Variációszámítás és optimális irányítás

Kánnai Zoltán, Szabó Imre, Tallos Péter[1]

[1]Készült a TÁMOP-4.1.2-08/2/A/KMR-2009-0045 program támogatásával

Előszó

Ez a kézikönyv matematikai bevezetésként kíván szolgálni a variációszámítás és az optimális irányítások elméletének tanulmányozásához, és elsősorban a közgazdasági mesterképzésben, illetve a doktori képzésben résztvevő hallgatók számára készült. Ennek megfelelően tárgyalja a legegyszerűbb lineáris irányítási rendszerek tulajdonságait, az alapvető variációs feladatokat és a nemlineáris rendszerekre vonatkozó Pontrjagin-féle maximumelvet.

A közgazdaságtan nemzetközi szakirodalmában számos olyan monográfia létezik, amely ezzel a területtel foglalkozik, a teljesség igénye nélkül utalunk *Aoki* illetve *Sydsaeter, Hammond, Seierstad*, vagy *Kamien, Schwartz* továbbá *Sethi, Thompson* munkáira (lásd az irodalomjegyzéket). Ezek a könyvek azonban a tételek megfogalmazására, azok interpretációjára és példákon keresztüli illusztrációira szorítkoznak. Nem adnak viszont képet az eredmények matematikai hátteréről. Számtalanszor szembesültünk azzal, hogy hallgatóink igénylik ennek tárgyalását is.

Igyekeztünk egy olyan könyvet az olvasó kezébe adni, amely rávilágít arra, hogy ezek a sokszor nagyon bonyolult dinamikus optimalizálási feladatok valójában roppant szemléletes és rendkívül egyszerű geometriai elveken nyugszanak. Ilyenek például a lineáris rendszerek esetében a legközelebbi pont elve és az ortogonalitási tétel Hilbert-terekben, illetve nemlineáris rendszerek esetében az érintősík fogalma és az ortogonalitási tétel normált terekben. Alapvető koncepciónk volt, hogy a dinamikus optimalizálás eredményeit ezen szemléletes geometriai elvek köré csoportosítsuk.

Úgy látjuk, hogy egy ilyen anyag hiánypótló lehet a közgazdaságtani szakirodalomban. Ezzel együtt bátorítjuk az olvasót az említett szakkönyvek nagyon igényes példaanyagának tanulmányozására is. A szükséges analízis ismeretek megtalálhatók *Kánnai* és *Magyarkuti* jegyzeteiben. Dinamikus optimalizálási feladatok egy szélesebb körének tárgyalását olvashatjuk *Kósa* könyveiben.

A könyv anyaga az évek során a Budapesti Corvinus Egyetem gazdaságmatematika szakán és a közgazdaságtani doktori iskola keretében tartott előadásaink rendszerezését tartalmazza. Nagyon bízunk azonban abban, hogy haszonnal forgathatják a természettudományi és műszaki képzésekben résztvevő hallgatók is.

Ezúton is szeretnénk köszönetet mondani hallgatóinknak, akik órai, vagy órán kívüli megjegyzéseikkel segítettek érthetőbbé tenni az anyagot. Hálásak vagyunk Varga Zoltánnak, akitől nagyon sok segítséget kaptunk munkánk során.

Budapest, 2013. június

Kánnai Zoltán, Szabó Imre, Tallos Péter

Tartalomjegyzék

I. rész

Globális optimalizálás: lineáris rendszerek

1. fejezet

Hilbert-terek geometriája

Ebben a fejezetben megismerkedünk a Hilbert-terek geometriájának alapjaival, a gyenge konvergenciával és egy igen általános globális optimalizálási elvvel. Ez az elv módszert is szolgáltat néhány érdekes optimalizálási feladat megoldására.

Az alábbiakban Hilbert-téren mindig a valós test fölötti teret értünk[1].

1.1. Variációs egyenlőtlenségek

A most következő absztrakt tétel a globális optimalizálási problémák talán legfontosabb alaptételének tekinthető.

1.1 Tétel. *Egy Hilbert-tér nem üres konvex, zárt részhalmazában egyetlen minimális normájú elem van.*

Bizonyítás. Legyen K konvex, zárt halmaz és

$$\alpha = \inf\{\|x\| : x \in K\}\,.$$

Válasszunk egy olyan K-beli x_n sorozatot, amelyre $\alpha = \lim \|x_n\|$. A parallelogramma-azonosság miatt

$$\begin{aligned}\left\|\frac{1}{2}(x_n - x_m)\right\|^2 &= \frac{1}{2}\left(\|x_n\|^2 + \|x_m\|^2\right) - \left\|\frac{1}{2}x_n + \frac{1}{2}x_m\right\|^2 \\ &\leq \frac{1}{2}\left(\|x_n\|^2 + \|x_m\|^2\right) - \alpha^2 \to 0\,,\end{aligned}$$

tehát x_n Cauchy-sorozat. Mivel a tér teljes és K zárt, azért x_n konvergens, és $x_n \to z \in K$. Mivel

$$0 \leq \big|\|x_n\| - \|z\|\big| \leq \|x_n - z\| \to 0\,,$$

[1] A Hilbert-terek elméletének alapjait illetően lásd *Kánnai: Analitikus módszerek a pénzügyben és a közgazdaságtanban, www.tankonyvtar.hu, 2013.*

világos, hogy $\|z\| = \alpha$. Másrészt, ha $y \in K$ is minimális normájú elem lenne, akkor ismét a parallelogramma-azonosság szerint

$$\left\|\frac{1}{2}(z-y)\right\|^2 = \alpha^2 - \left\|\frac{1}{2}z + \frac{1}{2}y\right\|^2 \leq 0\,,$$

azaz $z = y$. □

1.2 Tétel. *Legyen K a H Hilbert-tér konvex zárt részhalmaza, és $x \in H$. Akkor a K halmazban az x ponthoz egyetlen legközelebbi pont van.*

Bizonyítás. Alkalmazzuk az előző tételt a $K - x$ halmazra. □

A fenti állításokban csak az optimális elem létezését és egyértelműségét igazoltuk. A következőkben egy karakterizációt is megadunk. Az ilyen típusú állításokat variációs egyenlőtlenségeknek nevezzük.

1.3 Tétel. **(Tompaszög-tétel)** *Legyen K a H Hilbert-tér egy nem üres konvex, zárt részhalmaza, és $x \in H$. A $z \in K$ vektor akkor és csak akkor az x vektorhoz legközelebbi elem, ha*

$$\langle x - z, z - y\rangle \geq 0 \tag{1.1}$$

minden $y \in K$ esetén.

Bizonyítás. Legyen z az egyetlen legközelebbi elem. Bármely $y \in K$ és $0 \leq t \leq 1$ mellett $(1-t)z + ty \in K$. Tekintsük a

$$g(t) = \|x - (1-t)z - ty\|^2$$

másodfokú függvényt. Könnyen ellenőrizhető, hogy

$$g'(t) = 2\langle x - (1-t)z - ty, z - y\rangle\,,$$

illetve

$$g''(t) = 2\langle z - y, z - y\rangle = 2\|z - y\|^2\,.$$

Ha z legközelebbi elem, akkor $g'(0) \geq 0$, ami éppen az (1.1) egyenlőtlenség.

Fordítva, ha fennáll az (1.1) egyenlőtlenség valamely $z \in K$ vektorra, akkor egyrészt $g'(0) \geq 0$, másrészt $g''(t) \geq 0$ minden t-re. Ez azt jelenti, hogy $g(0) \leq g(1)$, azaz z valóban az x vektorhoz legközelebbi elem. □

A variációs egyenlőtlenség geometriai interpretációja igen szemléletes. Ezt úgy fogalmazhatjuk meg, hogy a z legközelebbi pontban az $x - z$ vektor bármely a K felé mutató $y - z$ vektorral tompaszöget zár be.

1.4 Definíció. Legyen K a H Hilbert-tér nem üres konvex, zárt részhalmaza. Tetszőleges $x \in H$ mellett jelentse $P_K(x)$ az x vektorhoz legközelebbi elemet K-ban. Ezt az x *projekciójának* nevezzük a K halmazra.

Zárt alterekhez (ilyenek például a véges dimenziós alterek) legközelebbi pont merőleges vetítéssel nyerhető, ezt fogalmazzuk meg az alábbi állításban.

1.5 Tétel. *Tekintsük a H Hilbert-tér egy L zárt alterét. Ekkor egy $x \in H$ vektor projekciója a következőképpen karakterizálható:*

$$\langle x - P_L(x), y \rangle = 0$$

minden $y \in L$ vektorra. Továbbá P_L lineáris leképezés.

Bizonyítás. Valóban, bármely $y \in L$ és valós t mellett $P_L(x) - ty \in L$, ennélfogva a

$$g(t) = \|x - P_L(x) + ty\|^2$$

függvénynek a $t = 0$ helyen minimuma van. Ezért $g'(0) = 0$, amiből adódik a karakterizáció.

Fordítva, ha valamely $z \in L$ vektorra $x - z$ merőleges az L minden elemére, akkor tetszőleges $y \in L$ vektorra

$$\|x - y\|^2 = \|x - z\|^2 + \|y - z\|^2 \,,$$

azaz z valóban az x-hez legközelebbi L-beli elem.

Másrészt nyilvánvaló, hogy $P_L(\alpha x) = \alpha P_L(x)$, így P_L homogén. Az additivitás abból adódik, hogy

$$\langle (x_1 + x_2) - P_L(x_1 + x_2), y \rangle = 0$$

minden $y \in L$ esetén, ha ez az x_1 és x_2 vektorokra egyaránt fennáll. Tehát $P_L(x_1 + x_2) = P_L(x_1) + P_L(x_2)$, azaz P_L valóban lineáris leképezés. □

1.6 Tétel. *Ha L a H Hilbert-tér zárt altere, akkor*

$$H = L \oplus L^{\perp} \,.$$

Bizonyítás. Valóban, bármely $x \in H$ felírható

$$x = P_L(x) + (x - P_L(x))$$

alakban, ahol $P_L(x) \in L$ és $x - P_L(x) \in L^{\perp}$. Ez a felbontás egyértelmű is az előző tételünk alapján. □

1.2. Riesz-reprezentáció

Ha $a \in H$ adott vektor a H Hilbert-térben, akkor a H-n értelmezett

$$f(x) = \langle x, a \rangle$$

lineáris függvény folytonos is, hiszen $|f(x) - f(y)| \leq \|a\| \cdot \|x - y\|$ a Cauchy-Schwarz-egyenlőtlenség szerint. Az alábbiakban megmutatjuk, hogy ez a példa tipikus, azaz egy Hilbert-tér minden folytonos lineáris funkcionálja megadható ilyen alakban.

1.7 Állítás. *Az f lineáris függvény a H Hilbert-téren akkor és csak akkor folytonos, ha van olyan α pozitív szám, amelyre*

$$|f(x)| \leq \alpha \cdot \|x\|$$

minden $x \in H$ vektorra.

Bizonyítás. Az elégségesség nyilvánvaló, ha ugyanis $x_n \to x$, akkor

$$|f(x_n) - f(x)| = |f(x_n - x)| \leq \alpha \cdot \|x_n - x\| \to 0\,.$$

A szükségességhez elég belátni, hogy

$$\alpha = \sup\{\frac{|f(x)|}{\|x\|} : x \neq 0\}$$

véges. Ha ugyanis ez végtelen lenne, akkor találhatnánk olyan x_n sorozatot, amelyre

$$y_n = \frac{x_n}{|f(x_n)|} \to 0\,,$$

jóllehet $|f(y_n)| = 1$ minden n mellett, ellentmondva a folytonosságnak. □

1.8 Tétel. **(Riesz-féle reprezentációs tétel)** *A H Hilbert-tér bármely f folytonos lineáris funkcionáljához egyértelműen található olyan $a \in H$ vektor, hogy*

$$f(x) = \langle x, a \rangle$$

minden $x \in H$ mellett.

Bizonyítás. Föltehető, hogy f nem azonosan nulla. Válasszunk egy y vektort, amelyre $f(y) \neq 0$. Jelölje z az y projekcióját a $\ker f$ zárt altérre, és legyen $b = y - z$. Ekkor b merőleges a $\ker f$ altérre, továbbá $f(b) = f(y)$. Bármely $x \in H$ vektor esetén

$$x - \frac{f(x)}{f(b)} b$$

a ker f altérben fekszik, és így merőleges a b vektorra, azaz

$$\langle x, b\rangle = \frac{f(x)}{f(b)}\langle b, b\rangle \,.$$

Ezt az

$$a = \frac{f(b)}{\|b\|^2} b$$

jelölés bevezetésével úgy is írhatjuk, hogy $f(x) = \langle x, a\rangle$ minden $x \in H$ esetén.

Az is nyilvánvaló, hogy ez az a vektor egyértelmű. Ha ugyanis b is ilyen tulajdonságú, akkor $\langle x, a-b\rangle = 0$ minden $x \in H$ vektorra, azaz $a = b$. □

1.9 Példa. Tekintsük az $L^2[0,T]$ Hilbert-teret. Riesz tétele értelmében e téren bármely folytonos lineáris funkcionál

$$\varphi \mapsto \int_0^T \varphi(t)\psi(t)\, dt$$

alakban adható meg, ahol $\psi \in L^2[0,T]$ alkalmas függvény.

Ellentétben a véges dimenziós esettel, egy végtelen dimenziós Hilbert-téren egy lineáris függvény nem feltétlenül folytonos.

1.3. Gyenge konvergencia

Egy végtelen dimenziós normált térben egy korlátos zárt halmaz nem feltétlenül kompakt. Ez azt is jelenti, hogy egy korlátos sorozatnak nem feltétlenül létezik konvergens részsorozata. Hilbert-terekben azonban megfogalmazhatunk Bolzano-Weierstrass-típusú tételt, ha a konvergencia fogalmát gyengítjük.

1.10 Definíció. Azt mondjuk, hogy a H Hilbert-térben az x_n sorozat gyengén tart x-hez, ha

$$\langle x_n, y\rangle \to \langle x, y\rangle$$

minden $y \in H$ mellett.

1.11 Állítás. *Ha $x_n \to x$ normában, akkor $x_n \to x$ gyengén is.*

Bizonyítás. Valóban, bármely $y \in H$ esetén

$$|\langle x_n, y\rangle - \langle x, y\rangle| = |\langle x_n - x, y\rangle| \le \|x_n - x\| \cdot \|y\| \to 0 \,.\Box$$

1.12 Tétel. *Egy Hilbert-térben bármely korlátos sorozatnak van gyengén konvergens részsorozata.*

Bizonyítás. Tekintsük a H Hilbert-tér valamely korlátos sorozatát, $\|x_n\| \leq \beta$. Jelölje M az x_n elemek által generált zárt alteret, és jelentse $M^\perp$ az ortogonális komplementerét. Mivel $\alpha_n = \langle x_1, x_n \rangle$ korlátos számsorozat, így kiválasztható belőle egy

$$\alpha_n^1 = \langle x_1, x_n^1 \rangle$$

konvergens részsorozat. Hasonlóan, $\langle x_2, x_n^1 \rangle$ is korlátos számsorozat, így ebből is kiválasztható egy

$$\alpha_n^2 = \langle x_2, x_n^2 \rangle$$

konvergens részsorozat. Az eljárást folytatva tekintsük az így keletkező x_n^n átlós sorozatot. Ekkor bármely k indexre $\langle x_k, x_n^n \rangle$ konvergens, hiszen $n > k$ mellett éppen az α_n^k konvergens sorozat részsorozatával van dolgunk. Megmutatjuk, hogy x_n^n gyengén konvergens.

Vezessük be az

$$f(x) = \lim_{n\to\infty} \langle x, x_n^n \rangle$$

függvényt, hacsak a limesz létezik. Ez a határérték nyilván létezik az x_n sorozat elemeiből alkotott véges lineáris kombinációkra, amelyek sűrű halmazt alkotnak az M altérben. Ebből következik, hogy f az egész M altéren értelmezhető, ha ugyanis $y \in M$ tetszőleges, akkor található az x_n lineáris kombinációinak olyan y_m sorozata, amelyre $y_m \to y$. Ezekre létezik az

$$f(y_m) = \lim_{n\to\infty} \langle y_m, x_n^n \rangle$$

határérték. Továbbá az

$$|\langle y_k - y_m, x_n^n \rangle| \leq \beta \|y_k - y_m\| \,, \tag{1.2}$$

egyenlőtlenség folytán $f(y_m)$ Cauchy-sorozat, tehát konvergens is, azaz $f(y_m) \to \alpha$ valamilyen α valós számra. Ez azt jelenti, hogy adott ε pozitív számhoz találhatók olyan m és N (az m-től függő) indexek, hogy bármely $n \geq N$ mellett

$$\begin{aligned} |\alpha - \langle y, x_n^n \rangle| &\leq |\alpha - f(y_m)| + |f(y_m) - \langle y_m, x_n^n \rangle| \\ &\quad + |\langle y_m, x_n^n \rangle - \langle y, x_n^n \rangle| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

az (1.2) reláció alapján. Ez pontosan azt jelenti, hogy $\langle y, x_n^n \rangle$ konvergens, következésképpen $f(y) = \alpha$.

Legyen ezenkívül f azonosan nulla az $M^\perp$ altéren, így nyilvánvalóan f értelmezve van az egész H téren. Továbbá f lineáris, és folytonos is, hiszen ha $y_m \to y$, akkor

$$|f(y_m - y)| = \lim_{n\to\infty} |\langle y_m - y, x_n^n \rangle| \leq \beta \|y_m - y\| \to 0 \,.$$

Riesz tétele folytán van olyan $a \in H$ vektor, amelyre $f(x) = \langle x, a \rangle$. Azonnal látható, hogy $x_n^n \to a$ gyengén. □

1.13 Példa. Tekintsünk egy olyan φ_n sorozatot az $L_X^2[0,T]$ térben, amelyre létezik olyan $\lambda \in L^2[0,T]$ valós függvény, hogy

$$\|\varphi_n(t)\| \leq \lambda(t)$$

minden n és majdnem minden $t \in [0,T]$ mellett. Ez a sorozat korlátos, így szükségképpen van gyengén konvergens részsorozta. Ez azt jelenti, hogy található olyan φ_{n_k} részsorozat, és olyan $\varphi \in L^2_X[0,T]$ függvény, hogy

$$\int_0^T \langle \varphi_{n_k}(t), \psi(t) \rangle \, dt \to \int_0^T \langle \varphi(t), \psi(t) \rangle \, dt$$

minden $\psi \in L^2_X[0,T]$ mellett.

A következő állításban azt látjuk, hogy a gyenge konvergenciából milyen további feltétel mellett következik a norma szerinti konvergencia.

1.14 Tétel. *Ha $x_n \to x$ gyengén és $\|x_n\| \to \|x\|$, akkor x_n tart x-hez normában is.*

Bizonyítás. Valóban,

$$\|x - x_n\|^2 = \|x\|^2 - 2\langle x, x_n \rangle + \|x_n\|^2 \to \|x\|^2 - 2\|x\|^2 + \|x\|^2 = 0\,,$$

hiszen $\langle x, x_n \rangle \to \|x\|^2$ a gyenge konvergencia miatt. □

1.4. Gyenge kompaktság

Ebben a szakaszban egy Hilbert-tér konvex zárt részhalmazairól látjuk be, hogy zártak a gyenge konvergenciára. Megvizsgáljuk a gyenge kompaktságot is.

1.15 Tétel. *Ha $x_n \to x$ gyengén a H Hilbert-térben, akkor*

$$x \in \operatorname{cl}\operatorname{co}\{x_1, x_2, \ldots\}\,,$$

azaz x a sorozat elemei által kifeszített konvex zárt halmazban fekszik.

Bizonyítás. Jelentse $K = \operatorname{cl}\operatorname{co}\{x_1, x_2, \ldots\}$ a sorozat elemeinek zárt konvex burkát. Indirekt módon tegyük fel, hogy $x \notin K$, és jelentse z az x projekcióját a K halmazra, azaz $z = P_K(x)$. Mivel $x_n - z$ is gyengén tart az $x - z$ vektorhoz, azért az 1.3 Tétel szerint

$$0 \geq \langle x_n - z, x - z \rangle \to \|x - z\|^2\,.$$

Ez azt jelenti, hogy $x = z$, azaz $x \in K$. □

1.16 Tétel. **(Mazur tétele)** *Ha $x_n \to x$ gyengén, akkor található az x_n elemek konvex kombinációiból álló olyan*

$$y_m = \sum_{j=1}^{k_m} \alpha_{m_j} x_{m_j}$$

sorozat, amelyre $y_m \to x$ normában.

Bizonyítás. Jelentse $C = \{x_1, x_2, \ldots\}$ a sorozat elemeinek halmazát. Mivel az előző állításunk szerint $x \in \operatorname{cl} \operatorname{co} C = K$, azért minden m természetes számra

$$x + \frac{1}{m} B \cap \operatorname{co} C \neq \emptyset\,.$$

Válasszunk tetszőlegesen egy y_m elemet a metszetből, az így nyert sorozat nyilván eleget tesz a tétel kívánalmának. □

1.17 Definíció. Az H Hilbert-tér valamely K részhalmazát *gyengén zártnak* nevezzük, ha bármely gyengén konvergens K-beli sorozat határértéke is a K halmazban fekszik.

Azt mondjuk továbbá, hogy K *gyengén kompakt*, ha bármely K- beli sorozatból kiválasztható a K-ban gyengén konvergens részsorozat.

Világos, hogy minden gyengén zárt halmaz egyúttal zárt is. Ennek megfordítása általában nem érvényes, tekintsük például azt az l^2-beli $K = \{x_1, x_2, \ldots\}$ halmazt, ahol az x_n sorozat elemeire $\xi_n^n = 1$, illetve $\xi_k^n = 0$, ha $k \neq n$. Ez a halmaz zárt, hiszen nincs is torlódási pontja, ugyanis $\|x_n - x_m\| = \sqrt{2}$ bármely két különböző indexre. Azonban nem gyengén zárt, hiszen könnyű ellenőrizni, hogy $x_n \to 0$ gyengén, de $0 \notin K$. (Lásd a 4. gyakorlatot.)

Az is világos, hogy a norma szerint kompakt halmazok egyúttal gyengén kompaktak is, hiszen bármely konvergens sorozat gyengén is konvergens. A megfordítás azonban már nem érvényes, tekintsük példaként az előző bekezdés x_n vektorainak halmazát a zérussal kiegészítve. Az így nyert halmaz gyengén kompakt (hiszen egy gyengén konvergens sorozat a limesszel kiegészítve), de normában nem lehet kompakt, ugyanis az eredeti sorozatnak nincs Cauchy-részsorozata.

1.18 Állítás. *Minden konvex zárt halmaz gyengén is zárt.*

Bizonyítás. Valóban, ha K konvex zárt halmaz és x_n a K elemeinek gyengén konvergens sorozata, akkor az 1.15 Tétel szerint a limesz is a K halmazban fekszik. □

Az 1.12 Tételt is figyelembe véve az alábbi következményt fogalmazhatjuk meg.

1.19 Következmény. *Egy Hilbert-tér bármely korlátos konvex zárt részhalmaza gyengén kompakt.*

1.20 Példa. Tekintsük a $[0, T]$ intervallumon olyan x_n abszolút folytonos függvények sorozatát, amelyekre $x_n(0) = x_0$, $x_n' \in L^2[0, T]$, és van olyan $\lambda \in L^2[0, T]$ függvény, amelyre

$$\|x_n'(t)\| \leq \lambda(t)$$

majdnem mindenütt minden n indexre. Ekkor található olyan x_{n_k} részsorozat és olyan x abszolút folytonos függvény a $[0,T]$ intervallumon, hogy $x_{n_k}(t) \to x(t)$ a $[0,T]$ intervallumon, továbbá

$$x'(t) \in \bigcap_{k=1}^{\infty} \operatorname{cl} \operatorname{co} \bigcup_{n=k}^{\infty} \{x'_n(t)\} . \tag{1.3}$$

Valóban, a fenti Következmény szerint kiválaszthatunk olyan x_{n_k} részsorozatot, amelyre x'_{n_k} gyengén konvergens $L^2[0,T]$-ben, nevezetesen $x'_{n_k} \to u$. Másrészt Mazur tétele szerint található e részsorozat elemeinek konvex kombinációiból álló y_m sorozat, amely normában tart u-hoz. Riesz tétele szerint ez utóbbi sorozatnak is van olyan részsorozata, amely majdnem mindenütt pontonként tart az u függvényhez. Ez azt jelenti, hogy minden k természetes számra

$$u(t) \in \operatorname{cl} \operatorname{co} \bigcup_{n=k}^{\infty} \{x'_n(t)\}$$

majdnem mindenütt. Vezessük be ezután az

$$x(t) = x_0 + \int_0^t u(s)\, ds$$

függvényt a $[0,T]$ intervallumon. Világos, hogy erre érvényes az (1.3) reláció. A pontonkénti konvergencia abból adódik, hogy

$$\|x(t) - x_{n_k}(t)\| \le \left\| \int_0^t (x'(s) - x'_{n_k}(s))\, ds \right\| \to 0$$

a gyenge konvergencia miatt.

Hilbert-téren egy folytonos függvény nem feltétlenül folytonos a gyenge konvergenciára nézve, így egy gyengén kompakt halmazon nem feltétlenül veszi fel a minimumát. Konvex függvényekről azonban többet állíthatunk. Legyen tehát H Hilbert-tér és tekintsünk egy $f : H \to \mathbb{R}$ függvényt.

