) \subset \mathcal{D}$.

$$\begin{aligned}\mu^{\mathsf{T}}(A(s+\delta)-b) =& \mu^{\mathsf{T}}(As + A\delta - b) = \mu^{\mathsf{T}}(b + A\delta - b) \\ =& \mu^{\mathsf{T}}A\delta = \sum_{i=1}^{\ell}(\mu^{\mathsf{T}}A)_i\delta_i \geq 0.\end{aligned}$$

- Ez teljesül $-\delta$-ra is, amiből $\mu^{\mathsf{T}}A = 0$.
- Kezdeti feltételünk (A teljes sorrangú) miatt $\mu = 0$.
- Így $n = (\lambda, \mu, \nu) = 0$, ami ellentmondás.

Szünet

Laza feltételek, gyenge dualitás emlékeztető

Jelölés

Legyen $x \in \mathcal{L}$. Azt mondjuk, hogy x-ben az i-edik egyenlőtlenség feltétel laza, ha
$$f_i(x) < 0.$$

- Legyen x^* primál optimális hely és (λ^*, μ^*) duál optimális hely. Speciálisan $\lambda^* \succeq 0$.
- A gyenge duálitás bizonyítását összefoglaljuk.
- Legyen $L(\lambda, \mu, x) = c(x) + \lambda^{\mathsf{T}} f(x) + \mu^{\mathsf{T}} g(x)$.
- Ekkor

$$\begin{aligned} d^* =& \widetilde{c}(\lambda^*, \mu^*) = \inf L(\lambda^*, \mu^*, x) = \inf_{x \in \mathcal{D}} (c(x) + (\lambda^*)^{\mathsf{T}} f(x) + (\mu^*)^{\mathsf{T}} g(x)) \\ \leq& c(x^*) + (\lambda^*)^{\mathsf{T}} f(x^*) + (\mu^*)^{\mathsf{T}} g(x^*) \leq c(x^*) = p^*. \end{aligned}$$

- Amennyiben erős dualitás van, akkor végig egyenlőség teljesül.

Gyenge dualitás: Második egyenlőtlenség analízise

$$
\begin{aligned}
d^* =& \widetilde{c}(\lambda^*, \mu^*) = \inf L(\lambda^*, \mu^*, x) = \inf_{x \in \mathcal{D}} (c(x) + (\lambda^*)^{\mathsf{T}} f(x) + (\mu^*)^{\mathsf{T}} g(x)) \\
\leq& c(x^*) + (\lambda^*)^{\mathsf{T}} f(x^*) + (\mu^*)^{\mathsf{T}} g(x^*) \leq c(x^*) = p^*.
\end{aligned}
$$

Definíció

Legyen x_0 primál lehetséges megoldás, azaz $f_i(x_0) \leq 0$, $g_i(x_0) = 0$. Legyen (λ_0, μ_0) duál lehetséges megoldás, azaz $(\lambda_0)_i \geq 0$. Ezen megoldáspár rendelkezik a *komplemetáris lazasági tulajdonsággal*, ha

(i) $f_i(x_0) < 0$, akkor $(\lambda_0)_i = 0$.

(ii) $(\lambda_0)_i > 0$, akkor $f_i(x_0) = 0$.

Észrevétel

Az első egyenlőtlenségben akkor és csak akkor van egyenlőség, ha x^* és (λ^*, μ^*) komplementárisan laza tulajdonságú.

Gyenge dualitás: Első egyenlőtlenség analízise

$$d^* = \widetilde{c}(\lambda^*, \mu^*) = \inf L(\lambda^*, \mu^*, x) = \inf_{x\in\mathcal{D}} (c(x) + (\lambda^*)^{\mathsf{T}} f(x) + (\mu^*)^{\mathsf{T}} g(x))$$
$$\leq c(x^*) + (\lambda^*)^{\mathsf{T}} f(x^*) + (\mu^*)^{\mathsf{T}} g(x^*) \leq c(x^*) = p^*.$$

Észrevétel

Ha az első egyenlőtlenség egyenlőség, akkor $c(x) + (\lambda^*)^{\mathsf{T}} f(x) + (\mu^*)^{\mathsf{T}} g(x)$ függvényeknek x^* egy optimális helye.