1.21 Tétel. *Ha K a H korlátos konvex zárt részhalmaza, és f folytonos konvex függvény, akkor f felveszi a minimumát a K halmazon.*

Bizonyítás. Legyen $\alpha = \inf\{f(x) : x \in K\}$, és tetszőleges n természetes szám mellett válasszunk egy olyan $x_n \in K$ elemet amelyre

$$f(x_n) \le \alpha + \frac{1}{n} .$$

Az 1.19 Következmény szerint az x_n sorozatból kiválasztható egy x_{n_k} gyengén konvergens részsorozat, jelölje x_0 e sorozat határértékét. Minden n természetes szám esetén vezessük be a

$$K_n = K \cap \{x \in H : f(x) \le \alpha + \frac{1}{n}\}$$

halmazt. Mivel egy konvex függvény alsó nívóhalmazai konvexek, és f folytonos, azért a K_n halmazok mindegyike konvex, korlátos és zárt, tehát gyengén is zárt. Ez azt jelenti, hogy $x_0 \in K_n$, azaz

$$f(x_0) \leq \alpha + \frac{1}{n}$$

minden n indexre. Ez csak úgy lehetséges, hogy $f(x_0) = \alpha$. $\square$

1.5. Extremális pontok

Ebben a szakaszban a gyenge konvergencia egy alkalmazását mutatjuk be az extremális pontok létezésének igazolására. Ez az állításunk az úgynevezett Krein-Milman-tétel speciális esete.

1.22 Definíció. Legyen K egy Hilbert-tér konvex részhalmaza. Azt mondjuk, hogy $E \subset K$ a K *extremális részhalmaza*, ha $x, y \in K$ és $0 < \alpha < 1$ esetén az $\alpha x + (1-\alpha)y \in E$ relációból következik, hogy $x, y \in E$. Az egyelemű extremális részhalmazok elemeit *extremális pontoknak* nevezzük.

1.23 Lemma. *Ha K egy Hilbert-tér nem üres korlátos, konvex zárt részhalmaza, akkor minden n természetes számhoz van olyan E_n extremális részhalmaz, amelyre* $\operatorname{diam} E_n < 1/n$.

Bizonyítás. Legyen $\alpha = \sup\{\|x\| : x \in K\}$, és legyen adott $\varepsilon > 0$. Válasszunk egy olyan $x_0 \in K$ vektort, amelyre $\|x_0\| > \alpha - \varepsilon$, és tekintsük a

$$K_0 = \{x \in K : \langle x - x_0, x_0 \rangle \geq 0\}$$

halmazt. Az 1.21 Tétel szerint az $x \to \langle x - x_0, x_0 \rangle$ függvény felveszi a maximumát a K halmazon. Világos, hogy a maximumhelyek halmaza olyan zárt extremális részhalmaz, amely a K_0 halmazban fekszik.

Ezután becslést adunk a K_0 átmérőjére. Tetszőleges $x \in K_0$ esetén

$$\begin{aligned} \|x\|^2 &= \|x - x_0\|^2 + \|x_0\|^2 + 2\langle x - x_0, x_0 \rangle \\ &\geq \|x - x_0\|^2 + \|x_0\|^2 . \end{aligned}$$

Ebből következik, hogy

$$\|x - x_0\|^2 \leq \|x\|^2 - \|x_0\|^2 \leq \alpha^2 - (\alpha - \varepsilon)^2 \leq 2\alpha\varepsilon .$$

Tehát a K_0 halmaz x és y pontjaira

$$\|x - y\| \leq \|x - x_0\| + \|y - x_0\| \leq 2\sqrt{2\alpha\varepsilon} \to 0 ,$$

ha $\varepsilon \to 0$, és éppen ezt kellett bizonyítani. $\square$

1.24 Tétel. **(Krein-Milman-tétel)** *Ha K egy Hilbert-tér nem üres korlátos, konvex, zárt részhalmaza, akkor van extremális pontja.*

Bizonyítás. Tekintsük a K halmaz zárt extremális részhalmazainak egy monoton fogyó olyan E_n sorozatát, amelyekre $\operatorname{diam} E_n < 1/n$. Az előző lemma szerint ilyen sorozat választható. A Cantor-féle metszettétel szerint e sorozat metszete

$$\bigcap_{n=1}^{\infty} E_n = \{\bar{x}\}$$

egyetlen elemből áll. Könnyű ellenőrizni, hogy ekkor $\bar{x}$ a K halmaz extremális pontja. $\square$

1.6. Az ortogonalitási tétel

Legyen a továbbiakban H valós Hilbert-tér, és Λ a H téren értelmezett, valamely Y Hilbert-térre képező folytonos lineáris leképezés.

1.25 Tétel. *Ha Λ ráképezés, akkor $(ker \Lambda)^{\perp} = im \Lambda^*$.*

Bizonyítás. Ha $a \in \operatorname{im} \Lambda^*$, akkor van olyan $y \in Y$, amelyre $a = \Lambda^* y$. Ha most $v \in \ker \Lambda$ tetszőleges, akkor $\langle a, v\rangle = \langle \Lambda^* y, v\rangle = \langle y, \Lambda v\rangle = 0$.

Fordítva, ha $a \in (\ker \Lambda)^{\perp}$, akkor minden $v \in \ker \Lambda$ esetén $\langle a, v\rangle = 0$. Tekintsük az alábbi $g : Y \to \mathbb{R}$ függvényt:

$$g(y) = \langle a, \Lambda^{-1}(y)\rangle$$

Ekkor a g függvény jól definiált, hiszen bármely $x_1, x_2 \in \Lambda^{-1}(y)$ esetén $x_1 - x_2 \in \ker \Lambda$, és így $\langle a, x_1 - x_2\rangle = 0$. Másrészt könnyen látható, hogy g folytonos és lineáris, azaz $g \in Y^*$.

A Riesz-féle reprezentációs tétel szerint ezért található olyan $b \in Y$ vektor, amelyre $\langle b, y\rangle = \langle a, \Lambda^{-1}(y)\rangle$ minden $y \in Y$ mellett. Bármely adott $y \in Y$ vektorhoz a feltételünk szerint választható olyan $x \in X$ vektor, amelyre $\Lambda x = y$. Ekkor

$$\langle a, x\rangle = \langle b, \Lambda x\rangle = \langle \Lambda^* b, x\rangle$$

minden $x \in X$ esetén. Tehát $a = \Lambda^* b$, azaz $a \in \operatorname{im} A^*$. $\square$

A tételünk fontos folyománya, hogy ilyenkor $\operatorname{im} \Lambda^*$ zárt altér (ez a ráképezési feltétel nélkül általában nem igaz), hiszen az ortogonális komplementer mindig zárt alteret eredményez. Érdemes megjegyezni, hogy a

$$\ker \Lambda = (\operatorname{im} \Lambda^*)^{\perp}$$

reláció a ráképezési feltétel nélkül is érvényes. Ennek ellenőrzését az olvasóra bízzuk (lásd a 14. gyakorlatot).

Tekintsünk ezután a H és X Hilbert-tereket. Adott $\Lambda \in L(H, X)$ folytonos, az X térre képező lineáris leképezés és $\bar{x} \in X$ vektor mellett keressük a

$$K = \{u \in H : \Lambda u = \bar{x}\} \tag{1.4}$$

halmaz minimális normájú elemét.

1.26 Állítás. *Ha $u \in K$ megoldás, akkor $u \in \operatorname{im} \Lambda^*$.*

Bizonyítás. Ha $u \in K$ minimális normájú elem, akkor a tompaszög-tétel (1.3 Tétel) szerint u merőleges a $K = \Lambda^{-1}(\bar{x})$ affin halmazra, azaz a $\ker \Lambda$ altérre is. Ezért $u \in (\ker \Lambda)^{\perp} = \operatorname{im} \Lambda^*$. □

1.7. Egy optimális irányítási feladat

A fejezetünk eredményeit összefoglalva egy egyszerű optimális irányítási feladatra mutatunk példát ebben a szakaszban. Az ilyen jellegű feladatok tipikusak a makroökonómiában.

Legyen X euklideszi tér, amelyben adott az x_0 pont. Keressük az

$$F(x, u) = \int_0^T \left(\|x(t)\|^2 + \|u(t)\|^2\right) dt$$

integrál minimumát, ahol

$$x'(t) = u(t)\,, \quad x(0) = x_0\,, \tag{1.5}$$

továbbá

$$\|u(t)\| \leq \lambda \quad \text{majdnem mindenütt} \tag{1.6}$$

valamely λ pozitív szám mellett. A feladat úgy interpretálható, hogy az u irányítás segítségével kívánjuk minimalizálni a kvadratikus veszteség-függvényt olyan módon, hogy közben az irányítás energiája korlátozott.

Az x és u függvényeket egyaránt az $L^2_X[0, T]$ térből választjuk, így az F leképezés értelmezési tartománya a

$$H = L^2_X[0, T] \times L^2_X[0, T]$$

szorzattér egy részhalmaza, amelyen a skaláris szorzatot az

$$\langle (x_1, u_1), (x_2, u_2) \rangle = \int_0^T \left(\langle x_1(t), x_2(t)\rangle + \langle u_1(t), u_2(t)\rangle\right) dt$$

egyenlőséggel értelmezzük. Világos, hogy H így Hilbert-tér. Ekkor

$$F(x, u) = \|(x, u)\|^2$$

a H-beli skaláris szorzatra nézve.

Jelentse ezután K mindazon (x,u) párok halmazát a H Hilbert-térből, amelyek kielégítik az (1.5) és az (1.6) korlátozó feltételeket. A feladatunk absztrakt módon úgy fogalmazható meg, hogy keresendő a K halmazban minimális normájú elem. Az 1.1 Tétel szerint ilyen elem egyértelműen létezik, ha K konvex zárt halmaz.

Egyszerűen ellenőrizhetjük, hogy K konvex halmaz. A zártság belátásához tegyük fel, hogy (x_n, u_n) olyan K halmazbeli sorozat, amelyre $(x_n, u_n) \to (x,u) \in H$. Azt kell megmutatnunk, hogy $(x,u) \in K$. Vezessük be az

$$y(t) = x_0 + \int_0^t u(s)\,ds$$

jelölést a $[0,T]$ intervallumon. Mivel az u_n sorozatnak van olyan részsorozata, amely majdnem mindenütt tart az u függvényhez, azért (1.6) fennáll. Elég tehát igazolni, hogy $x = y$. A Cauchy-Schwarz egyenlőtlenséget alkalmazva az 1 és az $u - u_n$ függvényekre azt kapjuk, hogy

$$\|y(t) - x_n(t)\|^2 \le t \cdot \int_0^t \|u(s) - u_n(s)\|^2\,ds \le T \cdot \|u - u_n\|_2^2\,.$$

Tehát mindkét oldalt integrálva a $[0,T]$ intervallumon

$$\int_0^T \|y(t) - x_n(t)\|^2\,dt \le T^2 \cdot \|u - u_n\|_2^2\,,$$

majd mindkét oldalból négyzetgyököt vonva $\|y - x_n\| \le T \cdot \|u - u_n\|$ adódik. Ebből következik, hogy

$$\|y - x\|_2 \le \|y - x_n\|_2 + \|x_n - x\|_2 \le T \cdot \|u - u_n\|_2 + \|x - x_n\|_2\,.$$

Mivel a jobboldalon mindkét tag nullához tart, ha $n \to \infty$, innen valóban következik, hogy $x = y$.

1.8. Gyakorlatok

1. Adjunk példát egy Hilbert-téren olyan folytonos függvényre, amely az egységgömbön nem veszi fel a maximumát.

2. Keressünk példát normált térben olyan K konvex zárt halmazra, amelynek az origóhoz legközelebbi pontja nem egyértelmű. Keressünk olyat is, amelyben nincs is legközelebbi pont. (Útmutatás: Az utóbbi példához vizsgáljuk meg a $C[0,1]$ térben az

$$\left\{ x \in C[0,1] : \int_0^{1/2} x(t)\,dt - \int_{1/2}^1 x(t)\,dt = \alpha \right\}$$

hipersíkot.)

3. Mutassuk meg, hogy egy Hilbert-tér bármely nem nulla folytonos lineáris funkcionálja pontosan egy pontban veszi fel az abszolút értéke maximumát az egységgömbön. (Ez Banach-térben már nem feltétlenül érvényes, lásd az előző példát.)

4. Tekintsük az l^2 térben azt az x_n sorozatot, amelyre $\xi_n^n = 1$ és $\xi_k^n = 0$ minden $k \neq n$ esetén. Mutassuk meg, hogy $x_n \to 0$ gyengén, de persze x_n normában nem konvergens.

5. Bizonyítsuk be, hogy egy Hilbert-téren az $f(x) = \|x\|$ függvény az origó kivételével mindenütt differenciálható. Keressünk példát arra, hogy ez általában normált terekben (már véges dimenzióban sem) nem érvényes.

6. Adjunk példát Hilbert-térben olyan x_n sorozatra, amelyre $\|x_n\| \to \alpha$, $x_n \to x$ gyengén, de $\|x\| \neq \alpha$.

7. Adjunk példát Hilbert-térben olyan korlátos, konvex és zárt halmazra, amelynek az origótól nincs legtávolabbi pontja. (Útmutatás: Jelölje B az $L^2[0,1]$ zárt egységgömbjét, és tekintsük az $y(t) = t \cdot x(t)$ függvények halmazát, midőn $x \in B$.)

8. Ellenőrizzük, hogy alterek esetében a projekció fogalma egybeesik a merőleges vetítéssel. Nevezetesen, ha L a H Hilber-tér tér altere, akkor $x - P_L(x)$ merőleges az L altérre.

9. Bizonyítsuk be, hogy ha K a H Hilbert-tér konvex, zárt részhalmaza, akkor a $P_K(x)$ vektor éppen a
$$\langle P_K(x) - x, P_K(x) - y\rangle \leq 0 \quad \forall y \in K$$
variációs egyenlőtlenség egyetlen megoldása.

10. Mutassuk meg, hogy ha K a H Hilbert-tér konvex, zárt részhalmaza, akkor a P_K leképezés minden $x, y \in H$ mellett kielégíti a
$$\|P_K(x) - P_K(y)\| \leq \|x - y\|$$
egyenlőtlenséget.

11. Igazoljuk, hogy ha K a H Hilbert-tér konvex, zárt részhalmaza, akkor P_K úgynevezett *monoton leképezés*, azaz
$$\langle P_K(x) - P_K(y), x - y\rangle \geq 0$$
bármely $x, y \in H$ mellett.

12. Ha K a H Hilbert-tér konvex, zárt kúpja, akkor érvényesek az alábbi relációk:
$$\langle x - P_K(x), P_K(x)\rangle = 0\,,$$
továbbá
$$P_K(\lambda x) = \lambda P_K(x)$$
minden $x \in H$, illetve $\lambda > 0$ mellett.

13. Legyen K a H Hilbert-tér nem üres konvex, zárt részhalmaza, és tekintsük a
$$g(x) = \inf\{\|x - y\|^2 : y \in K\}$$
függvényt. Igazoljuk, hogy g folytonosan differenciálható, és
$$g'(x) = 2(x - P_K(x))$$
minden $x \in H$ pontban.

14. Mutassuk meg, hogy ha Λ folytonos lineáris leképezés valamely Hilbert-térről egy másik Hilbert-térbe, akkor
$$\ker \Lambda = (\operatorname{im} \Lambda^*)^\perp$$

2. fejezet

Lineáris rendszerek

Ebben a fejezetben lineáris rendszerek alapvető tulajdonságaival foglalkozunk.

2.1. Lineáris irányítási rendszerek

Legyenek X és Y n, illetve m dimenziós euklideszi terek és $x_0 \in X$ adott vektor. Tekintsük az $A : [0,T] \to L(X)$ és $B : [0,T] \to L(Y,X)$ négyzetesen integrálható függvényeket, ahol tehát $A(t)$ $n \times n$, $B(t)$ pedig $n \times m$ méretű mátrix. A négyzetesen integrálhatóságon azt értjük, hogy A és B mérhetőek, továbbá a $t \mapsto \|A(t)\|$, valamint a $t \mapsto \|B(t)\|$ függvények az $L^2[0,T]$ tér elemei.

Lineáris irányítási rendszeren az

$$\begin{aligned} x' &= A(t)x + B(t)u \\ x(0) &= x_0 \end{aligned}$$

lineáris rendszert értjük. Itt az $u \in L^2_Y[0,T]$ függvények az irányítások, az X tér a rendszer állapottere, Y az irányítási tér. Az u irányítás megválasztása a rendszer pályáját már egyértelműen meghatározza. A rendszer megoldásain Caratheodory-értelemben vett megoldásokat értünk.

A fenti kezdeti érték feladat megoldása a Cauchy-formula szerint

$$\varphi(t) = \Phi(t,0)\left(x_0 + \int_0^t \Phi(0,s)B(s)u(s)\,ds\right),$$

ahol Φ a homogén rendszer mátrix-megoldása, amelyre $\Phi(0,0) = E$ az $n \times n$-es egységmátrix.

Vezessük be a $\Lambda : L^2[0,T] \to X$,

$$\Lambda u = \int_0^T \Phi(0,s)B(s)u(s)\,ds$$

folytonos lineáris leképezést. Ezzel a jelöléssel

$$\varphi(T) = \Phi(T,0)\,(x_0 + \Lambda u)\ .$$

A továbbiakban szükségünk lesz a Λ leképezés Λ^* adjungáltjára[1].

2.1 Állítás. *A $\Lambda^* : X \to L^2[0,T]$ leképezésre bármely $x \in X$ és $t \in [0,T]$ mellett*

$$(\Lambda^* x)(t) = B(t)^* \Phi(0,t)^* x$$

Bizonyítás. Valóban, bármely $u \in L^2[0,T]$ esetén

$$\begin{aligned}
\langle \Lambda^* x, u\rangle &= \langle x, \Lambda u\rangle = \left\langle x, \int_0^T \Phi(0,s)B(s)u(s)\,ds \right\rangle \\
&= \int_0^T \langle x, \Phi(0,s)B(s)u(s)\rangle\,ds = \int_0^T \langle B(s)^*\Phi(0,s)^* x, u(s)\rangle\,ds \\
&= \langle B(.)^*\Phi(0,.)^* x, u\rangle,
\end{aligned}$$

amiből adódik az állítás. (Figyelem, az integrálok mögött X-beli skaláris szorzat áll!) □

2.2. Irányíthatóság

Tekintsük az alábbi irányítási rendszert

$$\begin{aligned}
x' &= A(t)x + B(t)u \\
x(0) &= x_0\,,
\end{aligned} \tag{2.1}$$

a $[0,T]$ időintervallumon.

2.2 Definíció. Azt mondjuk, hogy a (2.1) rendszer *irányítható* a $[0,T]$ intervallumon, ha bármely $x_0, x_T \in X$ állapotokhoz található olyan u irányítás, hogy a megfelelő φ trajektóriára $\varphi(0) = x_0$, és $\varphi(T) = x_T$, azaz

$$x_T = \Phi(T,0)\left(x_0 + \int_0^T \Phi(0,s)B(s)u(s)\,ds\right)\ .$$

teljesül.

[1] Az adjungált leképezések tulajdonságait illetően lásd *Kánnai: Analitikus módszerek a pénzügyben és a közgazdaságtanban, www.tankonyvtar.hu, 2013.*

A mátrix-megoldás tulajdonságai alapján az irányíthatósági feltétel a következő ekvivalens alakban írható fel:

$$x_T = \Phi(T,0)x_0 + \int_0^T \Phi(T,s)B(s)u(s)\,ds\,.$$

Ez úgy is írható, hogy

$$\Phi(0,T)x_T - x_0 = \int_0^T \Phi(0,s)B(s)u(s)\,ds = \Lambda u\,.$$

Világos, hogy a (2.1) rendszer pontosan akkor irányítható, ha a Λ leképezés képtere az egész X tér, azaz $\operatorname{im}\Lambda = X$.

2.3 Definíció. Az alábbi $\Lambda\Lambda^*$ $n \times n$ méretű mátrixot

$$\Lambda\Lambda^* = \int_0^T \Phi(0,t)B(t)B(t)^*\Phi(0,t)^*\,dt$$

a (2.1) rendszer *Gram-féle irányíthatósági mátrixának* nevezzük.

Könnyen látható, hogy $\Lambda\Lambda^*$ szimmetrikus, pozitív szemidefinit mátrix.

2.4 Tétel. $\operatorname{im}\Lambda = \operatorname{im}\Lambda\Lambda^*$.

Bizonyítás. Az világos, hogy $\operatorname{im}\Lambda \supset \operatorname{im}\Lambda\Lambda^*$.

Fordítva, először belátjuk, hogy $\ker\Lambda^* = \ker\Lambda\Lambda^*$. Valóban, egyrészt nyilvánvaló, hogy $\ker\Lambda^* \subset \ker\Lambda\Lambda^*$. Másrészt, ha $x \in \ker\Lambda\Lambda^*$, akkor

$$0 = \langle x, \Lambda\Lambda^* x\rangle = \langle \Lambda^* x, \Lambda^* x\rangle = \|\Lambda^* x\|^2,$$

tehát $\Lambda^* x = 0$, azaz $x \in \ker\Lambda^*$.

Innen az 1.25 Ortogonalitási tétel és a Gram-mátrix szimmetriája folytán

$$\operatorname{im}\Lambda \subset (\ker\Lambda^*)^\perp = (\ker\Lambda\Lambda^*)^\perp = \operatorname{im}\Lambda\Lambda^*\,,$$

ami a tételünket igazolja. □

Az előző tételünk nyilvánvaló következményeként adódik az alábbi irányíthatósági kritérium. A tétel jelentősége abban áll, hogy a végtelen dimenziós $L^2[0,T]$ Hilbert-téren értelmezett Λ leképezés képterének meghatározása helyett egy $n \times n$-es mátrix invertálhatóságát kell csak ellenőriznünk.

2.5 Tétel. *A (2.1) rendszer akkor és csak akkor irányítható, ha a $\Lambda\Lambda^*$ Gram-féle irányíthatósági mátrixa nemszinguláris.*

A tétel használatát a következő példán illusztráljuk.

2.6 Példa. Tekintsük azt az irányítási rendszert, ahol

$$A(t) = \begin{bmatrix} 0 & \alpha \\ -\alpha & 0 \end{bmatrix}, \quad B(t) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Könnyen látható, hogy

$$\Phi(t,0) = \begin{bmatrix} \cos\alpha t & \sin\alpha t \\ -\sin\alpha t & \cos\alpha t \end{bmatrix},$$

és így nem nehéz ellenőrizni, hogy

$$\Lambda\Lambda^* = \begin{bmatrix} \frac{T}{2} - \frac{\sin 2\alpha T}{4\alpha} & \frac{\cos 2\alpha T - 1}{4\alpha} \\ \frac{\cos 2\alpha T - 1}{4\alpha} & \frac{T}{2} + \frac{\sin 2\alpha T}{4\alpha} \end{bmatrix}.$$

Megmutatjuk, hogy ez a mátrix nemszinguláris. Valóban, a determinánsa

$$\frac{T^2}{4} - \frac{\sin^2 2\alpha T}{16\alpha^2} - \frac{\cos^2 2\alpha T - 2\cos 2\alpha T + 1}{16\alpha^2},$$

ami pontosan akkor nulla, ha

$$4\alpha^2 T^2 + 2\cos 2\alpha T - 2 = 0.$$

Ha bevezetjük a $\beta = 2\alpha T > 0$ helyettesítést, akkor a fenti egyenlet ekvivalens a

$$\cos\beta = 1 - \frac{\beta^2}{2}$$

egyenlettel. Ennek azonban nyilván nincs megoldása, hiszen a

$$g(\beta) = \cos\beta - 1 + \frac{\beta^2}{2}$$

függvény deriváltja $g'(\beta) = -\sin\beta + \beta > 0$ minden pozitív β mellett, tehát g szigorúan monoton növő. Másrészt $g(0) = 0$, ezért a g függvénynek nem lehet zérushelye a pozitív félegyenesen. Tehát $\Lambda\Lambda^*$ valóban nemszinguláris, és így a rendszer irányítható.

2.3. Autonóm rendszerek irányíthatósága

Ebben a szakaszban olyan irányítási rendszerekkel foglalkozunk, ahol A és B állandó mátrixok. Tekintsük tehát az

$$\begin{aligned} x' &= Ax + Bu \\ x(0) &= x_0 \end{aligned} \tag{2.2}$$

rendszert, ahol A $n \times n$, és B $n \times m$ méretű mátrixok.

Jól ismert, hogy autonóm rendszerek esetében a mátrix-megoldás invariáns az időeltolásra, ezért $\Phi(t, 0)$ helyett a rövidebb $\Phi(t)$ jelölést használjuk. Ilyenkor

$$\Phi(t) = e^{At}, \qquad \text{és} \qquad \Phi(t)^{-1} = \Phi(-t) = e^{-At},$$

és ezért

$$\Lambda u = \int_0^T \Phi(-t) B(t) u(t) \, dt \,.$$

Mivel a mátrixok az időtől függetlenek, egy erősebb irányíthatósági fogalmat vezetünk be.

2.7 Definíció. Azt mondjuk, hogy a (2.2) rendszer *teljesen irányítható*, ha bármely $x_0, x_T \in X$ állapotokhoz és bármely $T > 0$ időponthoz van olyan u irányítás a $[0, T]$ intervallumon, hogy a megfelelő kezdeti feltételű φ trajektóriára $\varphi(0) = x_0$, $\varphi(T) = x_T$.

2.8 Definíció. Tekintsük az alábbi $n \times nm$ méretű mátrixot:

$$K = \left[B, AB, A^2B, \ldots, A^{n-1}B\right] \,.$$

Ezt a mátrixot a (2.2) rendszer *Kalman-féle irányíthatósági* mátrixának[2] nevezzük.

2.9 Állítás. *Bármely $T > 0$ időpontra* $\operatorname{im} \Lambda\Lambda^* = \operatorname{im} KK^*$.

Bizonyítás. Mivel $\Lambda\Lambda^*$ és KK^* egyaránt szimmetrikus mátrixok, elegendő megmutatni, hogy $\ker \Lambda\Lambda^* = \ker KK^*$, hiszen a képterek és a magterek az egymás ortogonális komplementerei az X térben.

Először tegyük föl, hogy $v \in \ker \Lambda\Lambda^*$, ekkor

$$\begin{aligned} 0 &= \langle v, \Lambda\Lambda^* v\rangle = \int_0^T \langle v, \Phi(-t) BB^* \Phi(-t)^* v\rangle \, dt \\ &= \int_0^T \left\| B^* \Phi(-t)^* v \right\|^2 \, dt \,, \end{aligned}$$

ahol Φ az autonóm rendszer mátrixmegoldása, amelyre $\Phi'(t) = A\Phi(t)$ a számegyenesen. Innen az adódik, hogy

$$B^* \Phi(-t)^* v = 0$$

[2]Kálmán Rudolf magyar származású amerikai mérnök és matematikus, a modern rendszerelmélet egyik megalkotója.

minden $t \in [0,T]$ esetén. Tekintsük ezt az egyenletet, illetve $(n-1)$-edrendig bezárólag a deriváltjait a $t=0$ pontban, akkor a $\Phi(0)=E$ egyenlőségre tekintettel azt kapjuk, hogy

$$B^*v = B^*A^*v = \cdots = B^*(A^*)^{n-1}v = 0\,. \tag{2.3}$$

Ez éppen azt jelenti, hogy $K^*v=0$, és így persze $KK^*v=0$.

Fordítva, tegyük föl, hogy $v \in \ker KK^*$. Ekkor $K^*v=0$, hiszen

$$0 = \langle v, KK^*v\rangle = \langle K^*v, K^*v\rangle = \|K^*v\|^2\,,$$

és így fennállnak a (2.3) egyenlőségek. Mivel az A mátrix $(n-1)$-nél magasabb hatványai mind kifejezhetők az $E, A, A^2, \ldots, A^{n-1}$ mátrixok lineáris kombinációjaként,

$$B^*(A^*)^k v = 0\,, \quad k=0,1,2,\ldots$$

Ez azt jelenti, hogy

$$\sum_{k=0}^{\infty} B^* \frac{(A^*t)^k}{k!} v = B^* e^{A^*t} v = B^*\Phi(t)^* v = 0$$

minden t időpillanatban. Innen ugyanúgy, mint fent

$$\int_0^T \langle v, \Phi(-t)BB^*\Phi(-t)^*v\rangle\, dt = \langle v, \Lambda\Lambda^* v\rangle = 0\,.$$

Mivel $\Lambda\Lambda^*$ szimmetrikus pozitív szemidefinit mátrix, ez csak úgy lehetséges, hogy $\Lambda\Lambda^*v=0$, azaz $v \in \ker \Lambda\Lambda^*$. □

2.10 Tétel. (Kalman-féle irányíthatósági feltétel) *A (2.2) rendszer akkor és csak akkor teljesen irányítható, ha a Kalman-féle K mátrix rangja n.*

Bizonyítás. Egyrészt már láttuk, hogy $\ker KK^* = \ker K^*$, ezért ezen alterek ortogonális komplementerei is megegyeznek, azaz

$$\operatorname{im} KK^* = \operatorname{im} K\,.$$

Másrészt az előző állítás alapján bármely $T>0$ esetén

$$\operatorname{im} \Lambda\Lambda^* = \operatorname{im} KK^*\,,$$

tehát $\Lambda\Lambda^*$ pontosan akkor nemszinguláris, ha K rangja n. □

Érdemes megjegyezni, hogy a fentiek szerint ha egy autonóm rendszer irányítható valamely $[0,T]$ intervallumon, akkor teljesen irányítható.

2.11 Példa. Tekintsük újra a 2.6 Példát, ahol a rendszer irányíthatóságát a Gram-mátrixával igazoltuk. Vizsgáljuk most meg a Kalman-féle irányíthatósági mátrixot. Kétdimenziós rendszerről van szó, így esetünkben

$$K = [B, AB] = \begin{bmatrix} 0 & \alpha \\ 1 & 0 \end{bmatrix}.$$

Világos, hogy $\alpha \neq 0$ esetén K rangja 2, azaz a rendszer teljesen irányítható.

Autonóm rendszerekre természetesen a Kalman-féle feltétel ellenőrzése általában sokkal egyszerűbb, mint a Gram-mátrix vizsgálata.

2.4. Gyakorlatok

1. Vizsgáljuk meg a Kalman-feltétel alkalmazásával, hogy az

$$x' = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$

teljesen irányítható-e?

2. Legyen U az X normált tér olyan konvex zárt részhalmaza, amely szimmetrikus, azaz $U = -U$, továbbá $0 \in \operatorname{int} U$. Tekintsük az

$$\|x\| = \inf\{t > 0 : x \in tU\}$$

úgynevezett Minkowski-normát. Igazoljuk, hogy ez valóban normát definiál az X téren. Vajon hogyan karakterizálható az $u(t) \in U$ irányítás?

3. Tekintsük a következő autonóm rendszert:

$$x' = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix} x + b \cdot u\,.$$

Vajon milyen $b \in \mathbb{R}^3$ vektor esetén lesz a rendszer teljesen irányítható?

4. Döntsük el, hogy az

$$x' = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

autonóm lineáris rendszer teljesen irányítható-e? Keressünk olyan irányítást, amely $x_0 = (0, 0)$, $x_1 = (1, 1)$ mellett megoldása az

$$\inf\{\|u\| : \Lambda u = \bar{x}\}$$

optimális irányítási feladatnak a $[0, 1]$ intervallumon.

3. fejezet

A Pontrjagin-féle maximumelv

Ebben a fejezetben megfogalmazzuk a Pontrjagin-féle maximumelvet néhány lineáris rendszerekre vonatkozó optimalizálási feladatban. Tételeinket példákkal illusztráljuk.

3.1. A minimális norma feladat

Vizsgáljuk meg a következő optimális irányítási feladatot. Tekintsük a $[0, T]$ intervallumon a (2.1) alatti lineáris rendszert:

$$x'(t) = A(t)x(t) + B(t)u(t)\,, \tag{3.1}$$

ahol A és B, illetve az u irányítás kielégítik a (2.1) rendszer feltételeit.

Legyenek adottak az x_0 és x_T állapotok az X térben. Keresendő olyan u négyzetesen integrálható irányítás, hogy a rendszer $x(0) = x_0$ kezdeti feltételű x megoldására

$$x(T) = x_T\,, \tag{3.2}$$

és amelyre

$$\int_0^T \|u(t)\|^2\,dt \to \min, \tag{3.3}$$

azaz az integrál minimális.

3.1 Definíció. Azt mondjuk, hogy az (x, u) pár *megengedett folyamat*, ha u olyan irányítás, amelyre a rendszer $x(0) = x_0$ kezdeti feltételű x megoldása kielégíti a (3.2) feltételt.

Azt mondjuk továbbá, hogy (x, u) *optimális folyamat*, ha olyan megengedett folyamat, amelyre a (3.3) integrál minimális. Ilyenkor u-t *optimális irányításnak* nevezzük.

Az eddigi jelöléseinkkel élve a (3.2) egyenlőség azt jelenti, hogy

$$x_T = \Phi(T,0)\left(x_0 + \int_0^T \Phi(0,t)B(t)u(t)\,dt\right),$$

azaz

$$\Phi(0,T)x_T - x_0 = \int_0^T \Phi(0,t)B(t)u(t)\,dt = \Lambda u\,.$$

Ha bevezetjük az $\bar{x} = \Phi(0,T)x_T - x_0$ jelölést, akkor a fenti egyenlőség a $\Lambda u = \bar{x}$ alakba írható át.

3.2 Tétel. *Ha a (3.1) rendszer irányítható a $[0,T]$ intervallumon, akkor a fenti normaminimalizálási feladatnak pontosan egy megoldása van.*

Bizonyítás. Világos, hogy valamely $(x,u) \in W_X^2[0,T] \times L_Y^2[0,T]$ akkor és csak akkor megengedett folyamat, ha $\Lambda u = \bar{x}$. Tekintsük a következő halmazt:

$$K = \{v \in L_Y^2[0,T] : \Lambda v = \bar{x}\}\,.$$

Az irányíthatósági feltétel miatt egyrészt K nem üres, másrészt könnyen látható, hogy konvex és zárt halmaz. Az u irányítás pontosan akkor megoldása a feladatnak, ha a K halmaz minimális normájú eleme. Az 1.1 Tétel szerint pontosan egy ilyen elem létezik, és ez éppen az optimális irányítás. □

Ezután rátérünk az optimalitás szükséges feltételének megfogalmazására. A (3.1) lineáris rendszer adjungált rendszerén az

$$y'(t) = -A(t)^* y(t)$$

homogén lineáris rendszert értjük. Az adjungált rendszer mátrix-megoldása

$$\left(\Phi(t,0)^{-1}\right)^* = \Phi(0,t)^*$$

Jól ismert, hogy az adjungált rendszer trajektóriái ortogonálisak az eredeti homogén rendszer trajektóriáira[1].

3.3 Tétel. **Pontrjagin-féle maximumelv** *Ha (x,u) optimális folyamat, akkor*

- *található olyan $p \in X$, amelyre $u = \Lambda^* p$,*
- *az $y'(t) = -A(t)^* y(t)$ adjungált rendszernek létezik olyan ψ megoldása, hogy*
$$\langle \psi(t), B(t)u(t)\rangle = \max_{\|v\| \le \|u(t)\|} \langle \psi(t), B(t)v\rangle$$
m.m. $t \in [0,T]$ esetén. Ha itt $u \neq 0$, akkor $\psi \neq 0$.

[1] Az adjungált rendszerről bővebben lásd: *Tallos: Dinamikai rendszerek alapjai, Aula, 1999.*

Bizonyítás. Az első állítás közvetlenül adódik az 1.26 Állításból. Ekkor valamely $p \in X$ mellett $u = \Lambda^* p$, és ezért $\Lambda\Lambda^* p = \bar{x}$.

Ha valamely másik $q \in X$ mellett ugyancsak $\Lambda\Lambda^* q = \bar{x}$, akkor

$$p - q \in \ker \Lambda\Lambda^* = \ker \Lambda^* ,$$

ezért $\Lambda^* q = u$.

Térjünk rá a második állítás igazolására. Legyen a $[0, T]$ intervallumon

$$\psi(t) = \Phi(0, t)^* p .$$

Ekkor egyrészt ψ megoldása az adjungált rendszernek, másrészt az

$$u(t) = (\Lambda^* p)(t) = B(t)^* \Phi(0, t)^* p$$

egyenlőség folytán $u(t) = B(t)^* \psi(t)$, innen adódik, hogy $\psi \neq 0$, ha $u \neq 0$.

Legyen most $v \in Y$ tetszőleges vektor. Ekkor a $[0, T]$ intervallum m.m. t pontjában

$$\frac{1}{2}\left(\|v\|^2 + \|u(t)\|^2\right) \geq \|u(t)\| \cdot \|v\| \geq \langle u(t), v\rangle .$$

Mindkét oldalból $\|u(t)\|^2$-et kivonva

$$\begin{aligned} \frac{1}{2}\left(\|v\|^2 - \|u(t)\|^2\right) &\geq \langle u(t), v - u(t)\rangle = \langle B(t)^*\psi(t), v - u(t)\rangle = \\ &= \langle \psi(t), B(t)v - B(t)u(t)\rangle . \end{aligned}$$

Az egyenlőtlenséget átrendezve azt kapjuk, hogy

$$-\frac{1}{2}\|u(t)\|^2 + \langle \psi(t), B(t)u(t)\rangle \geq -\frac{1}{2}\|v\|^2 + \langle \psi(t), B(t)v\rangle .$$

De $\|v\| \leq \|u(t)\|$ esetén $1/2\|u(t)\|^2 \geq 1/2\|v\|^2$. E két utóbbi egyenlőtlenséget összeadva adódik a tétel állítása. □

A *maximumelv* elnevezés a tétel második állításában szereplő maximum-tulajdonságból ered. Ezt a tulajdonságot egy egyszerű formalizmussal még könnyebben megjegyezhetővé tehetjük.

3.4 Definíció. A (3.1), (3.2), (3.3) alatti lineáris optimális irányítási feladat *Hamilton-függvényén* a $[0, T] \times X \times Y \times X$ halmazon értelmezett

$$H(t, x, u, p) = -\frac{1}{2}\|u\|^2 + \langle p, A(t)x + B(t)u\rangle$$

függvényt értjük.

Ezzel a formalizmussal a maximumelv a következő egyszerű alakba írható át.

3.5 Tétel. **(Pontrjagin-féle maximumelv, Hamilton-formalizmus)** *Ha (x, u) optimális folyamat, akkor az*

$$y'(t) = -\partial_2 H(t, x(t), u(t), y(t))$$

adjungált rendszernek létezik olyan ψ megoldása, amelyre

$$H(t, x(t), u(t), \psi(t)) = \max_{v \in Y} H(t, x(t), v, \psi(t))$$

majdnem minden $t \in [0, T]$ pontban.

Tételünk szemléletesen azt jelenti, hogy a Hamilton-függvény a maximumát az optimális irányítás mentén veszi fel. Ez esetünkben azt jelenti, hogy

$$\partial_3 H(t, x(t), u(t), \psi(t)) = 0$$

a $[0, T]$ intervallum majdnem minden pontjában.

3.6 Példa. Oldjuk meg a $[0, 2]$ időintervallumon a következő normaminimalizálási feladatot:

$$\begin{aligned} x_1' &= \frac{\pi}{4} x_2 \\ x_2' &= -\frac{\pi}{4} x_1 + u \end{aligned}$$

Állítsuk elő az optimális irányítást, ha a kezdő és végállapotok

$$x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{és} \quad x_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

és amelyre az alábbi integrál minimális:

$$\int_0^2 |u(t)|^2 \, dt \to \min .$$

Esetünkben

$$A = \begin{bmatrix} 0 & \frac{\pi}{4} \\ -\frac{\pi}{4} & 0 \end{bmatrix} \quad \text{és} \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

és ezért

$$[B, AB] = \begin{bmatrix} 0 & \frac{\pi}{4} \\ 1 & 0 \end{bmatrix} ,$$

tehát a rendszer a Kalman-feltétel szerint teljesen irányítható. Másrészt erre az autonóm rendszerre

$$\Phi(t) = e^{At} = \begin{bmatrix} \cos \frac{\pi}{4} t & \sin \frac{\pi}{4} t \\ -\sin \frac{\pi}{4} t & \cos \frac{\pi}{4} t \end{bmatrix}$$

és innen

$$\Phi(t)^{-1} B = e^{-At} B = \begin{bmatrix} -\sin \frac{\pi}{4} t \\ \cos \frac{\pi}{4} t \end{bmatrix} .$$

Ebből egyszerű integrálással

$$\begin{aligned}\Lambda\Lambda^* &= \int_0^2 e^{-At}BB^*e^{-A^*t}\,dt \\ &= \int_0^2 = \begin{bmatrix} \sin^2\frac{\pi}{4}t & -\frac{1}{2}\sin\frac{\pi}{2}t \\ -\frac{1}{2}\sin\frac{\pi}{2}t & \cos^2\frac{\pi}{4}t \end{bmatrix} = \begin{bmatrix} 1 & -\frac{2}{\pi} \\ -\frac{2}{\pi} & 1 \end{bmatrix}.\end{aligned}$$

Könnyen leellenőrizhető, hogy

$$\bar{x} = e^{-2A}x_2 - x_0 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

ezért a $\Lambda\Lambda^* p = \bar{x}$ egyenlet az alábbi alakot ölti:

$$\begin{bmatrix} 1 & -\frac{2}{\pi} \\ -\frac{2}{\pi} & 1 \end{bmatrix} \cdot p = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.$$

Innen azonnal

$$p = -\frac{\pi}{\pi-2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Végül az optimális irányítás az

$$\begin{aligned}u(t) &= (\Lambda^* p)(t) = B^*e^{-A^*t}p = -\frac{\pi}{\pi-2}\left[-\sin\frac{\pi}{4}t, \cos\frac{\pi}{4}t\right] \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \frac{\pi}{\pi-2}\left(\sin\frac{\pi}{4}t - \cos\frac{\pi}{4}t\right)\end{aligned}$$

formulával adható meg ($t \in [0,2]$).

3.2. Az időoptimum-feladat

Most egy olyan feladatot vizsgálunk meg, amelyben a $[0,T]$ intervallum nem rögzített. Legyen U az Y térnek olyan adott konvex részhalmaza, amelyre $0 \in U$. Csak olyan irányításokat tekintünk megengedettnek, amelyek az alábbi halmazhoz tartoznak:

$$\hat{U} = \{u \in L^2_Y[0,T] : u(t) \in U \text{ m.m.}\}.$$

Legyen adott az $x_0 \in X$ kezdeti állapot, és tekintsük az

$$x'(t) = Ax(t) + Bu(t)$$

autonóm lineáris irányítási rendszert az

$$x(0) = x_0\,, \quad \text{és} \quad x(T) = 0$$

peremfeltételek mellett. Keresendő olyan $u \in \hat{U}$ megengedett irányítás, amelyre teljesülnek a fenti peremfeltételek, és amelyre a T időpont minimális. Ezt a problémát *időoptimum-feladatnak* nevezzük.

3.7 Definíció. Adott $t > 0$ időpillanat mellett jelölje E_t azon $x_0 \in X$ pontok összességét, amelyekre vannak olyan $x(0) = x_0$ tulajdonságú $x \in W^2_X[0, T]$ és $u \in \hat{U}$ függvények, hogy a $[0, t]$ intervallumon $x' = Ax + Bu$, és $x(t) = 0$. (Tehát E_t a rendszer azon kezdőállapotainak halmaza, amelyekből az origó t idő alatt elérhető.) Legyen továbbá

$$E^0_t = \bigcup_{0<s<t} E_s\,.$$

Az alábbi állítás teljesen nyilvánvaló a definíció és a $0 \in U$ feltétel alapján.

3.8 Állítás. *Bármely $0 < s < t$ időpontokra $E_s \subset E_t$.*

3.9 Állítás. *Bármely $t > 0$ időpontra az E_t és E^0_t halmazok konvexek.*

Bizonyítás. Ha $x, y \in E_t$, akkor jelölje u és v azokat az irányításokat, amelyek az x illetve y állapotokat az origóba irányítják t idő alatt. Ekkor tetszőleges $0 < \alpha < 1$ skalárra az $\alpha u + (1 - \alpha)v$ irányítás az $\alpha x + (1 - \alpha)y$ kezdőállapotot t idő alatt az origóba viszi. Tehát $\alpha x + (1 - \alpha)y \in E_t$.

Az E^0_t halmaz konvexitása abból adódik, hogy előáll monoton növekedő konvex halmazok egyesítéseként. □

3.10 Állítás. *Tegyük fel, hogy (x, u) megengedett folyamat a $[0, T]$ intervallumon, és legyen ψ az*

$$y'(t) = -A^*y(t)$$

adjungált rendszer megoldása. Akkor

$$\langle\psi(T), x(T)\rangle - \langle\psi(0), x(0)\rangle = \int_0^T \langle\psi(t), Bu(t)\rangle\, dt$$

Bizonyítás. Valóban, a szorzat differenciálása alapján

$$\begin{aligned}
\frac{d}{dt}\langle\psi(t), x(t)\rangle &= \langle\psi'(t), x(t)\rangle + \langle\psi(t), x'(t)\rangle \\
&= \langle -A^*\psi(t), x(t)\rangle + \langle\psi(t), Ax(t) + Bu(t)\rangle \\
&= \langle\psi(t), -Ax(t)\rangle + \langle\psi(t), Ax(t) + Bu(t)\rangle \\
&= \langle\psi(t), Bu(t)\rangle\,.
\end{aligned}$$

Innen a Newton-Leibniz-formula alapján azonnal adódik az állítás. □

Megfogalmazzuk a Pontrjagin-féle maximumelvet az időoptimum-feladatra.

3.11 Tétel. *Ha (x, u) optimális folyamat a $[0, T]$ intervallumon, akkor az*

$$y'(t) = -A^* y(t)$$

adjungált rendszernek található olyan $\psi \neq 0$ megoldása, amelyre

$$\langle \psi(t), Bu(t) \rangle = \max_{v \in U} \langle \psi(t), Bv \rangle \tag{3.4}$$

majdnem minden $t \in [0, T]$ pillanatban.

Bizonyítás. Az optimalitás folytán $x_0 \in E_T$, de minden $0 < t < T$ időpillanatra $x_0 \notin E_t$, és ennélfogva $x_0 \notin E_T^0$. Ekkor az x_0 pont és az E_T^0 konvex halmaz hipersíkkal szeparálhatók, azaz létezik olyan nem zérus $a \in X$ vektor, amelyre

$$\langle a, z - x_0 \rangle \geq 0$$

minden $z \in E_T^0$ pontra.

Ha $z \in E_T$ tetszőlegesen adott pont, akkor jelölje φ_z a rendszer azon trajektóriáját, amelyre $\varphi_z(0) = z$, és $\varphi_z(T) = 0$. Ekkor bármely $0 < t < T$ mellett $\varphi_z(T - t) \in E_t \subset E_T^0$, és ezért

$$\langle a, \varphi_z(T - t) - x_0 \rangle \geq 0\,.$$

Itt a $t \to T - 0$ bal oldali határértékre térve $\varphi_z(T - t) \to z$, és így az

$$\langle a, z - x_0 \rangle \geq 0 \tag{3.5}$$

egyenlőtlenség minden $z \in E_T$ pontban is érvényes.

Tegyük föl, hogy ψ kielégíti az alábbi kezdetiérték-feladatot:

$$\begin{aligned} \psi'(t) &= -A^* \psi(t) \\ \psi(0) &= a \end{aligned}$$

ekkor $\psi \neq 0$. Indirekt módon tegyük fel nem érvényes a (3.4) maximum-tulajdonság. Ez azt jelenti, hogy található olyan $v \in U$ vektor és olyan $D \subset [0, T]$ pozitív mértékű halmaz, hogy

$$\langle \psi(t), Bu(t) \rangle < \langle \psi(t), Bv \rangle \tag{3.6}$$

a $t \in D$ pontokban. Vezessük be a következő irányítást:

$$\hat{u}(t) = \begin{cases} u(t) & \text{ha } t \in [0, T] \setminus D \\ v & \text{ha } t \in D \end{cases},$$

Legyen ezután $\hat{x}$ az alábbi feladat megoldása a $[0, T]$ intervallumon:

$$\begin{aligned} \hat{x}'(t) &= A\hat{x}(t) + B\hat{u}(t) \\ \hat{x}(T) &= 0\,. \end{aligned}$$

Ekkor $(\hat{x}, \hat{u})$ olyan megengedett folyamat, amely a rendszert az $\hat{x}(0)$ állapotból az origóba viszi, azaz $\hat{x}(0) \in E_T$. Ezért a (3.5) relációra tekintettel

$$\langle a, \hat{x}(0) - x_0 \rangle \geq 0\,. \tag{3.7}$$

A 3.10 Állítás szerint egyrészt

$$\langle \psi(T), x(T) \rangle - \langle \psi(0), x(0) \rangle = \int_0^T \langle \psi(t), Bu(t) \rangle \, dt\,,$$

másrészt

$$\langle \psi(T), \hat{x}(T) \rangle - \langle \psi(0), \hat{x}(0) \rangle = \int_0^T \langle \psi(t), B\hat{u}(t) \rangle \, dt\,.$$

Vonjuk ki ez utóbbi egyenlőséget az előbbiből, és vegyük figyelembe, hogy $x(T) = \hat{x}(T) = 0$, továbbá $x(0) = x_0$, és $\psi(0) = a$, ekkor azt kapjuk, hogy

$$\langle a, \hat{x}(0) - x_0 \rangle = \int_0^T \langle \psi(t), Bu(t) \rangle \, dt - \int_0^T \langle \psi(t), B\hat{u}(t) \rangle \, dt\,.$$

Ennek az egyenlőtlenségnek a bal oldala a (3.7) reláció folytán nem negatív, így

$$\int_0^T \langle \psi(t), Bu(t) \rangle \, dt \geq \int_0^T \langle \psi(t), B\hat{u}(t) \rangle \, dt\,.$$

Ez az $\hat{u}$ irányítás definíciójára tekintettel azt jelenti, hogy

$$\int_D \langle \psi(t), Bu(t) \rangle \, dt \geq \int_D \langle \psi(t), B\hat{u}(t) \rangle \, dt = \int_D \langle \psi(t), Bv \rangle \, dt\,,$$

ami ellentmond a (3.6) indirekt feltevésünknek. Tehát valóban

$$\langle \psi(t), Bu(t) \rangle \geq \langle \psi(t), Bv \rangle$$

minden $v \in U$ és majdnem minden $t \in [0, T]$ esetén. □

Ha bevezetjük a

$$H(x, u, p) = \langle p, Ax + Bu \rangle$$

Hamilton-függvényt, akkor tételünk az alábbi módon fogalmazható meg.

3.12 Tétel. *Ha (x, u) optimális folyamat a $[0, T]$ intervallumon, akkor az*

$$y'(t) = -\partial_1 H(x(t), u(t), y(t))$$

adjungált rendszernek létezik olyan $\psi \neq 0$ megoldása, amelyre

$$H(x(t), u(t), \psi(t)) = \max_{v \in U} H(x(t), v, \psi(t))$$

majdnem minden $t \in [0, T]$ pillanatban.

3.3. Elégséges feltétel optimalitásra

Ebben a szakaszban megmutatjuk, hogy a Pontrjagin-féle maximumelv egy további feltevés mellett elégséges feltétel is az időoptimum feladatban. A továbbiakban

$$\operatorname{lin}\{v_1, \ldots, v_n\}$$

a $v_1, \ldots, v_n$ vektorok lineáris burkát jelenti.

3.13 Tétel. *Legyen (x, u) olyan megengedett folyamat a $[0, T]$ intervallumon, amelyre az*

$$y'(t) = -A^* y(t)$$

adjungált rendszernek létezik olyan $\psi \neq 0$ megoldása, hogy m.m. $t \in [0, T]$ mellett

$$\langle \psi(t), Bu(t) \rangle = \max_{v \in U} \langle \psi(t), Bv \rangle .$$

Tegyük fel továbbá, hogy található olyan $a \in Y$ vektor, amelyre $[-a, a] \subset U$, és

$$\psi(T) \in \operatorname{lin}\{Ba, ABa, \ldots, A^{n-1}Ba\} .$$

Akkor (x, u) optimális folyamat.