Tegyük fel, hogy c és f_i függvények differenciálhatók. Ekkor

$$\nabla c(x^*) + (\lambda^*)^{\mathsf{T}} \nabla f(x^*) + (\mu^*)^{\mathsf{T}} \nabla g(x^*) = 0.$$

Tegyük fel, hogy c, f_i konvexek és g_i affin. Ekkor $c(x) + (\lambda^*)^{\mathsf{T}} f(x) + (\mu^*)^{\mathsf{T}} g(x)$ is konvex ($\lambda^* \succeq 0$). Ekkor a fenti feltétel szükséges és elégséges is a második egyenlőtlenség egyenlőségként való teljesüléséhez.

Karush—Kuhn—Tucher-tétel: Előzmények

- A tétel Karush MSc tézise volt a 30-as években.

- Később Kuhn és Tucker is felfedezik a tételt és ők tették ismertté az 50-es években.

Tegyük fel, hogy c, f_i, g_j differenciálhatóak. Továbbá c, f_i konvexek és g_j affin.

Definíció: Karash—Kuhn—Tucker-feltételek

$x^* \in \mathbb{R}^n$, $(\lambda^*, \mu^*) \in \mathbb{R}^k \times \mathbb{R}^\ell$ esetén

(KKT1) $f_i(x^*) \leq 0$ és $g_i(x^*) = 0$, azaz x primál lehetséges megoldás.

(KKT2) $\lambda_i^* \geq 0$, azaz (λ^*, μ^*) duál lehetséges megoldás.

(KKT3) x^* és (λ^*, μ^*) komplementárisan laza tulajdonságú.

(KKT4) $(\nabla c)(x^*) + (\lambda^*)^\mathsf{T} \nabla f(x^*) + (\mu^*)^\mathsf{T} \nabla g(x^*) = 0.$

Karush—Kuhn—Tucker-tétel: A Tétel

KKT tétel

Tegyük fel, hogy g_i affin függvény, c, f_i konvex és differenciálható függvények.

Ha van erős dualitás optimális helyekkel, akkor van olyan x_0 és (λ_0, μ_0), amelyek teljesítik a (KKT1), (KKT2), (KKT3), (KKT4) feltételeket.

Fordítva ha van x_0, (λ_0, μ_0), amelyek teljesítik a (KKT1), (KKT2), (KKT3), (KKT4) feltételeket akkor erős dualitás van és ezek primál és duál optimum helyek.

- A tétel első részét láttuk korábban.

Az elégség igazolása

$$\begin{aligned}
\widetilde{c}(\lambda_0, \mu_0) &= \inf(c(x) + \lambda_0^{\mathsf{T}} f(x) + \mu_0^{\mathsf{T}} g(x)) \\
&\underset{\text{(KKT4)}}{=} c(x_0) + \lambda_0^{\mathsf{T}} f(x_0) + \mu_0^{\mathsf{T}} g(x) \\
&\underset{\text{(KKT3)}}{=} c(x_0).
\end{aligned}$$

Hiszen KKT4 a differenciálhatósági és konvexitási feltétel miatt szükséges és elegendő feltétele annak, hogy x_0 optimum hely legyen.

Ezekután

$$d^* \underset{\text{(KKT2)}}{\geq} \widetilde{c}(\lambda_0, \mu_0) = c(x_0) \underset{\text{(KKT1)}}{\geq} p^* \underset{\text{gyenge dualitás}}{\geq} d^*.$$

Az egyenlőtlenség láncból kiolvasható, hogy végig egyenlőség teljesül, azaz erős dualitás van, x_0 primál optimál hely és (λ_0, μ_0) duál optimál hely.

KKT: Példa I

Minimalizáljuk	$\frac{1}{2}x^\mathsf{T}Px + q^\mathsf{T}x + r$-t
Feltéve, hogy	$Ax = b$,

ahol $P \in \mathcal{S}_+^n$.

• $c(x)$ konvex ($P \in \mathcal{S}_+^n$) és differenciálható, azaz KKT tétel alkalmazható.