Bizonyítás. Indirekt módon tegyük fel, hogy található egy $0 < \hat{T} < T$ időpont, és $(\hat{x}, \hat{u})$ megengedett folyamat a $[0, \hat{T}]$ intervallumon. A maximum feltétel miatt

$$\langle \psi(t), Bu(t) \rangle \geq \langle \psi(t), B\hat{u}(t) \rangle$$

majdnem mindenütt a $[0, \hat{T}]$ intervallumon. Ekkor itt a 3.10 Állítás alapján

$$\begin{aligned}
\langle \psi(\hat{T}), x(\hat{T}) \rangle - \langle \psi(0), x(0) \rangle \quad & - \quad \Big(\langle \psi(\hat{T}), \hat{x}(\hat{T}) \rangle - \langle \psi(0), \hat{x}(0) \rangle \Big) = \\
& = \int_0^{\hat{T}} \langle \psi(t), Bu(t) \rangle \, dt - \int_0^{\hat{T}} \langle \psi(t), B\hat{u}(t) \rangle \, dt \\
& = \int_0^{\hat{T}} \left(\langle \psi(t), Bu(t) \rangle - \langle \psi(t), B\hat{u}(t) \rangle \right) dt \geq 0 .
\end{aligned}$$

Mivel itt $x(0) = \hat{x}(0) = x_0$, és $\hat{x}(\hat{T}) = 0$, ez azt jelenti, hogy

$$\langle \psi(\hat{T}), x(\hat{T}) \rangle \geq 0 . \tag{3.8}$$

Másrészt a $0 \in U$ feltétel miatt a $[0, T]$ intervallumon

$$\langle \psi(t), Bu(t) \rangle \geq \langle \psi(t), 0 \rangle = 0 .$$

Innen az $x(T) = 0$ egyenlőség és a 3.10 Állítás figyelembe vételével

$$\begin{aligned}
-\langle \psi(\hat{T}), x(\hat{T}) \rangle &= \langle \psi(T), x(T) \rangle - \langle \psi(\hat{T}), x(\hat{T}) \rangle \\
&= \int_{\hat{T}}^{T} \langle \psi(t), Bu(t) \rangle \, dt \geq 0 .
\end{aligned}$$

Ha ezt összevetjük a (3.8) egyenlőtlenséggel, akkor azt kapjuk, hogy

$$0 = \langle \psi(\hat{T}), x(\hat{T}) \rangle = \int_{\hat{T}}^{T} \langle \psi(t), Bu(t) \rangle \, dt \,.$$

De itt az integrandus nem negatív, ezért a maximumfeltételre való tekintettel

$$0 = \langle \psi(t), Bu(t) \rangle = \max_{v \in U} \langle \psi(t), Bv \rangle$$

majdnem minden $t \in [\hat{T}, T]$ esetén. Tehát ezen az intervallumon a $\langle \psi(t), . \rangle$ skaláris szorzat a BU halmazon a maximumát az origóban veszi fel.

A feltételünk szerint létezik olyan $a \in Y$ vektor, amelyre $[-Ba, Ba] \subset BU$, és

$$\psi(T) \in \mathrm{lin}\,\{Ba, ABa, \ldots, A^{n-1}Ba\} \,. \tag{3.9}$$

Mivel a $[\hat{T}, T]$ intervallum minden pontjában

$$\max_{v \in [-a,a]} \langle \psi(t), Bv \rangle = \langle \psi(t), 0 \rangle = 0 \,,$$

azért ezen az intervallumon $\langle \psi(t), Ba \rangle = 0$.

Azonban ψ folytonosan differenciálható, ezért

$$\frac{d}{dt} \langle \psi(t), Ba \rangle = \langle -A^* \psi(t), Ba \rangle = -\langle \psi(t), ABa \rangle = 0 \,,$$

és teljesen hasonlóan a magasabbrendű deriváltakra

$$\langle \psi(t), A^k Ba \rangle = 0 \quad k = 1, 2, \ldots$$

a $[\hat{T}, T]$ intervallumon. Speciálisan a T pontban

$$\langle \psi(T), A^k Ba \rangle = 0 \quad k = 1, 2, \ldots$$

Ezt a (3.9) feltétellel összevetve az adódik, hogy $\psi(T) = 0$, ami lehetetlen, hiszen így ψ azonosan nulla lenne. Ez azt jelenti, hogy (x, u) valóban optimális folyamat. □

3.14 Következmény. *Legyen (x, u) olyan megengedett folyamat a $[0, T]$ intervallumon, amelyre az*

$$y'(t) = -A^* y(t)$$

adjungált rendszernek létezik olyan $\psi \neq 0$ megoldása, hogy m.m. $t \in [0, T]$ mellett

$$\langle \psi(t), Bu(t) \rangle = \max_{v \in U} \langle \psi(t), Bv \rangle \,.$$

Tegyük fel továbbá, hogy található olyan $a \in Y$ vektor, amelyre $[-a, a] \subset U$, és

$$\mathrm{lin}\,\{Ba, ABa, \ldots, A^{n-1}Ba\} = X \,.$$

Akkor (x, u) optimális folyamat.

Érdemes megjegyezni, hogy a $[-a, a] \subset U$ feltétel biztosan teljesül, ha az origó az U halmaz belső pontja, azaz $0 \in \mathrm{int}\, U$.

3.4. Példa időoptimális irányításra

A következő példában a szükséges, illetve az elégséges feltételünk használatát illusztráljuk.

3.15 Példa. Legyen U a $[-1,1]$ intervallum, és tekintsük az alábbi autonóm irányítási rendszert

$$x'(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t)$$

ahol $x(0) = x_0$, és $x(T) = 0$ Keresendő olyan u irányítás, amelyre $u(t) \in [-1,1]$ m.m., és amely az x_0 kezdeti állapotot az origóba viszi minimális T idő alatt.

Vizsgáljuk először az adjungált rendszert:

$$y'(t) = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} y(t)\,,$$

ennek általános megoldása

$$\psi(t) = \begin{bmatrix} \alpha \\ -\alpha t + \beta \end{bmatrix},$$

ahol α és β tetszőleges valós számok. Ha ezt beírjuk a maximumfeltételbe, akkor

$$\left\langle \begin{bmatrix} \alpha \\ -\alpha t + \beta \end{bmatrix}, \begin{bmatrix} 0 \\ u(t) \end{bmatrix} \right\rangle = \max_{v\in[-1,1]} \left\langle \begin{bmatrix} \alpha \\ -\alpha t + \beta \end{bmatrix}, \begin{bmatrix} 0 \\ v \end{bmatrix} \right\rangle .$$

A szorzást elvégezve azt kapjuk, hogy

$$(-\alpha t + \beta)u(t) = \max_{v\in[-1,1]} (-\alpha t + \beta)v\,.$$

Világos, hogy a maximumhely $-\alpha t + \beta$ előjelétől függ, nevezetesen

$$u(t) = \operatorname{sgn}(\beta - \alpha t)\,,$$

azaz $u(t) \in \{-1, 1\}$ az egész intervallumon. A két esetet szétválasztjuk.

Ha $u(t) \equiv 1$, akkor a rendszerünk az alábbi alakot ölti

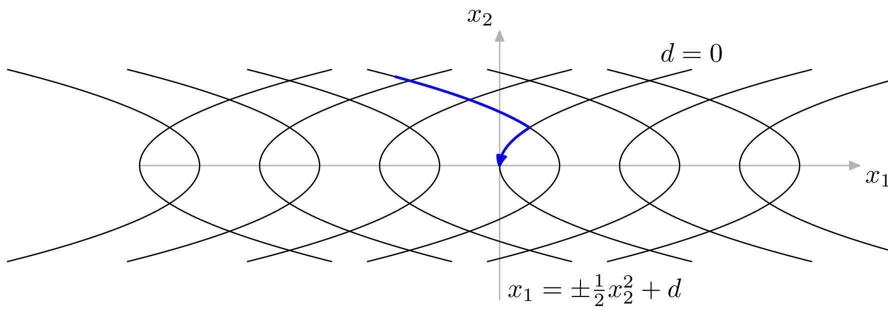
$$\begin{aligned} x_1' &= x_2 \\ x_2' &= 1\,, \end{aligned}$$

ennek általános megoldása $x_2(t) = t + c$ és $x_1(t) = 1/2(t + c)^2 + d$, azaz

$$x_1 = \frac{1}{2}x_2^2 + d$$

ahol d tetszőleges valós konstans. Ha pedig $u(t) \equiv -1$, akkor a rendszer

$$\begin{aligned} x_1' &= x_2 \\ x_2' &= -1\,, \end{aligned}$$



3.1. ábra. A rendszer fázisdiagramja

ennek általános megoldása $x_2(t) = c - t$ és $x_1(t) = 1/2(c-t)^2 + d$, tehát

$$x_1 = -\frac{1}{2}x_2^2 + d$$

ahol d tetszőleges valós állandó. Ezeket a görbéket a fázisdiagramon ábrázolva a 3.1 ábrához jutunk.

Ez azt mutatja, hogy egy teszőleges x_0 síkbeli kezdeti állapotból indulva először azon a parabolán haladunk, amelyen olyan parabolára juthatunk, amely átmegy az origón, az ábrán ekkor $u(t) \equiv 1$. A metszéspontban rátérünk arra a parabolára, amely az origóba visz, innentől $u(t) \equiv -1$.

Vizsgáljuk meg, hogy erre a folyamatra teljesül-e az elégséges feltételünk. Először is világos, hogy az origó az $U = [-1, 1]$ halmaz belső pontja, másrészt

$$\operatorname{im}[B1, AB1] = \operatorname{im}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \mathbb{R}^2 .$$

Tehát a fentiekben konstruált irányítás valóban optimális folyamatot definiál.

3.5. Gyakorlatok

1. Tekintsük a $[0, 1]$ intervallumon az

$$x'(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} u(t)$$

autonóm lineáris rendszert. Keressük meg azt az

$$\int_0^1 |u(t)|^2 \, dt \to \min$$

optimális irányítást, ahol

$$x_0 = \begin{bmatrix} -7 \\ 1 \end{bmatrix}, \quad \text{és} \quad x_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Mutassuk meg, hogy az optimális irányítás

$$u(t) = \begin{bmatrix} -6t+2 \\ 6 \end{bmatrix} \quad (t \in [0,1]).$$

2. Vizsgáljuk meg a $[0,1]$ intervallumon azt a normaminimalizálási feladatot, amelyben

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \quad \text{és} \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

és ahol

$$x_0 = \begin{bmatrix} -4 \\ -5+e^{-4} \end{bmatrix}, \quad \text{továbbá} \quad x_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Igazoljuk, hogy az optimális irányítás $u(t) = (24-16t)e^{-2t}$ a $[0,1]$ intervallumon.

3. Melyek azok az $x(0)$ vektorok, amelyekre az időoptimális irányítás konstans a $[0,T]$ intervallumon?

$$\begin{aligned} x'(t) &= \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t) \\ x(0) &= \begin{bmatrix} a \\ b \end{bmatrix}, \quad x(T) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned}$$

4. Keressük meg az időoptimális irányítást az alábbi feladatban:

$$x'(t) = \begin{bmatrix} 2 & 1 \\ -4 & -2 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ -3 \end{bmatrix} u(t),$$

ahol $U = [-1,1]$, továbbá

$$x(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad x(T) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Határozzuk meg T minimális értékét is.

4. fejezet

A bang-bang-elv

4.1. Bevezetés

Ebben a fejezetben a következő feladattal foglalkozunk. Legyenek X és Y n illetve m dimenziós euklideszi terek, $x_0 \in X$ adott, és tekintsük a (2.1) alatti

$$\begin{aligned} x'(t) &= A(t)x(t) + B(t)u(t) \\ x(0) &= x_0 \end{aligned}$$

lineáris irányítási rendszert a $[0, T]$ időintervallumon, ahol $A(t)$ $n \times n$ méretű, illetve $B(t)$ $n \times m$ méretű mátrixok, továbbá A és B négyzetesen integrálhatók.

Tegyük fel, hogy adott egy U nem üres, konvex, kompakt halmaz az Y térben, és a megengedett irányításokat korlátozzuk az

$$\hat{U} = \{u \in L^2_Y[0, T] : u(t) \in U \text{ m.m. }\} \tag{4.1}$$

halmazra.

4.1 Definíció. Egy u irányítást *megengedettnek* nevezünk, ha $u \in \hat{U}$.

4.2 Definíció. Azt mondjuk, hogy az x_T állapot elérhető az x_0 kezdőállapotból, ha található olyan u megengedett irányítás, hogy a rendszer megfelelő φ megoldására $\varphi(0) = x_0$, és $\varphi(T) = x_T$.

Kérdés, hogy egy elérhető x_T állapot vajon elérhető-e olyan irányításssal is, amely az értékeit az U extremális pontjaiban veszi fel. Az ilyen irányításokat *extremális* irányításoknak nevezzük. Ez különösen hasznos lehet olyan esetekben, amikor U poliéder, hiszen ekkor az extremális irányítások értékkészlete véges halmaz. Tehát a megengedett irányítások halmaza nagymértékben "ökonomizálható".

Az előző fejezetben bevezetett Λ leképezés segítségével a feladat formálisan úgy is felírható, hogy érvényes-e a

$$\Lambda U = \Lambda \operatorname{ex} \hat{U}$$

egyenlőség, ahol $\operatorname{ex} \hat{U}$ azon irányítások halmaza, amelyek értékkészlete az U extremális pontjainak részhalmaza.

Ez a fejezet nagymértékben támaszkodik a mértékelmélet eszköztárára[1].

4.2. Mérhető halmazértékű leképezések

Legyen X normált tér és tekintsünk egy $F : [0,T] \rightsquigarrow X$ zárt értékű halmazértékű leképezést. Ez azt jelenti, hogy minden $t \in [0,T]$ mellett $F(t)$ az X tér valamely nem üres, zárt részhalmaza.

Jelentse a továbbiakban $\mathcal{A}$ a $[0,T]$ intervallum Lebesgue-mérhető részhalmazainak σ-algebráját, továbbá jelölje μ a Lebesgue-mértéket.

4.3 Definíció. Azt mondjuk, hogy F *mérhető*, ha bármely $V \subset X$ nyílt halmazra

$$F^{-1}(V) = \{t \in [0,T] : F(t) \cap V \neq \emptyset\} \in \mathcal{A}\,,$$

azaz a nyílt halmazok inverzképei mérhetőek.

4.4 Állítás. *Ha X véges dimenziós, akkor az alábbi állítások ekvivalensek.*

(1) F mérhető.

(2) Bármely $K \subset X$ kompakt halmazra $F^{-1}(K) \in \mathcal{A}$.

(3) Bármely $M \subset X$ zárt halmazra $F^{-1}(M) \in \mathcal{A}$.

Bizonyítás. (1) $\Rightarrow$ (2). Ha $K \subset X$ kompakt halmaz, akkor tekintsük a

$$V_n = \{x \in X : d_K(x) < \frac{1}{n}\}$$

nyílt halmazokat. Világos, hogy $K = \cap_{n=1}^{\infty} V_n$, ezért

$$F^{-1}(K) = \bigcap_{n=1}^{\infty} F^{-1}(V_n) \in \mathcal{A}\,.$$

[1] A mértékelméleti segédeszközök megtalálhatók: *Magyarkuti: Mértékelmélet és dinamikus programozás, www.tankonyvtar.hu, 2013.*

(2) $\Rightarrow$ (3). Minden zárt halmaz megszámlálható sok kompakt halmaz egyesítéseként áll elő. Nevezetesen, ha M zárt, akkor a $K_n = M \cap nB$ halmazok kompaktak és

$$M = \bigcup_{n=1}^{\infty} K_n \, .$$

ahol B az X tér zárt egységgömbje. Ilyenkor tehát

$$F^{-1}(M) = \bigcup_{n=1}^{\infty} F^{-1}(K_n) \in \mathcal{A} \, .$$

(3) $\Rightarrow$ (1). Legyen most $V \subset X$ nyílt halmaz, és tekintsük az

$$M_n = \{x \in X : d_{V^c}(x) \geq \frac{1}{n}\}$$

zárt halmazokat. Mivel ekkor $V = \cup_{n=1}^{\infty} M_n$, azt kapjuk, hogy

$$F^{-1}(V) = \bigcup_{n=1}^{\infty} F^{-1}(M_n \cap nB) \in \mathcal{A} \, ,$$

amivel az állítást igazoltuk. $\Box$

Megjegyezzük, hogy (1) és (2) ekvivalenciája bármely normált térben érvényes.

4.5 Állítás. *Ha X véges dimenziós, továbbá F és G mérhető leképezések, akkor $F \cap G$ is mérhető.*

Bizonyítás. Először megmutatjuk, hogy $F \times G$ is mérhető leképezés. Valóban, ha $U \subset X \times X$ nyílt halmaz, akkor U előállítható

$$U = \bigcup_{n=1}^{\infty} (A_n \times B_n)$$

alakban, ahol A_n és B_n az X nyílt halmazai. Ekkor azonban

$$(F \times G)^{-1}(U) = \bigcup_{n=1}^{\infty} (F^{-1}(A_n) \cap G^{-1}(B_n)) \in \mathcal{A} \, ,$$

azaz $F \times G$ mérhető. Ha most $M \subset X$ zárt halmaz, akkor

$$(F \cap G)^{-1}(M) = \{t \in [0, T] : (F \times G)(t) \cap \Delta \cap (M \times M) \neq \emptyset\} \, ,$$

ahol $\Delta = \{(x, x) : x \in X\}$. Mivel $F \times G$ mérhető, azért ez utóbbi halmaz is mérhető, így az $F \cap G$ leképezés is mérhető. $\Box$

Könnyen ellenőrizhető a definíció alapján, hogy mérhető leképezések egyesítése is mérhető.

4.6 Példa. Tipikus példa mérhető halmazértékű leképezésre az alábbi konstrukció. Legyen u mérhető függvény a $[0, T]$ intervallumon, U az X kompakt részhalmaza, továbbá $F(t) = u(t) + U$. Adott $\varepsilon > 0$ mellett tekintsük a

$$G(t) = \{x \in F(t) : \|x - u(t)\| \geq \varepsilon\}$$

leképezést. Az előző állításunk alapján könnyű ellenőrizni, hogy G kompakt értékű mérhető leképezés. Valóban, G felírható a $G(t) = F(t) \cap H(t)$ alakban, ahol

$$H(t) = \{y \in X : \varepsilon \leq \|y\| \leq \alpha\},$$

és $\alpha = \sup\{\|u\| : u \in U\}$. Itt F és H mérhetőek, ugyanis Luzin tétele szerint a $[0, T]$ intervallumból elvehető egy tetszőlegesen kicsi pozitív mértékű rész, hogy a maradékon u folytonos. Ekkor nyílt halmaz inverzképe is nyílt, ebből adódik a teljes inverzkép mérhetősége. Igy G mérhetősége azonnal adódik az előző állításból.

4.3. Szelekciós tétel

Fontos kérdés, hogy egy mérhető halmazértékű leképezésben mikor halad egy mérhető függvény. Ezt válaszoljuk meg az alábbiakban. Igy arra is választ kapunk, hogy az irányítási rendszerek mikor írhatók fel differenciáltartalmazások alakjában. Az alábbi tétel (általánosabb formában) Kuratowski és Ryll-Nardzewski lengyel matematikusoktól származik.

4.7 Tétel. **(Szelekciós tétel)** *Ha X véges dimenziós és F zárt értékű mérhető leképezés, akkor F-nek van mérhető szelekciója, azaz olyan f mérhető függvény a $[0, T]$ intervallumon, amelyre*

$$f(t) \in F(t)$$

majdnem minden $t \in [0, T]$ esetén.

Bizonyítás. Jelölje $Q = \{r_1, r_2, \ldots\}$ a racionális koordinátájú pontok halmazát az X térben. Induktív módon definiáljuk mérhető függvények egy f_n sorozatát, amelyre

$$f_n(t) \in F(t) + \frac{1}{2^n} B, \tag{4.2}$$

valamint

$$\|f_n(t) - f_{n-1}(t)\| < \frac{1}{2^{n-1}} \tag{4.3}$$

minden t pontban és minden n természetes számra.

Legyen evégett $f_1(t) = r_1$ a $[0, T]$ intervallumon. Ha a sorozat első $n-1$ elemét már definiáltuk, úgy tekintsük a

$$C_i^n = \{t \in [0,T] : d_{F(t)}(r_i) < \frac{1}{2^n}\},$$

valamint a

$$D_i^n = \{t \in [0,T] : \|r_i - f_{n-1}(t)\| < \frac{1}{2^{n-1}}\}$$

halmazokat, és legyen $A_i^n = C_i^n \cap D_i^n$. Legyen $t \in [0, T]$ tetszőleges pont. Az indukciós feltevés miatt van olyan $x \in F(t)$, hogy $\|x - f_{n-1}(t)\| < 1/2^{n-1}$. Másrészt Q lezárása az egész X, ezért található olyan i index, amelyre $\|r_i - x\| < 1/2^n$, és $\|r_i - x\| + \|x - f_{n-1}(t)\| < 1/2^{n-1}$. Következésképpen $t \in A_i^n$. Így azt kaptuk, hogy

$$[0,T] = \bigcup_{i=1}^{\infty} A_i^n .$$

Vezessük be a

$$B_i^n = \{x \in X : \|x - r_i\| < \frac{1}{2^n}\}$$

jelölést. Ekkor B_i^n nyílt gömb, továbbá $C_i^n = F^{-1}(B_i^n)$, valamint $D_i^n = f_{n-1}^{-1}(B_i^{n-1})$. Ebből következik, hogy $C_i^n \in \mathcal{A}$ és $D_i^n \in \mathcal{A}$, tehát $A_i^n \in \mathcal{A}$. Vezessük be az $E_1^n = A_1^n$, illetve $i \geq 2$ esetén az

$$E_i^n = A_i^n \setminus \bigcup_{j=1}^{i-1} A_j^n$$

halmazokat. Világos, hogy ekkor az E_i^n halmazok páronként diszjunktak, mérhetőek, és egyesítésük kiadja a $[0, T]$ intervallumot. Értelmezzük az f_n függvényt az

$$f_n(t) = r_i \quad \text{ha} \quad t \in E_i^n$$

formulával. Könnyen látható, hogy f_n mérhető, továbbá a konstrukcióból adódóan fennállnak a (4.2) és (4.3) alatti egyenlőtlenségek.

Másrészt (4.3) szerint az f_n sorozat egyenletesen konvergens. Ha most $f = \lim f_n$, akkor f mérhető, továbbá a (4.2) relációra tekintettel az F halmazértékű leképezés szelekciója. □

A szelekciós tételt a későbbi szakaszokban extremális irányítások előállításához használjuk. Egy másik fontos alkalmazása az úgynevezett Filippov-féle implicitfüggvény-lemma, amely lehetővé teszi, hogy irányítási feladatokat differenciáltartalmazásokká írjunk át. Ennek egy speciális esetét mutatjuk most meg.

4.8 Tétel. **(Implicitfüggvény-lemma)** *Legyen $f : X \times Y \to X$ folytonos függvény, és $U \subset Y$ zárt halmaz. Tekintsük az*

$$F(x) = \{f(x,u) : u \in U\}$$

halmazértékű leképezést. Ha valamely $x : [0,T] \to X$ folytonos függvény mellett az $y : [0,T] \to X$ mérhető függvényre

$$y(t) \in F(x(t))$$

majdnem mindenütt, akkor található olyan $u : [0,T] \to U$ mérhető függvény, hogy

$$y(t) = f(x(t), u(t))$$

majdnem mindenütt.

Bizonyítás. Tekintsük a

$$G(t) = U \cap \{u \in Y : y(t) = f(x(t), u)\}$$

halmazértékű leképezést. Világos, hogy G értékei nem üres, zárt halmazok. Másrészt G mérhető is, hiszen a Luzin-tétel értelmében a $[0,T]$ intervallumból elvehető egy tetszőlegesen kicsi pozitív mértékű részhalmaz, hogy a maradékon y folytonos legyen. Tehát a szelekciós tétel szerint létezik egy u mérhető szelekciója. Ez éppen megfelel az állításunknak. □

Az implicitfüggvény-lemmából azonnal következik, hogy az

$$x'(t) = f(x(t), u(t)), \quad u \in \hat{U}$$

irányítási differenciálegyenletnek, illetve a fenti F mellett az

$$x'(t) \in F(x(t))$$

differenciáltartalmazásnak a megoldáshalmazai egybeesnek.

4.4. Extremális irányítások

A maximum-elv szerint az optimális irányítások értékkészlete az U irányítási tartomány határán fekszik. Ha történetesen U szigorúan konvex, akkor ezek a határpontok mind extremális pontok is.

4.9 Definíció. A (2.1) rendszer valamely u irányítását *extremálisnak* nevezzük, ha $u(t)$ majdnem mindenütt az U halmaz extremális pontja.

4.10 Lemma. *Az u pont akkor és csak akkor extremális pontja az U konvex halmaznak, ha az $U \cap (2u - U)$ halmaz az egyetlen u pontból áll.*

Bizonyítás. Mindenesetre, ha $u \in U$, akkor $u = 2u - u \in 2u - U$, és így $u \in U \cap (2u - U)$.

Tegyük fel, hogy $U \cap (2u - U)$ nem egyelemű, azaz található olyan $u_1 \neq u$, amelyre $u_1 \in U \cap (2u - U)$. Akkor egyrészt $u_1 \in U$, másrészt $u_2 = 2u - u_1 \in U$. Azonban U konvex, ezért $u = 1/2(u_1 + u_2)$, azaz u nem lehetne extremális pont.

Fordítva, tegyük fel, hogy u nem extremális pontja az U halmaznak. Ekkor találhatunk két különböző u_1 és u_2 pontot az U halmazból, valamint egy $0 < \alpha < 1$ skalárt, hogy $u = \alpha u_1 + (1 - \alpha)u_2$. Az általánosság csorbítása nélkül föltehető, hogy $\alpha > 1/2$. Ekkor

$$v = 2u - u_1 = (2\alpha - 1)u_1 + 2(1 - \alpha)u_2 \in U\,.$$

Könnyen látható, hogy ekkor u_1 és v egyaránt az $U \cap (2u - U)$ halmaz eleme, így az nem is lehet egyelemű. □

4.11 Tétel. *Tegyük fel, hogy U az Y konvex, kompakt részhalmaza, és tekintsük a megengedett irányítások (4.1) alatti $\hat{U}$ halmazát. Az u_0 irányítás akkor és csak akkor extremális, ha u_0 az $\hat{U}$ extremális pontja.*

Bizonyítás. Szükségesség. Indirekt módon tegyük fel, hogy u_0 nem extremális pontja az $\hat{U}$ halmaznak. Akkor található két különböző u_1 és u_2 megengedett irányítás, valamint egy $0 < \alpha < 1$ skalár, hogy

$$u_0 = \alpha u_1 + (1 - \alpha)u_2\,. \tag{4.4}$$

Mivel u_1 és u_2 különbözőek, van olyan pozitív mértékű D részhalmaza a $[0, T]$ intervallumnak, hogy $u_1(t) \neq u_2(t)$ minden $t \in D$ pontban. Ez azonban azt jelentené, hogy $u_0(t)$ nem lehet az U extremális eleme a D pontjaiban.

Elégségesség. Tegyük fel ismét indirekt módon, hogy létezik olyan pozitív mértékű D halmaz, hogy $u_0(t)$ nem extremális pontja az U halmaznak a D pontjaiban. Ekkor az előző lemma szerint a $W(t) = U \cap (2u_0(t) - U)$ halmaz nem is lehet egyelemű, ha $t \in D$. Találhatunk tehát olyan ε pozitív számot, és olyan $D_1 \subset D$ pozitív mértékű halmazt, amelyekre a

$$G(t) = \{x \in W(t) : \|x - u_0(t)\| \geq \varepsilon\} \neq \emptyset$$

a D_1 pontjaiban. A G leképezés mérhető (lásd a 4.6 Példát) és értékei zárt halmazok. A szelekciós tétel szerint ezért találhatunk egy olyan $\bar{u}$ mérhető függvényt, amelyre $\bar{u}(t) \in G(t)$ a D_1 halmazon. Világos a G értelmezéséből, hogy $2u_0(t) - \bar{u}(t) \in G(t)$ a D_1 pontjaiban. Tehát az

$$u_1(t) = \begin{cases} u_0(t) & \text{ha } t \notin D_1 \\ \bar{u}(t) & \text{ha } t \in D_1 \end{cases}$$

valamint az

$$u_2(t) = 2u_0(t) - u_1(t)$$

függvények egyaránt az $\hat{U}$ halmazban fekszenek. Mivel ekkor $u_0 = 1/2(u_1 + u_2)$, ez azt jelenti, hogy u_0 nem lehet az $\hat{U}$ extremális pontja. □

4.12 Lemma. *Ha H vektortér, $\Lambda : H \to X$ lineáris leképezés, $\hat{U} \subset H$ konvex halmaz és $\bar{x}$ a $\Lambda\hat{U}$ halmaz extremális pontja, akkor $\hat{U} \cap \Lambda^{-1}(\bar{x})$ az $\hat{U}$ extremális részhalmaza.*

Bizonyítás. Ha valamely $u_0 \in \hat{U} \cap \Lambda^{-1}(\bar{x})$ pontnak létezik olyan

$$u_0 = \alpha u_1 + (1-\alpha)u_2$$

előállítása, ahol $u_1, u_2 \in \hat{U}$ és $0 < \alpha < 1$, akkor

$$\bar{x} = \Lambda u_0 = \alpha \Lambda u_1 + (1-\alpha)\Lambda u_2 \,.$$

Mivel $\bar{x}$ extremális pont, ez csak úgy lehetséges, hogy $\Lambda u_1 = \Lambda u_2 = \bar{x}$. Ez azt jelenti, hogy $u_1, u_2 \in \hat{U} \cap \Lambda^{-1}(\bar{x})$. □

4.13 Tétel. *Legyen U az Y konvex kompakt részhalmaza, és tekintsük a megengedett irányítások (4.1) alatti $\hat{U}$ halmazát. Ha $\bar{x}$ a $\Lambda\hat{U}$ extremális pontja, akkor van olyan u_0 extremális irányítás, amelyre $\Lambda u_0 = \bar{x}$.*

Bizonyítás. Az előző lemma szerint mindenesetre $\hat{U} \cap \Lambda^{-1}(\bar{x})$ az $\hat{U}$ extremális részhalmaza. Másrészt $\hat{U} \cap \Lambda^{-1}(\bar{x})$ nyilván korlátos konvex zárt halmaz, ezért a 1.24 Tétel értelmében van extremális pontja. Mivel egy extremális részhalmaz extremális pontjai egyúttal az egész halmaz extremális pontjai is, találhatunk olyan $u_0 \in \hat{U}$ extremális pontot, amelyre $\Lambda u_0 = \bar{x}$. Ekkor azonban a 4.11 Tétel folytán u_0 extremális irányítás. □

4.5. Bang-bang elv

A következőkben azt fogalmazzuk meg, hogy ha egy lineáris rendszer valamely állapotba irányítható, akkor oda extremális irányítással is irányítható. Ezt az eredményt az irodalomban *bang-bang elvnek* nevezik, arra utalva, hogy az extremális irányítás az extremális pontok között ugrál.

4.14 Állítás. *Legyen $Y = \mathbb{R}$, $U = [0,1]$, és készítsük el a (4.1) alatti $\hat{U}$ halmazt. Tetszőleges $\psi_1, \dots, \psi_n \in L^2[0,T]$ függvények mellett az $L : \hat{U} \to \mathbb{R}^n$,*

$$Lu = \int_0^T (\psi_1(t), \dots, \psi_n(t))u(t)\, dt$$

leképezésre érvényes az

$$L\hat{U} = L(\mathrm{ex}\,\hat{U})$$

egyenlőség, ahol $\mathrm{ex}\,\hat{U}$ az extremális pontok halmaza.

Bizonyítás. Legyen $u_0 = 1/2$ a $[0,T]$ intervallumon, akkor $\hat{U} - u_0$ éppen az $1/2$ sugarú zárt gömb az $L^2[0,T]$ térben. Másrészt bármely $y \in L\hat{U}$ esetén $L^{-1}(y)$ zárt, ezért a 1.24 Tétel szerint az $\hat{U} \cap L^{-1}(y)$ halmaznak van egy u extremális pontja.

Megmutatjuk, hogy ekkor u az $\hat{U}$ extremális pontja is. Indirekt módon tegyük fel, hogy ez nem igaz, indukcióval belátjuk, hogy akkor u nem lehet az $\hat{U} \cap L^{-1}(y)$ extremális pontja sem.

Mindenesetre van olyan $\varepsilon > 0$ és olyan D pozitív mértékű halmaz, hogy a D pontjaiban

$$\varepsilon < u(t) < 1 - \varepsilon \,.$$

Tegyük fel, hogy az állításunk $(n-1)$-ig érvényes. Válasszunk két diszjunkt pozitív mértékű D_1 és D_2 halmazt a D-ben, és alkalmazzuk az indukciós feltevést a $\psi_1\chi_{D_1}, \ldots, \psi_{n-1}\chi_{D_1}$, illetve a $\psi_1\chi_{D_2}, \ldots, \psi_{n-1}\chi_{D_2}$ függvényekre. Ekkor találhatunk olyan $E_1 \subset D_1$ és $E_2 \subset D_2$ pozitív mértékű halmazokat, hogy

$$\int_0^T \psi_i(t)\chi_{E_j}(t)\,dt = \frac{1}{2}\int_0^T \psi_i(t)\chi_{D_j}(t)\,dt$$

az $i = 1, \ldots, n-1$, $j = 1, 2$ indexekre. Vezessük be az $\omega_j = 2\chi_{E_j} - \chi_{D_j}$ függvényeket, akkor

$$\int_0^T \psi_i(t)\omega_j(t)\,dt = 0$$

minden i, j mellett. Világos, hogy az ω_1 és ω_2 függvények szorzata nulla.

Ha most $\psi_n \in L^2[0,T]$ adott, akkor válasszunk olyan α és β nem zérus skalárokat, amelyekre $|\alpha| < \varepsilon$, $|\beta| < \varepsilon$, továbbá az $\omega = \alpha\omega_1 + \beta\omega_2$ függvényre

$$\int_0^T \psi_n(t)\omega(t)\,dt = 0$$

telejesül. Az α és β választásából adódik, hogy $u + \omega, u - \omega \in \hat{U}$, valamint $L\omega = 0$, következésképpen $L(u+\omega) = L(u-\omega) = Lu = y$. Ez azt jelentené, hogy u nem lehet $\hat{U} \cap L^{-1}(y)$ extremális pontja a feltevéssel ellentétben.

Ezzel megmutattuk, hogy minden $y \in L\hat{U}$ esetén $\mathrm{ex}(\hat{U} \cap L^{-1}(y)) \subset \mathrm{ex}\,\hat{U}$, ahonnan adódik az állításunk. $\square$

Megjegyezzük, hogy ezen állításunk a 4.11 Tétel értelmében az

$$L\hat{U} = L\{\chi_A : A \subset [0,1] \text{ mérhető}\}$$

alakban írható fel, ahol χ_A az A halmaz indikátorfüggvénye, azaz

$$\chi_A(t) = \begin{cases} 1, & \text{ha } t \in A \\ 0, & \text{ha } t \notin A \,. \end{cases}$$

Ha bevezetjük a

$$\nu(E) = (\nu_1(E), \ldots, \nu_n(E)) \tag{4.5}$$

jelölést, ahol

$$\nu_k(E) = \int_E \psi_k(t)\,dt \quad (\psi_k \in L^2[0,T])\,,$$

akkor ν a Lebesgue-mértékre vonatkozóan abszolút folytonos mértékekből összeállított úgynevezett *vektormérték.* A következő tételünk azt fogalmazza meg, hogy egy ilyen vektormérték értékkészlete konvex kompakt halmaz. Ez az állításunk Ljapunov egy igen általános tételének speciális alakja.

4.15 Tétel. **(Ljapunov tétele)** *Legyen $M \subset [0,T]$ pozitív mértékű, és tekintsük a (4.5) alatti ν vektormértéket. Akkor az*

$$R = \{\nu(E) : E \subset M \ \ \textit{mérhető részhalmaz}\ \}$$

értékkészlet az $\mathbb{R}^n$ kompakt konvex részhalmaza.

Bizonyítás. Bármely $E \subset M$ mérhető részhalmazra

$$\nu(E) = \left(\int_E \psi_1(t)\chi_E(t)\,dt, \ldots, \int_E \psi_n(t)\chi_E(t)\,dt \right),$$

ezért az előző álításunk jelöléseit használva $R = L(\mathrm{ex}\,\hat{U})$. Ekkor azonban az előző állítás folytán $R = L\hat{U}$, és ez utóbbi pedig konvex kompakt halmaz. □

4.16 Tétel. **(Dvoretzki-Wald-tétel)** *Tegyük fel, hogy $\nu^1, \ldots, \nu^m$ a Lebesgue-mértékre abszolút folytonos $\mathbb{R}^n$-értékű vektormértékek a $[0,T]$ intervallumon. Legyenek $\alpha_1, \ldots, \alpha_m$ olyan nemnegatív számok, amelyek összege 1. Akkor bármely $E \subset [0,T]$ mérhető halmaz felírható $E = E_1 \cup \cdots \cup E_m$ diszjunkt unió alakban, ahol*

$$\nu^k(E_k) = \alpha_k \nu^k(E)$$

minden $k = 1, \ldots, m$ esetén.

Bizonyítás. A bizonyítást indukcióval végezzük. Legyen először $m = 2$ és $\alpha_2 = 1 - \alpha_1$. Tekintsük a $2n$-dimenziós $\nu(E) = (\nu^1(E), \nu^2(E))$ vektormértéket. Ljapunov tétele szerint van olyan $E_1 \subset E$ mérhető halmaz, hogy $\nu(E_1) = \alpha_1\nu(E)$. Ez azt jelenti, hogy

$$\nu^1(E_1) = \alpha_1\nu^1(E) \quad \text{és} \quad \nu^2(E_1) = \alpha_1\nu^2(E)\,.$$

Legyen $E_2 = E \setminus E_1$, ez megfelel az álllításunknak.

Tegyük fel, hogy $m-1$-ig bezárólag az állításunk igaz, és tekintsünk m számú $\nu^1, \ldots, \nu^m$ vektormértéket. Vezessük be a

$$\nu(E) = (\nu^1(E), \ldots, \nu^m(E))$$

nm-dimenziós vektormértéket. Ljapunov tétele értelmében választhatunk olyan $E_1 \subset E$ mérhető halmazt, hogy $\nu(E_1) = \alpha_1\nu(E)$, azaz

$$\nu^k(E_1) = \alpha_1\nu^k(E)$$

minden k indexre. Legyen $E' = E \setminus E_1$, és alkalmazzuk az indukciós feltevést a $\nu^2, \dots, \nu^m$ mértékekre. Azt kapjuk, hogy $E' = E_2 \cup \dots \cup E_m$ írható, ahol

$$\nu^k(E_k) = \frac{\alpha_k}{1-\alpha_1}\nu^k(E')$$

minden $k = 2, \dots, m$ mellett. Innen következik, hogy

$$\nu^k(E_k) = \frac{\alpha_k}{1-\alpha_1}(1-\alpha_1)\nu^k(E) = \alpha_k \nu_k(E)\,,$$

amit igazolnunk kellett. □

Tekintsük ezután a (2.1) rendszert, legyen U az X konvex kompakt részhalmaza, és legyen $\hat{U}$ a (4.1) alatti megengedett irányítások halmaza.

4.17 Tétel. **(Bang-bang-elv)** $\Lambda\hat{U} = \Lambda(\operatorname{ex}\hat{U})$.

Bizonyítás. Legyen $\bar{x} \in \Lambda\hat{U}$. Mivel $\Lambda\hat{U}$ konvex kompakt halmaz az X n-dimenziós térben, azért a Caratheodory-tétel szerint találhatunk olyan $e_0, \dots, e_n$ extremális pontokat, hogy $\bar{x}$ előállítható az

$$\bar{x} = \alpha_0 e_0 + \dots + \alpha_n e_n$$

konvex kombinációval. Akkor a 4.13 Tétel alapján találhatunk olyan $u_0, \dots, u_n$ extremális irányításokat, amelyekre $\Lambda u_k = e_k$ minden indexre.

Tekintsük ezután a

$$\nu^k(E) = \int_E \Phi(0,t)B(t)u_k(t)\,dt$$

vektormértékeket, mindegyikükre $\nu_k([0,T]) = e_k$. A Dvoretzki-Wald-tétel szerint $[0,T] = E_0 \cup \dots \cup E_n$ diszjunkt unió írható, ahol

$$\nu^k(E_k) = \alpha_k \nu^k([0,T]) = \alpha_k e_k$$

minden k esetén. Definiáljuk az

$$\bar{u} = u_0 \chi_{E_0} + \ldots + u_n \chi_{E_n}$$

irányítást. Az értelmezésből világos, hogy $\bar{u}$ extremális irányítás. Másrészt

$$\begin{aligned}
\Lambda\bar{u} &= \int_0^T \Phi(0,t)B(t)\bar{u}(t)\,dt = \sum_{k=0}^n \int_{E_k} \Phi(0,t)B(t)u_k(t)\,dt \\
&= \sum_{k=0}^n \nu^k(E_k) = \sum_{k=0}^n \alpha_k e_k = \bar{x}\,,
\end{aligned}$$

amit igazolnunk kellett. □

4.6. Gyakorlatok

1. Legyen X véges dimenziós tér, és tekintsünk egy $[0, T]$ intervallumon értelmezett olyan F halmazértékű leképezést, amely az X nem üres, zárt részhalmazai közé képez. Bizonyítsuk be, hogy F akkor és csak akkor mérhető, ha bármely $x \in X$ mellett a $t \mapsto d_{F(t)}(x)$ függvény mérhető.

2. Igazoljuk, hogy ha $U : [0, T] \rightsquigarrow Y$ felülről félig folytonos konvex kompakt értékű leképezés, akkor az
$$\hat{U} = \{u \in L^2_Y[0, T] : u(t) \in U(t) \text{ m.m. } \}$$
halmaz korlátos, konvex és zárt az $L^2_Y[0, T]$ Hilbert-térben.

3. Legyen U az X olyan konvex zárt részhalmaza amely szimmetrikus, azaz $U = -U$, továbbá $0 \in \operatorname{int} U$. Tekintsük az
$$\|x\| = \inf\{t > 0 : x \in tU\}$$
úgynevezett Minkowski-normát. Mutassuk meg, hogy ez valóban normát definiál az X téren. Vajon hogyan karakterizálható az $u(t) \in U$ irányítás?

II. rész

Lokális optimalizálás: nemlineáris rendszerek

5. fejezet

Differenciálszámítás normált terekben

5.1. Differenciálhatóság

Ebben a szakaszban bevezetjük a normált tereken értelmezett függvények deriváltjának fogalmát. Normált téren mindig a valós test fölötti teret értünk. Célunk a normált terekben értelmezett függvények szélsőértékhelyeinek megkeresése.

5.1 Definíció. Legyenek X és Y normált terek, és tekintsük az X egy részhalmazán értelmezett az $F : X \to Y$ leképezést, amely értelmezve van az $x \in X$ pont egy környezetében. Azt mondjuk, hogy F *differenciálható* az x pontban, ha található olyan $A \in L(X,Y)$ folytonos lineáris leképezés, hogy bármely $v \in X$, $x + v \in D_F$ esetén

$$F(x+v) = F(x) + Av + r(v) ,$$

ahol $\lim_{v\to 0} \|r(v)\|/\|v\| = 0$. Ebben az esetben az A leképezést az F deriváltjának nevezzük az x pontban. Jelölése $A = F'(x)$.

Ezt a fogalmat az irodalomban néha Fréchet-differenciálhatóságnak is nevezik. Megjegyezzük, hogy az érintővel való közelítéshez hasonlóan a fenti definíció azt fogalmazza meg, hogy az x pont egy környezetében az F függvény megváltozása jól, azaz kis ordó nagyságrendben közelíthető az A lineáris leképezéssel. Világos ugyanis a definícióból, hogy az $r : X \to Y$ függvény kis ordó nagyságrendű a 0 egy környezetében. Vegyük észre, hogy a derivált függ az X és Y terek normáitól.

Nem világos a definícióból, hogy a derivált egyértelműen meghatározott, azaz csak egyetlen olyan A lineáris leképezés létezhet, amely kielégíti a fenti definíciót. Erre ad választ az alábbi állítás.

5.2 Állítás. *A derivált egyértelműen meghatározott.*

Bizonyítás. Tegyük fel, hogy az A és B lineáris leképezések egyaránt eleget tesznek a definíció követelményeinek, azaz, ha $x+v \in D_F$, úgy

$$\begin{aligned} F(x+v) &= F(x) + Av + r(v) \\ F(x+v) &= F(x) + Bv + q(v) \ , \end{aligned}$$

ahol r és q kis ordó függvények. Ekkor a $C = B - A$ jelöléssel a $Cv = r(v) - q(v) = o(v)$ egyenlőséghez jutunk, amely ugyancsak kis ordó függvény. Tehát tetszőleges $v \neq 0$ vektor mellet

$$\frac{\|Cv\|}{\|v\|} = \frac{\|C(\frac{1}{n}v)\|}{\|\frac{1}{n}v\|} = \frac{\|o(\frac{1}{n}v)\|}{\|\frac{1}{n}v\|} \to 0 \ ,$$

ha $n \to \infty$. Ez azt jelenti, hogy $Cv = 0$, azaz $C = B - A = 0$. $\square$

Nyilvánvaló, hogy ha F differenciálható az x pontban, akkor ott folytonos is. (Lásd az 1. gyakorlatot.) Azt is belátjuk, hogy minden lineáris leképezés differenciálható, és a deriváltja minden pontban saját maga.

5.3 Állítás. *Ha $A \in L(X,Y)$, akkor az $F(x) = Ax$ leképezés minden $x \in X$ pontban differenciálható, és $F'(x) = A$.*

Bizonyítás. Valóban, alkalmazzuk a definíciót az $F = A$, $r = 0$ szereposztás mellett. $\square$

5.4 Példa. Tekintsük az X Hilbert térben az $F : X \to \mathbb{R}$, $F(x) = \langle x, Bx\rangle$ kvadratikus alakot, ahol $B \in L(X)$ önadjungált operátor. Megmutatjuk, hogy F minden $x \in X$ pontban differenciálható, éspedig $F'(x) = 2Bx$.

Valóban, bármely $v \in X$ vektor mellett

$$\begin{aligned} F(x+v) - F(x) &= \langle x+v, B(x+v)\rangle - \langle x, Bx\rangle \\ &= \langle v, Bx\rangle + \langle x, Bv\rangle + \langle v, Bv\rangle \\ &= \langle v, 2Bx\rangle + \langle v, Bv\rangle \ , \end{aligned}$$

hiszen B önadjungált. Állításunk igazolásához tehát elég megmutatni, hogy $\langle v, Bv\rangle$ kis ordó nagyságrendű. Ez azonban egyszerűen látható a

$$|\langle v, Bv\rangle| \leq \|B\| \cdot \|v\|^2$$

Cauchy-Schwarz-féle egyenlőtlenségből.

Megjegyezzük, hogy ebben a példában $2Bx$ azt az $L(X,\mathbb{R}) = X^*$ duális térbeli lineáris funkcionált jelenti, amelyre

$$2Bx(v) = \langle v, 2Bx\rangle$$

bármely $v \in X$ esetén. Hilbert-terekben a Riesz-féle reprezentációs tétel szerint az X^* duális tér azonosítható az X térrel.

Az alábbiakban összefoglaljuk a derivált legfontosabb tulajdonságait. A következő állítás egyszerűen adódik a definícióból.

5.5 Állítás. *Tegyük fel, hogy az F és G függvények egyaránt differenciálhatók az $x \in X$ pontban, és legyen $\lambda \in \mathbb{R}$ tetszőleges. Akkor $F+G$, illetve λF is differenciálhatók az x pontban, és*

$$\begin{aligned} (F+G)'(x) &= F'(x) + G'(x) \\ (\lambda F)'(x) &= \lambda F'(x) \end{aligned}$$

Az alábbi tétel az összetett függvény deriválási szabályát általánosítja normált terekre. Vegyük észre azonban, hogy e tétel bizonyítása szinte szó szerint megegyezik az első éves analízisben tanulttal.

Legyenek tehát X, Y és Z normált terek, és tekintsük az $F : X \to Y$, valamint a $G : Y \to Z$ függvényeket. Tegyük fel, hogy x belső pontja az F értelmezési tartományának, és $F(x)$ is belső pontja a G értelmezési tartományának.

5.6 Tétel. *Ha F differenciálható az x pontban, továbbá G differenciálható az $F(x)$ pontban, akkor $G \circ F$ is differenciálható az x pontban, éspedig*

$$(G \circ F)'(x) = G'(F(x))F'(x) \ .$$

Bizonyítás. A feltételeink azt jelentik, hogy

$$F(x+v) = F(x) + F'(x)v + r(v) \ ,$$

illetve

$$G(F(x)+u) = G(F(x)) + G'(F(x))u + q(u) \ ,$$

ahol r és q egyaránt kis ordó nagyságrendűek. Ha most $v \in X$ tetszőleges, akkor az $u = F(x+v) - F(x)$ jelöléssel

$$\begin{aligned} G(F(x+v)) - G(F(x)) &= G'(F(x))u + q(u) \\ &= G'(F(x))(F(x+v) - F(x)) + q(u) \\ &= G'(F(x))(F'(x)v + r(v)) + q(u) \\ &= G'(F(x))F'(x)v + G'(F(x))r(v) + q(u) \ . \end{aligned}$$

Azt kell igazolni, hogy $G'(F(x))r(v)+q(u)$ kis ordó nagyságrendű v szerint. Ezt tagonként mutatjuk meg. Az első tagra ez a megállapítás nyilvánvaló, hiszen

$$\lim_{v\to 0} \frac{\|G'(F(x))r(v)\|}{\|v\|} \leq \|G'(F(x))\| \lim_{v\to 0} \frac{\|r(v)\|}{\|v\|} = 0 \, .$$

A második tag kis ordó nagyságrendű u szerint. Ez azonban v szerint is igaz, ugyanis

$$\frac{\|q(u)\|}{\|v\|} = \begin{cases} 0 \, , & \text{ha } F(x+v) - F(x) = 0 \\ \frac{\|q(u)\|}{\|u\|} \frac{\|F(x+v)-F(x)\|}{\|v\|} \, , & \text{ha } F(x+v) - F(x) \neq 0 \end{cases}$$

Mivel az F folytonossága miatt $v \to 0$ esetén $u \to 0$ is fennáll, azért

$$\lim_{v\to 0} \frac{\|q(u)\|}{\|v\|} = 0 \, ,$$

hiszen az

$$\frac{\|F(x+v) - F(x)\|}{\|v\|} = \frac{\|F'(x)v + r(v)\|}{\|v\|} \leq \|F'(x)\| + \frac{\|r(v)\|}{\|v\|}$$

tört korlátos. □

5.2. Iránymenti deriváltak

Tekintsünk egy $F : X \to Y$ függvényt, és legyen $v \in X$ egy rögzített, nem zérus vektor.

5.7 Definíció. Azt mondjuk, hogy F differenciálható az $x \in X$ belső pontban a v irányban, ha létezik és véges a

$$\lim_{t\to 0+} \frac{1}{t}(F(x+tv) - F(x)) = D_v F(x)$$

határérték. Ilyenkor a $D_v F(x)$ határértéket az F *iránymenti deriváltjának* nevezzük az x pontban a v irányban.