• Olyan x_0, μ_0 értékeket kell keresnünk, amelyek mind a négy Karush—Kuhn-Tucker-feltételt teljesítik:
(KKT1): $Ax_0 = b$.
(KKT2): $\emptyset$.
(KKT3): $\emptyset$.
(KKT4): $\nabla c(x_0) + \mu_0^\mathsf{T}\nabla(Ax - b)|_{x=x_0} = 0$, azaz

$$Px_0 + q + A^\mathsf{T}\mu_0 = 0.$$

KKT: Példa I (folytatás)

- A keresett x_0, μ_0-k tulajdonságai összefoglalva:

$$\begin{pmatrix} P_{n\times n} & A^{\mathsf{T}}_{n\times k} \\ A_{k\times n} & 0 \end{pmatrix} \begin{pmatrix} x_0 \\ \mu_0 \end{pmatrix} = \begin{pmatrix} -q \\ b \end{pmatrix}.$$

- Ezen egyenletrendszer megoldhatóságának diszkussziója, illetve megoldhatóság esetén a megoldás megkeresése egyszerű lineáris algebrai feladat.

KKT: Példa II

Példa

Minimalizáljuk	$2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2$-t
Feltéve, hogy	$x_1^2 + x_2^2 \leq 5$,
	$3x_1 + x_2 \leq 6$.

- Esetünkben $\mathcal{D} = \mathbb{R}^2$.

- Könnyen ellenőrizhető, hogy a célfüggvény konvex, az egyenlőtlenség feltételek f_i függvényei is konvexek.

- Minden előforduló függvény differenciálható.

KKT: Példa II (folytatás)

- A KKT olyan primál/duál $x_1, x_2, \lambda_1, \lambda_2$ helyett keres, amelyek eleget tesznek a primál/duál feltételeknek ((KKT1) és (KKT2)):

$$x_1^2 + x_2^2 \leq 5, \quad 3x_1 + x_2 \leq 6, \quad \lambda_1 \geq 0, \quad \lambda_2 \geq 0.$$

Továbbá (KKT4) is teljesül.

- Ehhez:

$$\nabla(2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2) = \begin{pmatrix} 4x_1 + 2x_2 - 10 \\ 2x_1 + 2x_2 - 10 \end{pmatrix},$$

$$\nabla(x_1^2 + x_2^2 - 5) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}, \nabla(3x_1 + x_2 - 6) = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

- Azaz (KKT4) teljesülése kiírva:

$$4x_1 + 2x_2 - 10 + 2\lambda_1 x_1 + 3\lambda_2 = 0, \qquad 2x_1 + 2x_2 - 10 + 2\lambda_1 x_2 + \lambda_2 = 0.$$

KKT: Példa II (folytatás)

- Amit még tudnia kell számnégyesünknek, az a komplementáris hézagosság tulajdonsága.
- Ez négyféle módon teljesülhet:

$$I: \quad x_1^2 + x_2^2 = 5 \text{ és } \lambda_1 \geq 0, \quad 3x_1 + x_2 = 6 \text{ és } \lambda_2 \geq 0.$$

$$II: \quad x_1^2 + x_2^2 < 5 \text{ és } \lambda_1 = 0, \quad 3x_1 + x_2 < 6 \text{ és } \lambda_2 = 0.$$

$$III: \quad x_1^2 + x_2^2 < 5 \text{ és } \lambda_1 = 0, \quad 3x_1 + x_2 = 6 \text{ és } \lambda_2 \geq 0.$$

$$IV: \quad x_1^2 + x_2^2 = 5 \text{ és } \lambda_1 \geq 0, \quad 3x_1 + x_2 < 6 \text{ és } \lambda_2 = 0.$$

KKT: Példa II (folytatás)

• Elemi módszerekkel megállapítható, hogy I, II és III nem vezet megfelelő számnégyeshez.

• A IV lehetőség viszont, elvezet a

$$x_1 = 1, x_2 = 2, \lambda_1 = 1, \lambda_2 = 0$$

megoldáshoz.

• Ebből adódik,hogy $(1, 2)$ egy primál optimális megoldás, $(1, 0)$ egy duál optimális megoldás. Továbbá erős dualitás áll fenn.

Köszönöm a figyelmet!