Egyszerűen belátható, hogy a $v \to D_v F(x)$ leképezés pozitív-homogén, de általában nem lineáris. Példa erre a valós értékű $F(x) = |x|$ függvény az $x = 0$ pontban. Arra az esetre amikor a linearitás is teljesül, új elnevezést vezetünk be.

5.8 Definíció. Azt mondjuk, hogy az $F : X \to Y$ függvény *Gâteaux-értelemben differenciálható* az x pontban, ha F minden irányban differenciálható az x pontban, és található olyan $A \in L(X, Y)$ lineáris leképezés, amelyre

$D_vF(x) = Av$ minden $v \in X$ vektor mellett. Ezt az A leképezést az F *Gâteaux-deriváltjának* nevezzük az x pontban. Jelölése: $A = DF(x)$.

5.9 Állítás. *Ha F differenciálható az x pontban, akkor ott Gâteaux-differenciálható is, és $DF(x) = F'(x)$.*

Bizonyítás. Legyen $v \in X$ egy tetszőleges nem zérus vektor. Akkor a differenciálhatóság folytán

$$F(x + tv) = F(x) + F'(x)(tv) + o(tv) \ ,$$

azaz átrendezve és határértékre térve $D_vF(x) = F'(x)v$. Ebből már közvetlenül adódik az állítás. $\square$

Az állításunk megfordítása már nem érvényes, ezt mutatja az alábbi példánk.

5.10 Példa. Vizsgáljuk a következő függvényt a síkon:

$$F(x,y) = \begin{cases} 1 & \text{ha } \ (x-1)^2 + y^2 = 1 \, , \text{ és } \ (x,y) \neq (0,0) \\ 0 & \text{különben} \ . \end{cases}$$

Nyilvánvaló, hogy bármely origón átmenő egyenesnek van olyan, az origót is tartalmazó szakasza, amelyen e függvény zérus. Ez azt jelenti, hogy F minden v irányban differenciálható az origóban, éspedig $D_vF(0,0) = 0$, ezért Gâteaux-differenciálható is, méghozzá $DF(0,0) = [0,0]$. Azonban F még csak nem is folytonos az origóban, hiszen annak minden környezetében felveszi a 0 és az 1 értéket is. Ezért ott nem lehet differenciálható sem.

5.3. Folytonos differenciálhatóság

Folytonos differenciálhatóság esetén a Gâteaux-differenciálhatóságból már következik a Fréchet-differenciálhatóság

5.11 Tétel. (Lagrange-féle középérték-tétel) *Tegyük fel, hogy $F : X \to Y$ Gâteaux-értelemben folytonosan differenciálható az x pont egy környezetében. Akkor*

$$F(x + v) - F(x) = \int_0^1 DF(x + tv)v \, dt \, .$$

Nevezetesen

$$\|F(x + v) - F(x)\| \leq \sup_{0 \leq t \leq 1} \|DF(x + tv)\| \cdot \|v\|$$

az x pont környezetében.

Bizonyítás. Vezessük be a $\varphi(t) = F(x+tv)$ függvényt a $[0,1]$ intervallumon, akkor φ folytonosan differenciálható, és $\varphi'(t) = DF(x+tv)v$. Innen a Newton-Leibniz-formula alapján következik az első állítás. A második állítás az integrál triviális becslése szerint

$$\|F(x+v) - F(x)\| \leq \int_0^1 \|DF(x+tv)v\| \, dt \leq \sup_{0 \leq t \leq 1} \|DF(x+tv)\| \cdot \|v\|$$

bármely v mellett. $\square$

5.12 Tétel. *Tegyük fel, hogy F Gâteaux-differenciálható egy környezetben, és itt $x \to DF(x)$ folytonos. Akkor itt F differenciálható is, és $F'(x) = DF(x)$*

Bizonyítás. Valóban, a folytonosság miatt bármely ε pozitív számhoz található olyan $\delta > 0$, hogy minden $\|v\| < \delta$ mellett

$$\|DF(x+tv) - DF(x)\| < \varepsilon \, .$$

Másrészt a középérték-tétel szerint

$$F(x+v) - F(x) = DF(x)v + \int_0^1 (DF(x+tv) - DF(x))v \, dt \, .$$

Itt a jobboldalon a második tag kis ordó nagyságrendű, hiszen $\|v\| < \delta$ esetén

$$\left\| \int_0^1 (DF(x+tv) - DF(x))v \, dt \right\| \leq \sup_{0 \leq t \leq 1} \|DF(x+tv) - DF(x)\| \cdot \|v\| \leq \varepsilon \cdot \|v\| \, .$$

Ez azt jelenti, hogy F differenciáható, és $F'(x) = DF(x)$ a környezet minden pontjában. $\square$

5.4. Példák differenciálhatóságra

Az alábbiakban példákat mutatunk normált téren értelmezett függvények differenciálhatóságára. Ezek a függvények fontos szerepet játszanak a későbbi fejezetekben.

5.13 Példa. Legyen f a $[0,T] \times X$ halmazon értelmezett olyan folytonos függvény, amely folytonosan differenciálható a második változójában, azaz $f(t,.)$ folytonosan differenciálható minden $t \in [0,T]$ mellett az X téren. Tekintsük a $C[0,T]$ téren az

$$F(x) = \int_0^T f(s, x(s)) \, ds$$

függvényt. Megmutatjuk, hogy F Gâteaux-differenciálható a $C[0,T]$ minden pontjában. Valóban, tetszőleges $v \in C[0,T]$ esetén vezessük be a valós változós

$$F_v(t) = \int_0^T f(s, x(s) + tv(s)) \, ds$$

függvényt. A paraméteres integrál differenciálhatósága szerint F_v differenciálható, és

$$F_v'(t) = \int_0^T \langle \partial_2 f(s, x(s) + tv(s)), v(s) \rangle \, ds$$

minden t pontban. Másrészt a $t = 0$ pontban az integrál folytonos lineáris funkcionált definiál a $C[0,T]$ téren, ezért

$$DF(x)v = F_v'(0) = \int_0^T \langle \partial_2 f(s, x(s)), v(s) \rangle \, ds \, .$$

Mivel a jobb oldalon álló integrál a feltételünk szerint folytonos is az x változóban, azért az 5.12 Tétel alapján

$$F'(x)v = DF(x)v = \int_0^T \langle \partial_2 f(s, x(s)), v(s) \rangle \, ds \, ,$$

azaz F folytonosan differenciálható is az x pontban.

5.14 Példa. Legyen ezután f a $[0,T] \times X \times Y$ halmazon értelmezett olyan folytonos függvény, amely folytonosan differenciálható a második és harmadik változójában, azaz $f(t,.,.)$ folytonosan differenciálható minden $t \in [0,T]$ mellett az $X \times Y$. Tekintsük a $C[0,T] \times L^2[0,T]$ szorzattéren az

$$F(x,u) = \int_0^T f(s, x(s), u(s)) \, ds$$

függvényt. Megmutatjuk, hogy F Gâteaux-differenciálható a $C[0,T] \times L^2[0,T]$ minden pontjában. Valóban, tetszőleges $v \in C[0,T]$ és $w \in L^2[0,T]$ esetén vezessük be a valós változós

$$F_{v,w}(t) = \int_0^T f(s, x(s) + tv(s), u(s) + tw(s)) \, ds$$

függvényt. A paraméteres integrál differenciálhatósága szerint $F_{v,w}$ differenciálható, és

$$\begin{aligned} F_{v,w}'(t) &= \int_0^T (\langle \partial_2 f(s, x(s) + tv(s), u(s) + tw(s)), v(s) \rangle \\ &\quad + \langle \partial_3 f(s, x(s) + tv(s), u(s) + tw(s)), w(s) \rangle) \, ds \end{aligned}$$

minden t pontban. Másrészt a $t = 0$ pontban az integrál folytonos lineáris funkcionált definiál a $C[0,T] \times L^2[0,T]$ téren, ezért

$$\begin{aligned} DF(x,u)(v,w) = F_{v,w}'(0) &= \int_0^T (\langle \partial_2 f(s, x(s), u(s)), v(s) \rangle \\ &\quad + \langle \partial_3 f(s, x(s), u(s)), w(s) \rangle) \, ds \, . \end{aligned}$$

Mivel a jobb oldalon álló integrál a feltételünk szerint folytonos is az (x,u) változóban, azért az 5.12 Tétel alapján

$$\begin{aligned} F'(x,u)(v,w) = DF(x)(v,w) &= \int_0^T (\langle \partial_2 f(s,x(s),u(s)), v(s)\rangle \\ &\quad + \langle \partial_3 f(s,x(s),u(s)), w(s)\rangle)\, ds\,, \end{aligned}$$

azaz F folytonosan differenciálható is az (x,u) pontban.

5.15 Példa. Tekintsük most azon x abszolút folytonos függvények $W^2[0,T]$ vektorterét, amelyekre $x' \in L^2[0,T]$. Lássuk el ezt a vektorteret az

$$\|x\| = \max_{[0,T]} \|x(s)\| + \|x'\|_{L^2}$$

normával, világos, hogy így $W^2[0,T]$ normált tér. Tekintsük ezen a téren az

$$F(x) = \int_0^T f(s,x(s),x'(s))\, ds$$

függvényt. Megmutatjuk, hogy F minden $x \in W^2[0,T]$ pontban Gâteaux-differenciálható. Tekintsük ugyanis egyrészt az előző példában szereplő

$$\hat{F}(x,u) = \int_0^T f(s,x(s),u(s))\, ds$$

differenciálható függvényt. Ez a függvény nyilván a $W^2[0,T]$ tér erősebb topológiájában is differenciálható. Másrészt a

$$D : W^2[0,T] \to L^2[0,T], \quad Dx = x'$$

lineáris leképezés folytonos, így differenciálható is. Ezért az összetett függvény differenciálhatóságáról szóló 5.6 Tétel értelmében

$$F(x) = \hat{F}(x,Dx) = \int_0^T f(s,x(s),x'(s))\, ds$$

is differenciálható, éspedig

$$F'(x)v = \int_0^T (\langle \partial_2 f(s,x(s),x'(s)), v(s)\rangle + \langle \partial_3 f(s,x(s),x'(s)), v'(s)\rangle)\, ds$$

minden $v \in W^2[0,T]$ esetén. Könnyen ellenőrizhető, hogy ebben a normában F folytonosan is differenciálható.

5.5. Szélsőérték

5.16 Tétel. *Tegyük fel, hogy az $F : X \to \mathbb{R}$ függvénynek lokális minimuma van az x pontban. Ha F differenciálható valamely v irányban az x pontban, akkor ott $D_v F(x) \geq 0$.*

Bizonyítás. Valóban, ha x lokális minimumhely, akkor

$$F(x + tv) - F(x) \geq 0$$

minden elég kicsi t értékre, és innen adódik az állítás. □

Nyilván analóg állítást fogalmazhatunk meg a lokális maximum esetére.

5.17 Következmény. *Tegyük fel, hogy az $F : X \to \mathbb{R}$ függvénynek lokális minimuma van az x pontban. Ha F Gâteaux-differenciálható az x pontban, akkor ott $DF(x) = 0$.*

Bizonyítás. Az előző állításunk alapján bármely v irányban $D_v F(x) = DF(x)v \geq 0$. Ez azonban $DF(x)$ linearitása miatt csak úgy lehetséges, hogy $DF(x) = 0$. □

A függvény értelmezési tartományának azon pontjait, ahol a függvény differenciálható, és a derivált zérus, *kritikus pontoknak* nevezzük.

5.18 Példa. Legyen $B \in L(X)$ az X Hilbert-tér önadjungált transzformációja, és tekintsük az $F(x) = \langle x, Bx \rangle$ kvadratikus alakot. Keressük meg az F szélsőértékeit. Az 5.4 Példa szerint F differenciálható, és $F'(x) = 2Bx$. Az 5.16 Tétel alapján a kritikus pontok az

$$F'(x) = 2Bx = 0$$

homogén lineáris egyenlet megoldásai, azaz a $\ker B$ altér elemei. Világos, hogy minden kritikus pont (globális) minimumhely, ha B pozitív szemidefinit, mindegyik (globális) maximumhely, ha B negatív szemidefinit, illetve egyik sem lokális szélsőértékhely, ha B indefinit. Ez utóbbi esetben az F kritikus pontjait *nyeregpontoknak* nevezzük.

5.6. Monotonitás és konvexitás

Ebben a szakaszban konvex függvények esetében elégséges feltételt fogalmazunk meg a minimumhely létezésére.

5.19 Állítás. *Ha F konvex függvény a $D \subset X$ konvex halmazon, $x \in D$, és v olyan vektor, hogy $x + v \in D$, akkor F differenciálható az x pontban a v irányban.*

Bizonyítás. Tekintsük a $[0,1]$ intervallumon a

$$\varphi(t) = F(x+tv) - F(x)$$

függvényt. Nyilvánvaló, hogy φ konvex, így minden pontban létezik a jobb oldali deriváltja. Nevezetesen a 0 pontban a

$$\varphi'_+(0) = \lim_{t \to 0+} \frac{1}{t}(F(x+tv) - F(x)) = D_v F(x)$$

határérték létezik és véges. □

5.20 Állítás. *Tekintsük az F függvényt a $D \subset X$ konvex halmazon. Az F akkor és csak akkor konvex, ha*

$$D_v F(y) - D_v F(x) \geq 0$$

bármely $x, y \in D$ és $v = y - x$ esetén.

Bizonyítás. Szükségesség. Legyenek $x, y \in D$ tetszőlegesek, és $v = y - x$. Ha F konvex, akkor

$$\varphi(t) = F(x+tv)$$

konvex a $[0,1]$ intervallumon. Ezért φ'_+ létezik és monoton növő, tehát

$$0 \leq \varphi'_+(1) - \varphi'_+(0) = D_v F(y) - D_v F(x)\,.$$

Elégségesség. Legyen $0 \leq t \leq 1$ és $x, y \in X$. Ekkor a $v = y - x$ jelöléssel

$$\varphi(t) = F(ty + (1-t)x) = F(x + t(y-x)) = F(x+tv)\,.$$

A feltételünk alapján tetszőleges $t, s \in [0,1]$ mellett

$$(\varphi'_+(t) - \varphi'_+(s))(t-s) = D_v F(x+tv) - D_v(x+sv))(t-s) \geq 0\,,$$

azaz φ'_+ monoton növő. Ezért φ konvex a $[0,1]$ intervallumon. Tehát

$$F(x+tv) = \varphi(t) \leq t\varphi(1) + (1-t)\varphi(0) = tF(y) + (1-t)F(x)$$

bármely $0 \leq t \leq 1$ esetén, és így F konvex az X téren. □

A fenti állítás alapján érdemes bevezetni a monotonitás alábbi általánosabb fogalmát.

5.21 Definíció. Tekintsük az X normált teret. Az $A : X \to X^*$ leképezést *monoton leképezésnek* nevezzük, ha

$$(A(y) - A(x))(y - x) \geq 0$$

minden $x, y \in X$ esetén.

Tekintsünk ezután egy olyan $F : X \to \mathbb{R}$ leképezést, amely az X téren Gâteaux-differenciálható.

5.22 Következmény. *Az $F : X \to \mathbb{R}$ leképezés akkor és csak akkor konvex, ha $DF : X \to X^*$ monoton leképezés.*

A következő tételünk azt fogalmazza meg, hogy konvex függvények esetében az 5.16 Tétel szükséges és elégséges feltételt ad a minimumhely létezésére.

5.23 Tétel. *Ha F konvex a D konvex halmazon, továbbá az $x \in D$ pontban $D_vF(x) \geq 0$ minden olyan v irányban, amelyre $x + v \in D$, akkor x az F minimumhelye a D halmazon.*

Bizonyítás. Valóban, minden $0 < t < 1$ esetén az F konvexitása folytán

$$F((1-t)x + t(x+v)) \leq (1-t)F(x) + tF(x+v)$$

minden olyan v irányra, amelyre $x + v \in D$. Innen

$$\frac{1}{t}(F(x+tv) - F(x)) \leq F(x+v) - F(x)\,.$$

Az előző állításunk miatt az iránymenti derivált létezik, ezért

$$0 \leq D_vF(x) \leq F(x+v) - F(x)\,,$$

azaz x valóban minimumhely. □

5.24 Következmény. *Ha az F konvex függvény Gâteaux-differenciálható az x pontban és ott $DF(x) = 0$, akkor x az F minimumhelye.*

Megjegyezzük, hogy konvex függvények esetében minden lokális minimumhely egyúttal globális minimumhely is egy konvex halmazon.

Nyilvánvaló, hogy analóg állításokat fogalmazhatunk meg a maximumhelyekre vonatkozóan konkáv függvények esetében.

5.7. Gyakorlatok

1. Közvetlenül a definíció alapján ellenőrizzük, hogy ha az $F : X \to Y$ függvény differenciálható az $x \in X$ pontban, akkor ott folytonos is.

2. Differenciálható-e az $F(x) = \langle x, Bx \rangle$ kvadratikus alak akkor is, ha B nem önadjungált operátor? Ha igen, adjuk meg a deriváltját. (Vesd össze az 5.4 Példával.)

3. Végezzük el az 5.5 Állítás bizonyítását.

4. Mutassuk meg, hogy ha az x pontban az F függvény differenciálható a v irányban, akkor $D_{\lambda v}F(x) = \lambda D_v F(x)$ bármely $\lambda > 0$ esetén.

5. A Gâteaux-differenciálhatóságból nem következik a differenciálhatóság. Tekintsük a síkon az

$$F(x,y) = \begin{cases} 1 & \text{ha } y = x^2, \;\; (x,y) \neq (0,0) \\ 0 & \text{különben} \end{cases}$$

függvényt. Igazoljuk, hogy F Gâteaux-differenciálható a 0 pontban, és $DF(0,0) = [0,0]$. Azonban F még csak nem is folytonos az origóban, hiszen annak bármely környezetében egyaránt felveszi a 0 és az 1 értékeket is.

6. Mutassuk meg, hogy ha egy normált téren értelmezett függvény differenciálható, akkor differenciálható marad bármely azzal ekvivalens normában is. Változik-e vajon a derivált, ha ekvivalens normára térünk át?

6. fejezet

Variációszámítás

Ebben a fejezetben a klasszikus variációszámítás legegyszerűbb feladattípusát tárgyaljuk. Ez valójában egy függvénytéren értelmezett szélsőérték-feladat, amelyben az optimalitás szükséges feltételét a normált téren vett derivált szolgáltatja.

6.1. A Lagrange-feladat

Legyen X véges dimenziós euklideszi tér, x_0 és x_T adott pontok az X térben. Legyen továbbá $f : [0,T] \times X \times X \to \mathbb{R}$ olyan folytonos függvény, amelyre $f(t,.,.)$ folytonosan differenciálható az $X \times X$ téren minden $t \in [0,T]$ esetén.

Vezessük be a következő függvényteret:

$$W^2[0,T] = \{x : [0,T] \to X,\ x \text{ abszolút folytonos},\ x' \in L^2[0,T]\}\,,$$

Lássuk el ezt a teret az

$$\|x\| = \|x\|_0 + \|x'\|_2$$

normával, ahol jobb oldalon az első tag a szokásos maximum norma a $C[0,T]$ térben, a második tag pedig az $L^2[0,T]$ tér szokásos normája. Tekintsük a $W^2[0,T]$ téren az alábbi függvényt:

$$F(x) = \int_0^T f(t, x(t), x'(t))\, dt\,, \tag{6.1}$$

amelynek értelmezési tartománya a

$$D = \{x \in W^2[0,T] : x(0) = x_0, x(T) = x_T\} \tag{6.2}$$

halmaz. Keressük meg az F függvény lokális minimumhelyét a D halmazon a $W^2[0,T]$ tér normájára nézve. Az ilyen alakú feladatokat *Lagrange-feladatnak* nevezzük. Az alábbiakban szükséges feltételt keresünk a minimumhelyre.

Vezessük be a $W^2[0,T]$ tér

$$W_0^2[0,T] = \{x \in W^2[0,T] : x(0) = x(T) = 0\}$$

alterét. A szükséges feltételünk a következő lemmán múlik.

6.1 Lemma. **(Du Bois-Reymond-lemma)** *Tegyük föl, hogy $y \in L_X^2[0,T]$ olyan függvény, amelyre*

$$\int_0^T \langle y(t), v'(t) \rangle \, dt = 0$$

minden $v \in W_0^2[0,T]$ esetén. Akkor y konstans a $[0,T]$ intervallumon majdnem mindenütt.

Bizonyítás. A Newton-Leibniz-formula miatt világos, hogy bármely $a \in X$ vektor mellett

$$\int_0^T \langle y(t) - a, v'(t) \rangle \, dt = 0$$

minden $v \in W_0^2[0,T]$ esetén. Nevezetesen válasszuk az

$$a = \frac{1}{T} \int_0^T y(t) \, dt$$

vektort az X térből, és tekintsük a

$$v(t) = \int_0^t (y(s) - a) \, ds$$

függvényt. Könnyen ellenőrizhető, hogy ekkor $v \in W_0^2[0,T]$, továbbá

$$0 = \int_0^T \langle y(t) - a, v'(t) \rangle \, dt = \int_0^T \langle y(t) - a, y(t) - a \rangle \, dt = \int_0^T \|y(t) - a\|^2 \, dt \, .$$

Ez éppen azt jelenti, hogy $y(t) = a$ majdnem mindenütt. □

6.2. Az Euler-Lagrange-egyenlet

Az eddig vizsgált szélsőértékfeladatokkal szemben a Lagrange-feladat minimalizálandó függvénye nem véges dimenziós téren, hanem egy függvénytér részhalmazán van definiálva. Gondolhatunk arra, hogy a szélsőérték helyeket az

$$F'(x) = 0$$

egyenlet megoldásai között keressük, ez a derivált azonban nem létezik, hiszen F nem nyílt halmazon van értelmezve. Használhatjuk azonban az iránymenti derivált fogalmát.

Az 5.15 Példa szerint a (6.1) alatti F függvény iránymenti deriváltjára az alábbi adódik.

6.2 Állítás. *Minden $x \in D$ pontban az F függvény differenciálható bármely $v \in W_0^2[0,T]$ irányban, éspedig*

$$D_v F(x) = \int_0^T \left(\langle \partial_2 f(t, x(t), x'(t)), v(t) \rangle + \langle \partial_3 f(t, x(t), x'(t)), v'(t) \rangle \right) dt$$

6.3 Tétel. *Ha x a Lagrange-feladat megoldása, akkor*

$$\partial_3 f(t, x(t), x'(t)) = \int_0^t \partial_2 f(s, x(s), x'(s))\, ds + \text{const} \tag{6.3}$$

majdnem minden $t \in [0,T]$ pontban.

Bizonyítás. Ha x megoldás, akkor az 5.16 Állítás folytán $D_v F(x) \geq 0$ minden $v \in W_0^2[0,T]$ irányban. Mivel ez az egyenlőtlenség egy altéren teljesül, és $v \to D_v F(x)$ lineáris az előző állítás szerint, ez azonban csak úgy lehetséges, ha $D_v F(x) = 0$. Ez újra az előző állítás szerint azt jelenti, hogy

$$\int_0^T \left(\langle \partial_2 f(t, x(t), x'(t)), v(t) \rangle + \langle \partial_3 f(t, x(t), x'(t)), v'(t) \rangle \right) dt = 0\,.$$

Parciálisan integrálva az első tagot azt kapjuk, hogy

$$\int_0^T \left\langle \partial_3 f(t, x(t), x'(t)) - \int_0^t \partial_2 f(s, x(s), x'(s))\, ds, v'(t) \right\rangle dt = 0$$

minden $v \in W_0^2[0,T]$ esetén. A Du Bois-Reymond-lemma miatt tehát

$$\partial_3 f(t, x(t), x'(t)) - \int_0^t \partial_2 f(s, x(s), x'(s))\, ds = \text{const}$$

majdnem mindenütt a $[0,T]$ intervallumon. $\square$

6.4 Definíció. A (6.3) egyenletet *Euler-Lagrange-egyenletnek* nevezzük. Az egyenlet megoldásait *stacionárius függvényeknek* nevezzük.

Megjegyezzük, hogy ha f kétszer folytonosan differenciálható a második és harmadik változójában, úgy x akkor és csak akkor extremális, ha

$$\partial_2 f(t, x(t), x'(t)) = \frac{d}{dt} \partial_3 f(t, x(t), x'(t))$$

majdnem mindenütt. Ezt az egyenletet Euler-Lagrange-féle differenciálegyenletnek nevezzük.

6.5 Példa. Tekintsük az

$$F(x) = \int_0^1 (4x(t) + x'(t)^2)\,dt \to \min, \quad x_0 = 0\,, \quad x_1 = 1$$

Lagrange-feladatot. Ebben az esetben $f(t, x, u) = 4x + u^2$, ezért az Euler-Lagrange-egyenlet

$$\frac{d}{dt}2x'(t) = 2x''(t) = 4$$

alakú. A peremfeltételeket figyelembe véve ennek egyetlen megoldása van, méghozzá $x(t) = t^2$ a $[0, 1]$ intervallumon.

Megmutatjuk, hogy ez a stacionárius függvény megoldása a feladatnak. Valóban, bármely $v \in W_0^2[0, 1]$ mellett

$$\begin{aligned} F(x+v) &= \int_0^1 (4(x(t) + v(t)) + (x'(t) + v'(t))^2\,dt \\ &= F(x) + \int_0^1 4v(t)\,dt + \int_0^1 2x'(t)v'(t)\,dt + \int_0^1 v'(t)^2\,dt \\ &= F(x) + 2\int_0^1 x''(t)v(t)\,dt + 2\int_0^1 x'(t)v'(t)\,dt + \int_0^1 v'(t)^2\,dt\,, \end{aligned}$$

hiszen $x''(t) = 2$. Mivel a jobb oldalon az első két integrálban

$$x''(t)v(t) + x'(t)v'(t) = \frac{d}{dt}(x'(t)v(t))\,,$$

ezért a Newton-Leibniz-formulára tekintettel

$$\int_0^1 x''(t)v(t)\,dt + \int_0^1 x'(t)v'(t)\,dt = [x'(t)v(t)]_0^1 = 0\,.$$

Innen azonnal adódik, hogy

$$F(x+v) = F(x) + \int_0^1 v'(t)^2\,dt \geq F(x)\,,$$

azaz $x(t) = t^2$ valóban a Lagrange-feladat megoldása a $[0, 1]$ intervallumon.

6.3. Elégséges feltétel

Megmutatjuk ebben a szakaszban, hogy konvexitási feltétel mellett az Euler-Lagrange-egyenlet az optimalitás szükséges és elégséges feltétele.

6.6 Tétel. *Tegyük fel, hogy a (6.1), (6.2) alatti Lagrange-feladatban az f függvény konvex az (x, u) változóban. Ekkor x akkor és csak akkor a feladat megoldása, ha x extremális, azaz kielégíti az Euler-Lagrange egyenletet.*

Bizonyítás. A feltételünkből következik, hogy ilyenkor F konvex, ezért az 5.23 Állítás alapján a tételünk abból adódik, hogy x pontosan akkor szélsőértékhely, ha $D_v F(x) \geq 0$ minden $v \in W_0^2[0,T]$ irányban. □

6.7 Példa. **(Egy termelésirányítási modell)** Egy vállalat rendelést kap x_T egységnyi termék leszállítására a T időpontig. Ezt a rendelést a cég minimális költség mellett kívánja teljesíteni. Föltételezzük, hogy a termelési költség a termelési ütem megváltozásának valamely c függvénye, továbbá a kész termék raktározási költsége arányos az eltelt idővel és a termékmennyiséggel. Jelölje α ezt az arányossági tényezőt.

Legyen 0 a kezdeti időpont, x_0 a $t = 0$ kezdeti időpontban rendelkezésre álló raktárkészlet az adott termékből, és jelentse $x(t)$ a t időpontig felhalmozott készletet. Ekkor a t időpillanatbeli költség

$$c(x'(t)) + \alpha x(t)\,.$$

A teljes költség minimalizálása a $[0,T]$ intervallumon azt jelenti, hogy olyan x függvényt keresünk, amelyre

$$\int_0^T \left(c(x'(t)) + \alpha x(t)\right)\, dt \;\to\; \min\,,$$

és amely kielégíti a feltételeinket, azaz

$$x(0) = x_0\,, \quad x(T) = x_T\,, \quad x'(t) \geq 0\,.$$

Erre a Lagrange-feladatra az Euler-Lagrange-egyenlet a következő alakot ölti

$$\frac{d}{dt} c'(x'(t)) = \alpha$$

Ha például $c(u) = u^2$, akkor a feladat az alábbi egyszerű másodrendű differenciálegyenletre redukálódik

$$2x''(t) = \alpha,$$

amelynek általános megoldása

$$x(t) = \frac{\alpha}{4} t^2 + \beta t + \gamma\,.$$

Itt a β és γ valós állandók az $x(0) = x_0$ és $x(T) = x_T$ peremfeltételekből egyszerűen meghatározhatók. Az integrandus konvexitása folytán ez a függvény valóban a feladat egyetlen megoldását szolgáltatja.

6.4. Szabad végpontú feladatok

Tekintsük most újra a (6.1) és (6.2) alatti Lagrange-feladatot de a (6.2) peremfeltételből hagyjuk el az $x(T) = x_T$ feltételt, azaz az $x(T)$ végpont szabad. Keressünk szükséges feltételt a minimumhelyre ebben az esetben.

Először megfogalmazzuk a Du Bois-Reymond-lemma megfelelő változatát.

6.8 Lemma. *Tegyük föl, hogy $y \in L^2_X[0,T]$ olyan függvény, amelyre*

$$\int_0^T \langle y(t), v'(t)\rangle \, dt = 0$$

minden olyan $v \in W^2[0,T]$ mellett, amelyre $v(0) = 0$. Akkor y azonosan nulla a $[0,T]$ intervallumon majdnem mindenütt.

Bizonyítás. Valóban, könnyen látható, hogy adott y mellett a

$$v(t) = \int_0^t y(s)\, ds$$

függvény kielégíti a feltételt. Ezt a feltételbe behelyettesítve

$$0 = \int_0^T \langle y(t), v'(t)\rangle \, dt = \int_0^T \|y(t)\|^2 \, dt \,,$$

ahonnan azonnal adódik az állítás. □

6.9 Tétel. *Ha x a szabad végpontú feladat megoldása, akkor x kielégíti a (6.3) Euler-Lagrange-egyenletet, továbbá*

$$\partial_3 f(T, x(T), x'(T)) = 0 \,. \tag{6.4}$$

Bizonyítás. Valóban, egyrészt $D_v F(x) = 0$ minden $v \in W^2_0[0,T]$ esetén, ezért fennáll az Euler-Lagrange-egyenlet.

Másrészt ilyenkor $D_v F(x) = 0$ bármely olyan $v \in W^2[0,T]$ mellett is, amelyre $v(0) = 0$ és $v(T)$ tetszőleges, hiszen ilyenkor $x + v$ kielégíti a kezdeti feltételt. Ez az előző lemmánkra tekintettel azt jelenti, hogy

$$\int_0^t \partial_2 f(s, x(s), x'(s))\, ds - \partial_3 f(t, x(t), x'(t)) = 0$$

a $[0,T]$ intervallumon. Tehát parciális integrálással

$$\begin{aligned}
0 = D_v F(x) \quad &= \quad \int_0^T \left(\langle \partial_2 f(t, x(t), x'(t)), v(t)\rangle + \langle \partial_3 f(t, x(t), x'(t)), v'(t)\rangle\right) dt \\
&= \quad \left[\int_0^t \langle \partial_2 f(s, x(s), x'(s))\, ds, v(t)\rangle \, dt\right]_0^T - \\
&- \quad \int_0^T \left\langle \int_0^t \partial_2 f(s, x(s), x'(s))\, ds - \partial_3 f(t, x(t), x'(t)), v'(t)\right\rangle dt \,.
\end{aligned}$$

Az előző egyenlőség szerint ekkor

$$
\begin{aligned}
0 = D_v F(x) &= \left[\int_0^t \langle \partial_2 f(s, x(s), x'(s))\, ds, v(t) \rangle \, dt \right]_0^T \\
&= \left\langle \int_0^T \partial_2 f(s, x(s), x'(s))\, ds, v(T) \right\rangle
\end{aligned}
$$

tetszőleges $v(T) \in X$ mellett. Tehát a módosított Du Bois-Reymond-lemma szerint

$$
\int_0^T \partial_2 f(s, x(s), x'(s))\, ds = \partial_3 f(T, x(T), x'(T)) = 0\,,
$$

amit igazolnunk kellett. □.

6.10 Definíció. A szabad végpontú feladatok esetén a (6.4) feltételt *transzverzalitási feltételnek* nevezzük.

Nem nehéz végig gondolni, hogy ha a szabad végpontú feladatban $x(T) = x_T$ és $x(0)$ szabad, akkor a (6.4) transzverzalitási feltétel a

$$
\partial_3 f(0, x(0), x'(0)) = 0 \tag{6.5}
$$

egyenletre módosul (lásd a 11 gyakorlatot).

Az előző szakasz elégséges feltételéhez teljesen hasonlóan gondolható meg a következő tétel.

6.11 Tétel. *Ha az f függvény konvex a második és harmadik változójában, továbbá x stacionárius, azaz megoldása a (6.3) Euler-Lagrange-egyenletnek, valamint kielégíti a (6.4) transzverzalitási feltételt, akkor x megoldása a szabad végpontú feladatnak.*

6.12 Példa. **(Egy egyszerű makroökonómiai feladat)** Jelentse $y(t)$ a gazdaság állapotvektorát a t időpillanatban a $[0, T]$ időintervallumon, és a gazdaságot az x_0 kezdeti állapotból az x_T állapotba kívánjuk vezérelni minimális költség mellett. Az állapot megváltoztatása költséges, és az összköltséget az

$$
\int_0^T \left((y(t) - x_T)^2 + \alpha^2 y'(t)^2 \right) \, dt
$$

integrál adja, ahol α pozitív valós szám. Ha bevezetjük az $x(t) = y(t) - x_T$ jelölést a $[0, T]$ intervallumon, akkor a következő variációs feladathoz jutunk:

$$
\int_0^T \left(x(t)^2 + \alpha^2 x'(t)^2 \right) \, dt \to \min
$$

ahol

$$x(0) = x_0 - x_T \quad \text{és} \quad x(T) = 0\,.$$

Könnyen látható, hogy a feladat Euler-Lagrange-egyenlete

$$x''(t) - \frac{1}{\alpha^2}x(t) = 0\,.$$

Az $r = 1/\alpha$ jelölés bevezetésével ennek a másodrendű differenciálegyenletnek az általános megoldása

$$x(t) = A \cdot e^{rt} + B \cdot e^{-rt}\,,$$

ahol A és B tetszőleges valós állandók. Innen A és B a peremfeltételek behelyettesítésével egyszerűen meghatározhatók, nevezetesen

$$x(t) = \frac{x_0 - x_T}{e^{rT} - e^{-rT}}\left(e^{r(T-t)} - e^{-r(T-t)}\right)\,.$$

Az integrandus konvexitása miatt ez a feladat egyetlen megoldása.

Vizsgáljuk most ezt a feladatot úgy is, hogy az $x(T)$ végpont szabad. Ilyenkor az Euler-Lagrange-egyenlet általános megoldása az $x(0) = x_0 - x_T$ kezdeti feltétel figyelembe vételével

$$x(t) = Ae^{rt} + (x_0 - x_T - A)e^{-rt}\,.$$

A transzverzalitási feltételből azt kapjuk, hogy $x'(T) = 0$. Innen A már egyszerű elsőfokú egyenlettel meghatározható, nevezetesen

$$x(t) = \frac{x_0 - x_T}{e^{rT} - e^{-rT}}\left(e^{r(T-t)} - e^{-r(T-t)}\right)$$

az előző esethez hasonlóan. Érdemes megjegyezni, hogy itt

$$x(T) = \frac{2(x_0 - x_T)}{e^{rT} - e^{-rT}} \to 0$$

ha $T \to \infty$. Hasonló módon, ha $\alpha \to 0$, akkor $r \to \infty$ és ekkor rögzített T mellett

$$\lim_{r\to\infty} x(t) = \lim_{r\to\infty} \frac{e^{r(T-t)} + e^{-r(T-T)}}{e^{rT} + e^{-rT}} = 0$$

minden rögzített t időpontban. Ez nem meglepő, hiszen ha $\alpha \to 0$, akkor az állapotváltozás költsége elhanyagolhatóvá válik, és így x nullához tart.

6.5. A haszonmaximalizálási feladat

6.13 Példa. **(Élethosszig tartó haszonmaximalizálás)** Jelentse S a felhalmozott megtakarításunkat, amelyet az adott $[0, T]$ időintervallumon a legnagyobb hasznosságot biztosító módon kívánunk felhasználni. Ez azt jelenti, hogy rendelkezünk egy, a mikroökonómiában szokásos tulajdonságokkal rendelkező U hasznossági függvénnyel.

Ha a $t \in [0, T]$ időpontban $r(t)$ jelenti a felhasznált vagyont, akkor a feladatunk úgy írható le, hogy keressük azt az $r(t)$ függvényt, amelyre az

$$\int_0^T e^{-\beta t} U(r(t))\, dt$$

integrál maximális, ahol $\beta > 0$ a jövőbeli érték diszkont faktora.

Jelölje $x(t)$ a rendelkezésre álló megtakarítást a t időpillanatban. Tegyük föl, hogy vagyonunk névérteke fix α kamatlábbal növekszik, azaz

$$x'(t) = \alpha x(t) - r(t)$$

Ekkor feladatunk az alábbi Lagrange-feladatként fogalmazható meg:

$$\int_0^T e^{-\beta t} U(\alpha x(t) - x'(t))\, dt \to \max$$

ahol a peremfeltételek

$$x(0) = S \quad \text{és} \quad x(T) = 0\,.$$

A feladat Euler-Lagrange-egyenlete

$$\frac{d}{dt} e^{-\beta t} U'(\alpha x(t) - x'(t)) = -\alpha e^{-\beta t} U'(\alpha x(t) - x'(t))\,,$$

ahonnan a következő szétválasztható változójú differenciálegyenlet adódik:

$$U''(\alpha x(t) - x'(t)) = (\beta - \alpha) U'(\alpha x(t) - x'(t))\,.$$

Innen az $r(t)$ függvényt visszahelyettesítve

$$U'(r(t)) = U'(r(0)) e^{(\beta - \alpha)t}\,.$$

Ha például speciálisan a hasznossági függvényünk

$$U(r) = 2\sqrt{r}$$

alakú, akkor azt kapjuk, hogy

$$\alpha x(t) - x'(t) = r(t) = r(0) e^{2(\alpha - \beta)t}$$

a $[0, T]$ intervallumon. Az x függvény ennek a lineáris differenciálegyenletnek a megoldásával nyerhető.

Felhívjuk a figyelmet arra, hogy az r felhasználás monoton növő, ha az α kamatláb nagyobb a β diszkont faktornál, míg ellenkező esetben monoton fogyó. Természetesen konstans felhasználást kapunk, ha a két állandó megegyezik.

6.6. A Ramsey-féle növekedési modell

6.14 Példa. **(A gazdasági növekedés Ramsey-féle modellje, 1928)** Vizsgáljuk meg egy homogén, egyszektoros gazdaság működését a $[0, T]$ időintervallumon.

A gazdaság $t \in [0, T]$ időpontbeli kibocsátását jelölje $Y(t)$, a fogyasztását $C(t)$, a beruházásásra fordított hányadot pedig $I(t)$. Ekkor a következő összefüggés áll fenn:

$$Y(t) = C(t) + I(t) \tag{6.6}$$

A gazdaságban a $t \in [0, T]$ időpontban rendelkezésre álló tőkemennyiség legyen $K(t)$, és jelentse $L(t) > 0$ a rendelkezésre álló munkamennyiséget ebben az időpontban.

Feltesszük a továbbiakban, hogy a munkamennyiség exponenciális növekedést követ, azaz

$$\frac{L'(t)}{L(t)} = \nu > 0$$

adott állandó.

A beruházás mértékét az

$$I(t) = K'(t) + \mu \cdot K(t) \tag{6.7}$$

egyenlet határozza meg ahol a $\mu > 0$ paraméter az amortizációs normát jelenti.

A gazdaság bruttó kibocsátását egy a tőkétől és a munkától függő $F(K, L)$ termelési függvénnyel származtatjuk, határozza meg, nevezetesen

$$Y(t) = F(K(t), L(t)) \,. \tag{6.8}$$

a $[0, T]$ időintervallum minden pontjában. Feltesszük, hogy a termelési függvény eleget tesz a szokásos feltételeknek, azaz az F függvény

1. pozitív homogén: minden $\lambda > 0$ esetén $F(\lambda K, \lambda L)) = \lambda F(K, L)$.
2. szigorúan monoton növő,
3. konkáv,
4. érvényesek a következő relációk:

$$\lim_{(K,L)\to(0,0)+} \partial_1 F(K, L) = \infty \,, \qquad \lim_{(K,L)\to(0,0)+} \partial_2 F(K, L) = \infty \,,$$
$$\lim_{(K,L)\to(\infty,\infty)} \partial_1 F(K, L) = 0 \,, \qquad \lim_{(K,L)\to(\infty,\infty)} \partial_2 F(K, L) = 0 \,.$$

A (6.6), (6.7) és (6.8) összefüggések együttesen azt jelentik, hogy

$$C(t) + K'(t) + \mu \cdot K(t) = F(K(t), L(t)) \,.$$

A kezelhetőbb alak érdekében osszuk el az egyenletet a $L(t)$ pozitív számmal:

$$\frac{C(t)}{L(t)} + \frac{K'(t)}{L(t)} + \mu \frac{K(t)}{L(t)} = F\left(\frac{K(t)}{L(t)}, 1\right) , \tag{6.9}$$

ahol felhasználtuk, hogy F pozitív homogén. Jelölje továbbá

$$c(t) \frac{C(t)}{L(t)}$$

az egy főre jutó átlagfogyasztást, valamint

$$k(t) = \frac{K(t)}{L(t)}$$

, az egy főre jutó átlagos tőkét (tőkefelszereltséget) a t időpontban. Vezessük be az

$$f(k) = F(k, 1) = F\left(\frac{K}{L}, 1\right)$$

módosított termelési függvényt.

Mivel

$$k'(t) = \left(\frac{K(t)}{L(t)}\right)' = \frac{K'(t)}{L(t)} - k(t)\frac{L'(t)}{L(t)} = \frac{K'(t)}{L(t)} - \nu k(t) ,$$

ezért a (6.9) összefüggés a

$$c(t) + k'(t) + \nu \cdot k(t) + \mu \cdot k(t) = f(k(t)) ,$$

alakot ölti. Ha bevezetjük a $\lambda = \mu + \nu$ paramétert, a következő differenciálegyenlethez jutunk:

$$k'(t) = f(k(t)) - \lambda \cdot k(t) - c(t) ,$$

amelyet a gazdasági növekedés differenciálegyenletének nevezünk. Tegyük fel, hogy teljesülnek a $k(0) = k_0$ és $k(T) = k_T$ peremfeltételek.

A gazdaságot egy, a fogyasztástól függő u hasznossági függvény jellemzi, amelyről feltesszük, hogy kétszer differenciálható, és rendelkezik a szokásos tulajdonságokkal, azaz

$$u'(c) > 0 \quad \text{és} \quad u''(c) < 0$$

valamint

$$\lim_{c \to 0+} u'(c) = +\infty \quad \text{és} \lim_{c \to +\infty} u'(c) = 0 .$$

Válasszuk meg a fogyasztást leíró c függvényt úgy, hogy az

$$\int_0^T e^{-\beta t} u(c(t))\, dt$$

összhasznosság maximális legyen, ahol $\beta > 0$ a diszkont kamatlábat jelenti.

Ha a differenciálegyenletből az átlagfogyasztást kifejezzük:

$$c(t) = f(k(t)) - \lambda k(t) - k'(t)\,, \tag{6.10}$$

és ezt az integrálba behelyettesítjük, akkor a következő Lagrange-feladatot kapjuk:

$$\int_0^T e^{-\beta t} u(f(k(t)) - \lambda k(t) - k'(t))\,dt \to \max\,, \tag{6.11}$$

ahol a peremfeltételek $k(0) = k_0$ és $k(T) = k_T$.

A feladat Euler-Lagrange-differenciálegyenlete:

$$\frac{d}{dt}(e^{-\beta t}u'(f(k(t)) - \lambda k(t) - k'(t))) = e^{-\beta t}u'(f(k(t)) - \lambda k(t) - k'(t))(\lambda - f'(k(t))).$$

A (6.10) reláció alapján ez rövidebben úgy írható, hogy

$$-\frac{d}{dt}(e^{-\beta t}u'(c(t)) = e^{-\beta t}u'(c(t))(f'(k(t)) - \lambda)\,.$$

Elvégezve a deriválást azt kapjuk, hogy:

$$-\beta t e^{-\beta t}u'(c(t)) + e^{-\beta t}u''(c(t))c'(t) = e^{-\beta t}u'(c(t))(f'(k(t)) - \lambda)\,,$$

amiből

$$c'(t) = \frac{u'(c(t))}{u''(c(t))}(\beta + \lambda - f'(k(t)))\,. \tag{6.12}$$

A hasznossági függvény tulajdonságait figyelembe véve könnyen látható, hogy a fogyasztás pontosan akkor monoton növő, ha

$$f'(k(t)) > \beta + \lambda = \beta + \mu + \nu\,.$$

A (6.12) egyenlet mindkét oldalát a $c(t)$-vel osztva azt kapjuk, hogy

$$\frac{c'(t)}{c(t)} = \frac{u'(c(t))}{c(t)u''(c(t))}(\beta + \lambda - f'(k(t)))\,,$$

ahol az $u'(c(t))/c(t)u''(c(t))$ együttható éppen az u' függvény elaszticitásának reciproka.

Végül határozzuk meg az optimális k függvényt. A (6.10) egyenletből

$$c'(t) = f'(k(t))k'(t) - \lambda k'(t) - k''(t)\,,$$

ezt a (6.12) egyenletbe helyettesítve az alábbi másodrendű differenciálegyenlethez jutunk:

$$k''(t) = f'(k(t))k'(t) + \frac{u'(c(t))}{u''(c(t))}(f'(k(t)) - \beta - \lambda)\,. \tag{6.13}$$

A k függvény ismeretében a (6.10) egyenletből nyerjük a $c(t)$ átlagfogyasztást.

1. Speciális eset: Tekintsük először a $\lambda = \mu = \nu = 0$ legegyszerűbb esetet, azaz az amortizációs és a növekedési állandó is zérus. Ekkor a (6.13) egyenlet:

$$k''(t) = f'(k(t))k'(t) + \frac{u'(c(t))}{u''(c(t))}(f'(k(t)) - \beta)\,. \tag{6.14}$$

Ha itt speciálisan az

$$f(k) = mk \quad \text{és} \quad u(c) = \frac{c^{1-\gamma}}{1-\gamma}\,,$$

függvényeket tekintjük, ahol $m > 0$ és $0 < \gamma < 1$ adott állandók, akkor a (6.14) differenciálegyenlet a következő alakú lesz:

$$k''(t) = \left(m + \frac{m-\beta}{\gamma}\right)k'(t) - m\frac{m-\beta}{\gamma}k(t)\,.$$

Ennek általános megoldása

$$k(t) = \delta_1 e^{-mt} + \delta_2 e^{-\frac{m-\beta}{\gamma}t}\,,$$

ahol a δ_1 és δ_2 állandók a peremfeltételeket jelentő

$$\begin{aligned} k_0 &= \delta_1 + \delta_2 \\ k_T &= \delta_1 e^{-mT} + \delta_2 e^{-\frac{m-\beta}{\gamma}T} \end{aligned}$$

egyenletrendszerből határozhatók meg.

2. Speciális eset: Legyen most $\beta = 0$, továbbá legyen

$$f(k) = mk \quad \text{és} \quad u(c) = -\alpha(c - c_0)^2\,,$$

ahol m és α pozitív állandók, továbbá $c_0 \geq 0$ adott. Ekkor a (6.14) differenciálegyenletre

$$k''(t) = m^2 k(t) - mc_0\,.$$

adódik. Ennek általános megoldása

$$k(t) = \delta_1 \cdot e^{mt} + \delta_2 e^{-mt} + \frac{c_0}{m}\,,$$

ahol a δ_1 és δ_2 állandók újra a

$$\begin{aligned} k_0 &= \delta_1 + \delta_2 \\ k_T &= \delta_1 e^{mT} + \delta_2 e^{-mT} + \frac{c_0}{m} \end{aligned}$$

egyenletrendszerből számíthatók ki.

6.7. Monopólium árazási problémája

6.15 Példa. (Árazási feladat) Egy monopólium a $[0,T]$ időszakban egyféle árut termel, és maximális profitra törekszik.

A keresletet a $t \in [0,T]$ időpontban az áru $p(t)$ árától és annak $p'(t)$ megváltozásától függő $d(p(t)), p'(t))$ keresleti függvény szabályozza, ezért $p(t)$ ár mellett a monopólium termelése (kibocsátása)

$$q(t) = d(p(t)), p'(t)). \tag{6.15}$$

Tegyük fel, hogy q mennyiségű áru termelésének a költsége $c(q)$. Eszerint a monopólium profitja a $t \in [0,T]$ időpontban

$$p(t) \cdot q(t) - c(q(t)) = p(t) \cdot d(p(t)), p'(t)) - c(d(p(t)), p'(t))).$$

Ennek megfelelően a monopólium feladata olyan árfüggvény kialakítása, amely mellett a $[0,T]$ időszakbeli

$$\int_0^T p(t) \cdot d(p(t)), p'(t)) - c(d(p(t)), p'(t)))\, dt.$$

profitja maximális. Tegyük fel, hogy a kezdeti ár p_0, a T időpontbeli pedig p_T.

A fentiek a következő variációszámítási feladatra vezetnek:

$$\begin{cases} \int_0^T p(t) \cdot d(p(t)), p'(t)) - c(d(p(t)), p'(t)))\, dt \to \max \\ p(0) = p_0,\ p(T) = p_T \end{cases}. \tag{6.16}$$

A feladat megoldása:

A feladat Euler-Lagrange-differenciálegyenlete:

$$\begin{aligned} & \frac{d}{dt}\partial_3(p(t) \cdot d(p(t)), p'(t)) - c(d(p(t)), p'(t)))) \\ = \quad & \partial_2(p(t) \cdot d(p(t)), p'(t)) - c(d(p(t)), p'(t)))), \end{aligned}$$

a parciális deriválásokat elvégezve:

$$\begin{aligned} & \frac{d}{dt}(p(t) \cdot \partial_2 d(p(t)), p'(t)) - c'(d(p(t)), p'(t))) \cdot \partial_2 d(p(t)), p'(t)) \\ = \quad & d(p(t)), p'(t)) + p(t) \cdot \partial_1 d(p(t)), p'(t)) - c'(d(p(t)), p'(t)) \cdot \partial_1 d(p(t)), p'(t)), \end{aligned}$$

felhasználva a (6.15) összefüggés ez rövidebben

$$\begin{aligned} & \frac{d}{dt}(p(t) \cdot \partial_2 d(p(t)), p'(t)) - c'(q(t)) \cdot \partial_2 d(p(t)), p'(t)) \\ = \quad & q(t) + p(t) \cdot \partial_1 d(p(t)), p'(t)) - c'(q(t)) \cdot \partial_1 d(p(t)), p'(t)). \end{aligned} \tag{6.17}$$

Speciális eset: Tegyük fel, hogy a költségfüggvény valamely

$$c(q) \doteq \alpha q^2 + \beta q + \gamma$$

másodfokú polinom, ahol $\alpha, \beta, \gamma > 0$ számok, a keresleti függvény pedig

$$d(p, p') \doteq -Ap + Bp' + C$$

alakú, ahol $A, B, C > 0$ számok. Mivel ekkor

$$\begin{aligned}
&c'(q) = 2\alpha q + \beta, \\
&\partial_1 d(p, p') = -A, \ \partial_2 d(p, p') = B, \\
&q(t) = -Ap(t) + Bp'(t) + C \\
&q'(t) = -Ap'(t) + Bp''(t),
\end{aligned}$$

ezért a (6.16) egyenlet ebben az esetben a következő:

$$\begin{aligned}
\frac{d}{dt}(p(t) \cdot B - (2\alpha q(t) + \beta) \cdot B) &= q(t) + p(t) \cdot (-A) - 2\alpha q(t) \cdot (-A), \\
Bp'(t) - 2\alpha Bq'(t)) &= -Ap(t) + (2\alpha A + 1)q(t),
\end{aligned}$$

amibe $q(t)$-t és $q'(t)$-t behelyettesítve a következő másodrendű differenciálegyenletet kapjuk:

$$p''(t) = \frac{A(\alpha A + 1)}{\alpha B} \cdot p(t) + \frac{\alpha A + C}{2\alpha B},$$

azaz

$$p''(t) = \mu \cdot p(t) + \nu,$$

ahol $\mu \doteq \frac{A(\alpha A+1)}{\alpha B}$ és $\nu \doteq \frac{\alpha A+C}{2\alpha B}$.

6.8. Gyakorlatok

1. Tekintsük az alábbi Lagrange-feladatot:

   $$F(x) = \int_0^1 x'(t)^2 \, dt \to \min, \quad x_0 = x_1 = 0\,,$$

   és igazoljuk, hogy $x = 0$ az egyetlen minimumhely.

2. Oldjuk meg az alábbi Lagrange-feladatot:

   $$\int_0^2 (4 - 3x(t)^2 - 16x'(t) - 4x'(t)^2) \, dt \to \max\,,$$

   ahol

   $$x(0) = -\frac{8}{3} \quad \text{és} \quad x(2) = \frac{1}{3}\,.$$

   Az elégséges feltétellel igazoljuk, hogy az Euler-Lagrange-egyenlet valóban megoldást ad.

3. Vizsgáljuk meg az

$$\int_0^1 x'(t)^3\,dt \to \min, \quad x(0)=0 \quad \text{és} \quad x(1)=1$$

Lagrange-feladatot. Igazoljuk, hogy az Euler-Lagrange-egyenletnek végtelen sok megoldása van, de csak $x(t)=t$ elégíti ki a peremfeltételeket.

4. Vizsgáljuk meg az

$$F(x)=\int_0^1 t^2 x'(t)^2\,dt \to \min, \quad x_0=0, \quad x_1=1$$

Lagrange-feladatot és mutassuk meg, hogy az Euler-Lagrange egyenletnek nincs megoldása.

5. Igazoljuk közvetlenül, az Euler-Lagrange egyenlet vizsgálata nélkül, hogy az előző gyakorlatban a Lagrange-feladatnak nem lehet megoldása. Tekintsük az

$$x_n(t)=\begin{cases} nt & \text{ha } 0\le t\le \frac{1}{n} \\ 1 & \text{ha } \frac{1}{n}<t\le 1 \end{cases}$$

függvénysorozatot. Mutassuk meg, hogy itt $F(x_n)\to 0$, és ugyanakkor $F(x)>0$ minden $x\in W^\infty[0,1]$ mellett.

6. Keressük meg az x_0 és az x_1 pontokat összekötő legrövidebb görbét. Nevezetesen oldjuk meg az alábbi variációs feladatot:

$$F(x)=\int_0^1 \sqrt{1+x'(t)^2}\,dt \to \min,$$

ahol $x(0)=x_0$ és $x(1)=x_1$.

7. Mutassuk meg, hogy az

$$F(x)=\int_0^1 t^{\frac{2}{3}} x'(t)^2\,dt \to \min, \quad x_0=0\,, \quad x_1=1$$

Lagrange-feladatnak nincs megoldása a $W^\infty[0,1]$ térben.

8. Ellenőrizzük, hogy az

$$F(x)=\int_{-1}^1 x(t)^2(1-x'(t))^2\,dt \to \min, \quad x_0=0\,, \quad x_1=1$$

Lagrange-feladatnak végtelen sok megoldása van. Ezek közül azonban csak egy optimális, méghozzá

$$x(t)=\begin{cases} 0, & \text{ha } -1\le t\le 0 \\ t, & \text{ha } 0<t\le 1\,. \end{cases}$$

Folytonosan differenciálható megoldás azonban nem létezik.

9. Adott egy $a \in L^2[0,1]$ függvény. Keressünk olyan legfeljebb n-edfokú p polinomot, amelyre $\int_0^1 p(t)\,dt = 0$, és amelyre
$$\int_0^1 |a(t) - p(t)|^2\,dt$$
minimális. Igazoljuk, hogy a feladat egyértelműen megoldható.

10. Mutassuk meg, hogy az előző feladat megoldható úgy, hogy először előállítunk egy olyan legfeljebb n-edfokú q polinomot, amelyre
$$\int_0^1 |a(t) - q(t)|^2\,dt$$
minimális, majd ezután megkeressük azt a legfeljebb n-edfokú p polinomot, amelyre $\int_0^1 p(t)\,dt = 0$, továbbá az
$$\int_0^1 |q(t) - p(t)|^2\,dt$$
integrál minimális.

11. Igazoljuk az $x(0)$ szabad kezdőpont esetére a (6.5) transzverzalitási feltételt.

12. Oldjuk meg az alábbi Lagrange-feladatot (Ügyeljünk a transzverzalitási feltételre és az elégséges feltételre!)
$$\int_0^1 (tx' + (x')^2)\,dt \to \min \qquad x(0) = 1 \qquad x(1) \text{ szabad}$$

13. Oldjuk meg a következő variációs feladatot:
$$\int_0^1 (1 - x(t)^2 - x'(t)^2)\,dt \to \max$$
ahol
$$x(0) = 1 \quad \text{és} \quad x(1) \text{ szabad.}$$
Ügyeljünk a transzverzalitási feltételre, és használjuk az elégséges feltételt.

7. fejezet

Lagrange-multiplikátorok

Ebben a fejezetben egy absztrakt, függvényterekben felírt feltételes szélsőérték-feladatot vizsgálunk meg, és a feladat megoldásának szükséges feltételeként megfogalmazzuk a Lagrange-féle multplikátorelvet.

A tárgyalásban erősen támaszkodunk a funkcionálanalízis[1]. alapvető fogalmaira.

7.1. Faktorterek

Legyen az alábbiakban X valós normált tér, és jelentse L az X valamely zárt alterét. Vezessük be a következő jelölést:

$$X/L = \{x + L : x \in X\}.$$

Az X/L halmazon az alábbi műveleteket definiáljuk

$$\begin{aligned}(x+L)+(y+L) &= (x+y)+L \\ \lambda(x+L) &= \lambda x + L.\end{aligned}$$

Könnyen ellenőrizhető, hogy X/L ezekkel a műveletekkel vektorteret alkot a valós test fölött (lásd az 1. gyakorlatot).

Adott $x + L \in X/L$ esetén vezessük be a következő normát:

$$\|x + L\| = \inf_{u \in L} \|x + u\|. \tag{7.1}$$

Egyszerűen belátható, hogy ez az egyenlőség valóban normát definál (lásd a 2. gyakorlatot.

7.1 Definíció. A fent definiált X/L valós normált teret az X tér L altér szerinti *faktorterének* nevezzük.

[1]Lásd például *Kánnai Zoltán: Analitikus módszerek a pénzügyben és a közgazdaságtanban, www.tankonyvtar.hu, 2013.*

7.2 Példa. Legyenek X és Y valós normált terek, és tekintsünk egy olyan $A \in L(X,Y)$ folytonos lineáris leképezést, amely ráképezés, azaz

$$\operatorname{im} A = Y\,.$$

Ilyenkor $\ker A$ az X zárt altere, továbbá minden $y \in Y$ mellett

$$A^{-1}(y) \in X/\ker A\,.$$

Ez azt jelenti, hogy az az A^{-1} lineáris leképezés, amelyre

$$A^{-1} : Y \to X/\ker A\,, \quad y \mapsto A^{-1}(y)$$

algebrai izomorfizmus. Valóban, egyrészt a linearitás nyilvánvaló, másrészt ha $A^{-1}(y_1) = A^{-1}(y_2)$, akkor találhatók olyan x_1 és x_2 X-beli vektorok, amelyekre $Ax_1 = y_1$, és $Ax_2 = y_2$, és ezért

$$x_1 + \ker A = x_2 + \ker A\,.$$

Innen az adódik, hogy $x_1 - x_2 \in \ker A$, azaz $y_1 = y_2$.

Ezt az A^{-1} leképezést az A *általánosított inverz operátorának* nevezzük (jóllehet az A maga nem feltétlenül kölcsönösen egyértelmű, így az eredeti értelemben nem invertálható).

7.3 Állítás. *Ha X és Y Banach-terek, és $A \in L(X,Y)$ ráképezés, akkor az*

$$A^{-1} : Y \to X/\ker A$$

inverz operátor folytonos izomorfizmus.

Bizonyítás. A Banach-féle nyílt-leképezés tétel értelmében található olyan δ pozitív szám, hogy

$$\delta B_Y \subset A(B_X)\,,$$

ahol B_Y és B_X az Y, illetve X terek zárt egységgömbjeit jelentik.

Ha most $y \in Y$ nem zérus vektor, akkor létezik olyan $x \in X$, amelyre $Ax = y$, valamint olyan $v \in B_X$ vektor, hogy

$$Av = \frac{\delta}{\|y\|}y\,.$$

Ekkor egyrészt $A^{-1}(y) = x + \ker A$, másrészt

$$\|A^{-1}(y)\| = \inf_{u \in \ker A} \|x+u\| = \inf_{Aw=y} \|w\| \le \frac{\|y\|}{\delta}\|v\| \le \frac{\|y\|}{\delta}\,.$$

Innen azonnal adódik, hogy A^{-1} folytonos, és $\|A^{-1}\| \le 1/\delta$. $\square$

7.2. Az általánosított inverzfüggény-tétel

Az alábbi tételünk a klasszikus inverzfüggvény-tétel[2] általánosításának tekinthető abban az értelemben, hogy az inverzfüggvény létezésének igazolásához nem tesszük fel, hogy a derivált egy pontban invertálható legyen. Pusztán annyit teszünk fel, hogy a derivált ráképezés, nem feltétlenül invertálható. Igy azt kapjuk, hogy a leképezés környezetet környezetre képez.

7.4 Tétel. **(Inverzfüggvény-tétel)** *Legyenek X és Y Banach-terek, és tekintsünk egy olyan $H : X \to Y$ folytonosan differenciálható függvényt, amelyre az x_0 pontban a derivált ráképezés, azaz*

$$\operatorname{im} H'(x_0) = Y\,.$$

Ekkor a $H(x_0)$ pontnak található olyan V környezete, és olyan $\alpha > 0$ szám, hogy minden $y \in V$ ponthoz létezik olyan $x \in X$, amelyre

$$\begin{aligned} H(x) &= y \\ \|x - x_0\| &\leq \alpha \|y - H(x_0)\| \end{aligned}$$

Bizonyítás. Az elősző szakasz megállapításai alapján az $A = H'(x_0)$ jelöléssel az $A : Y \to X/\ker A$ leképezés folytonos izomorfizmus. Legyen $\alpha = 4\|A^{-1}\|$.

A H' függvény x_0-beli folytonossága miatt az $\varepsilon = 1/\alpha$ pozitív számhoz található olyan $\delta > 0$, hogy minden $x \in x_0 + \delta B$ pontra

$$\|H'(x)\| - A\| \leq \frac{1}{\alpha}\,.$$

Vezessük be a

$$V = H(x_0) + \frac{\delta}{\alpha} B_Y$$

jelölést, és rögzítsünk egy tetszőleges $y \in V$ vektort. Értelmezzünk indukcióval egy u_n X-beli sorozatot a következő módon.

Első lépésként az $A^{-1}(y - H(x_0))$ halmaznak válasszunk ki egy olyan u_1 elemét, amelyre

$$\|u_1\| \leq 2\|A^{-1}(y - H(x_0))\| \leq 2\|A^{-1}\| \cdot \|y - H(x_0)\| \leq 2\frac{\alpha}{4}\frac{\delta}{\alpha} = \frac{\delta}{2}\,.$$

Ezután tegyük fel, hogy valamely $n \geq 2$ indexre az $u_1, \ldots, u_{n-1}$ vektorokat már úgy definiáltuk, hogy

$$\|u_k\| \leq \frac{1}{2^{k-1}}$$

[2]lásd: *Tallos Péter: Dinamikai rendszerek alapjai, Aula, 1999.*

bármely $k = 1, \ldots, n-1$ indexre. Világos, hogy ilyenkor

$$\sum_{k=1}^{n-1} \|u_k\| \leq 2\|u_1\| \leq \delta\,.$$

Válasszuk ki most az

$$A^{-1}\left(y - H\left(x_0 + \sum_{k=1}^{n-1} u_k\right)\right) \in X/\ker A$$

affin halmaznak egy olyan u_n elemét, amelyre

$$\|u_n\| \leq \left\|A^{-1}\left(y - H\left(x_0 + \sum_{k=1}^{n-1} u_k\right)\right)\right\|\,.$$

A választás miatt természetesen

$$Au_n = y - H\left(x_0 + \sum_{k=1}^{n-1} u_k\right) \tag{7.2}$$

és ezért az inverzképre térve (itt használjuk ki, hogy A ráképezés)

$$u_n + \ker A = A^{-1}\left(y - H\left(x_0 + \sum_{k=1}^{n-1} u_k\right)\right)\,. \tag{7.3}$$

Innen egyrészt

$$\|u_n + \ker A\| = \left\|A^{-1}\left(y - H\left(x_0 + \sum_{k=1}^{n-1} u_k\right)\right)\right\| \geq \frac{1}{2}\|u_n\|\,, \tag{7.4}$$

másrészt (7.3)-vel teljesen analóg módon n helyett $n-1$-re

$$u_{n-1} + \ker A = A^{-1}\left(y - H\left(x_0 + \sum_{k=1}^{n-2} u_k\right)\right)\,. \tag{7.5}$$

A (7.3) és (7.5) egyenlőségek alapján

$$\begin{aligned} u_n - u_{n-1} + \ker A &= u_n + \ker A - (u_{n-1} + \ker A) \\ &= A^{-1}\left(H\left(x_0 + \sum_{k=1}^{n-2} u_k\right) - H\left(x_0 + \sum_{k=1}^{n-1} u_k\right)\right)\,, \end{aligned}$$

tehát más módon

$$u_n + \ker A = u_{n-1} + A^{-1}\left(H\left(x_0 + \sum_{k=1}^{n-2} u_k\right) - H\left(x_0 + \sum_{k=1}^{n-1} u_k\right)\right)\,.$$

Megmutatjuk, hogy az így definiált u_n vektorra $\|u_n\| \leq \frac{1}{2}\|u_{n-1}\|$.

Valóban, a

$$\hat{H}(x) = x - A^{-1}(H(x)) = A^{-1}(Ax - H(x))$$

folytonosan differenciálható függvény deriváltja $\hat{H}'(x) = A^{-1}(A - H'(x))$, és így a Lagrange-féle középértéktételt (lásd az 5.11 Tételt) alkalmazva az

$$\left[x_0 + \sum_{k=1}^{n-2} u_k, x_0 + \sum_{k=1}^{n-1} u_k\right] \subset x_0 + \delta B_X$$

halmazon azt kapjuk, hogy

$$\begin{aligned}
\frac{1}{2}\|u_n\| &\leq \left\|A^{-1}\left(y - H\left(x_0 + \sum_{k=1}^{n-1} u_k\right)\right)\right\| = \|u_n + \ker A\| \\
&= \left\|u_{n-1} + A^{-1}\left(H\left(x_0 + \sum_{k=1}^{n-2} u_k\right) - H\left(x_0 + \sum_{k=1}^{n-1} u_k\right)\right)\right\| \\
&= \left\|\hat{H}\left(x_0 + \sum_{k=1}^{n-1} u_k\right) - \hat{H}\left(x_0 + \sum_{k=1}^{n-2} u_k\right)\right\| \\
&\leq \sup_{\left[x_0 + \sum_{k=1}^{n-2} u_k, x_0 + \sum_{k=1}^{n-1} u_k\right]} \|\hat{H}'(x)\| \cdot \|u_{n-1}\| \\
&\leq \sup_{\|x - x_0\| \leq \delta} \|A^{-1}(A - H'(x))\| \cdot \|u_{n-1}\| \leq \|A^{-1}\| \frac{1}{\alpha} \|u_{n-1}\| \\
&\leq \frac{1}{4}\|u_{n-1}\|\,.
\end{aligned}$$

Ebből következik, hogy valóban

$$\|u_n\| \leq \frac{1}{2}\|u_{n-1}\|\,,$$

és az indukciós eljárást is figyelembe véve

$$\|u_n\| \leq \frac{1}{2^{n-1}}\|u_1\|\,.$$

Innen a végtelen sor összegére azt kapjuk, hogy

$$\left\|\sum_{k=1}^{\infty} u_k\right\| \leq 2\|u_1\| \leq \delta\,,$$

továbbá az $n \to \infty$, $u_n \to 0$ határátmenet azt eredményezi, hogy

$$0 = y - H\left(x_0 + \sum_{k=1}^{\infty} u_k\right)\,.$$

Tehát az

$$x = x_0 + \sum_{k=1}^{\infty} u_k$$

választással az adódik, hogy $H(x) = y$, valamint

$$\|x - x_0\| = \left\| \sum_{k=1}^{\infty} u_k \right\| \leq 2\|u_1\| \leq 4\|A^{-1}\| \cdot \|y - H(x_0)\| = \alpha \|y - H(x_0)\| ,$$

amit igazolnunk kellett. □

7.3. Az ortogonalitási tétel

Ebben a szakaszban igazoljuk az 1.25 ortogonalitási tétel Banach-terekre érvényes változatát. A teljesség kedvéért először néhány elnevezést vezetünk be. A Banach-terek geometriájának részletesebb tárgyalását illetően itt utalunk a korábban javasolt tankönyvre[3].

Ha X valamely valós Banach-tér, akkor az X tér X^* dualisán az összes X-en értelmezett folytonos lineáris funkcionál terét értjük. Ismert, hogy X^* Banach-teret alkot a

$$\|p\| = \sup\{|\langle p, x\rangle| : x \in X,\ \|x\| = 1\}, \quad p \in X^*$$

normával ellátva. Ha L az X tér valamely altere, úgy legyen

$$L^{\perp} = \{p \in X^* : \langle p, x\rangle = 0\ \forall x \in L\}$$

az L annullátora. Világos, hogy $L^{\perp}$ zárt altér az X^* térben.

Legyenek X és Y valós Banach-terek, és tekintsünk egy $A \in L(X, Y)$ folytonos lineáris leképezést. Az A leképezés adjungáltján azt az $A^* : Y^* \to X^*$ leképezést értjük, amelyre

$$\langle A^* q, x\rangle = \langle q, Ax\rangle$$

minden $x \in X$ és $q \in Y^*$ mellett. Ilyenkor A^* folytonos lineáris leképezés, és $\|A^*\| = \|A\|$. Vegyük észre, hogy véges dimenziós X és Y vektorterek esetében ez a fogalom - rögzített mátrixreprezentáció mellett - éppen egybeesik a transzponált mátrix fogalmával.

7.5 Tétel. *Ha* $\operatorname{im} A = Y$, *akkor* $(\ker A)^{\perp} = \operatorname{im} A^*$.

Bizonyítás. Először legyen $p \in \operatorname{im} A^*$ tetszőleges. Akkor található olyan $q \in Y^*$, amelyre $p = A^* q$, és így bármely $x \in \ker A$ esetén

$$\langle p, x\rangle = \langle A^* q, x\rangle = \langle Ax, q\rangle = 0 .$$

[3]Kánnai Zoltán: *Analitikus módszerek a pénzügyben és a közgazdaságtanban, 2013. www.tankonyvtar.hu*

Ez éppen azt jelenti, hogy $p \in (\ker A)^{\perp}$, azaz $\operatorname{im} A^* \subset (\ker A)^{\perp}$.

Fordítva, tegyük most fel, hogy $p \in (\ker A)^{\perp}$. Adott $y \in \operatorname{im} A$ mellett válasszunk egy $x \in X$ elemet, amelyre $Ax = y$, és értelmezzük az f függvényt az

$$f(y) = \langle p, x \rangle$$

formulával. Könnyen látható, hogy f jól definiált, hiszen ha az $\bar{x}$ vektorra is $A\bar{x} = y$, akkor $x - \bar{x} \in \ker A$, és ezért $\langle p, x - \bar{x} \rangle = 0$. Persze f lineáris függvény, amely a 7.4 Tétel A-ra való alkalmazásával folytonos is, emiatt valamely $q \in Y^*$-ra

$$f(y) = \langle q, y \rangle$$

minden $y \in Y$ mellett. Ezért

$$\langle A^* q, x \rangle = \langle q, Ax \rangle = \langle q, y \rangle = \langle p, x \rangle$$

bármely $x \in X$ vektorra. Tehát $p = A^* q$, így $(\ker A)^{\perp} \subset \operatorname{im} A^*$, amit bizonyítanunk kellett. □

7.4. A Lagrange-elv

Tekintsük a következő absztrakt feltételes szélsőérték-feladatot. Legyenek X és Y valós Banach-terek, továbbá $F : X \to \mathbb{R}$ és $G : X \to Y$ folytonosan differenciálható függvények.

Keressük az

$$\begin{cases} F(x) \to \min \\ G(x) = 0 \end{cases} \tag{7.6}$$

feltételes szélsőérték feladat megoldását.

7.6 Lemma. *Legyen x_0 a (7.6) feladat megoldása, és tegyük még fel, hogy $\operatorname{im} G'(x_0) = Y$. Akkor*

$$\ker G'(x_0) \subset \ker F'(x_0)\,.$$

Bizonyítás. Indirekt módon tegyük fel, hogy létezik olyan $u \in \ker G'(x_0)$, amelyre $F'(x_0)u \neq 0$, és tekintsük azt a $H : X \to \mathbb{R} \times Y$ leképezést, amelyre

$$H(x) = (F(x), G(x))\,.$$

A feltételünk szerint ekkor

$$\operatorname{im} H'(x_0) = \mathbb{R} \times Y\,.$$

Valóban, tetszőleges $\alpha \in \mathbb{R}$ és $y \in Y$ esetén van olyan $v \in X$, hogy $G'(x_0)v = y$, ezért az

$$x = \frac{\alpha - F'(x_0)v}{F'(x_0)u} u + v$$

vektorra könnyen látható, hogy $F'(x_0)x = \alpha$ és $G'(x_0)x = y$, azaz $H'(x_0)x = (\alpha, y)$.

Tehát a 7.4 Tétel alapján a $H(x_0)$ pontnak van olyan V környezete, hogy minden $(\alpha, y) \in V$ ponthoz található olyan $x \in X$ vektor, amelyre $H(x) = (\alpha, y)$. Ha most a δ pozitív szám elegendően kicsi, akkor

$$(F(x_0) - \delta, 0) \in V\,,$$

Ezért alkalmas x vektorra $H(x) = (F(x_0) - \delta, 0)$, nevezetesen

$$F(x) = F(x_0) - \delta\,, \quad \text{és} \quad G(x) = 0\,.$$

Ez ellentmond annak, hogy x_0 a (7.6) feladat megoldása. □

7.7 Tétel. (Lagrange-multiplikátorok) *Tegyük fel, hogy x_0 a (7.6) feladat megoldása, és $\operatorname{im} G'(x_0) = Y$. Akkor van olyan $q \in Y^*$ funkcionál, hogy*

$$F'(x_0) + qG'(x_0) = 0\,.$$

Bizonyítás. Az előző lemmánk és a 7.5 ortogonalitási tétel értelmében

$$F'(x_0) \in (\ker G'(x_0))^{\perp} = \operatorname{im} G'(x_0)^*\,.$$

Innen következik, hogy létezik olyan $q \in Y^*$, amelyre $F'(x_0) = -G'(x_0)^* q$. Ez az adjungált leképezés definíciója szerint éppen azt jelenti, hogy

$$F'(x_0) + qG'(x_0) = 0\,,$$

amit igazolnunk kellett. □

Ha bevezetjük a (7.6) feladat Lagrange-függvényét az alábbi formulával:

$$L(x, q) = F(x) + \langle q, G(x) \rangle\,,$$

akkor a tételünk állítása úgy is megfogalmazható, hogy ha x_0 a feladat megoldása, akkor van olyan $q \in Y^*$ funkcionál, amelyre a Lagrange-függvény deriváltja eltűnik az x_0 helyen, azaz

$$\partial_1 L(x_0, q) = 0\,.$$

Ebben a terminológiában a q funkcionált *Lagrange-multiplikátornak* nevezzük.

7.5. Az izoperimetrikus probléma

Ezt a feladatot Euler fogalmazta meg elsőként, és a probléma megoldása Euler és Lagrange munkássága révén nagyban hozzájárult a variációszámítás kifejlődéséhez. (*Izoperimetrikus = azonos kerületű.*)

7.8 Példa. Adottak a síkon a $(-1, 0)$ és $(1, 0)$ koordinátájú pontok. Keressünk olyan, a $[-1, 1]$ intervallumon értelmezett folytonosan differenciálható x függvényt a síkon, amely grafikonjának ívhossza egy adott c pozitív szám, végpontjai az adott pontok, továbbá a grafikon és a vízszintes tengely által közrezárt terület a lehető legnagyobb.

A probléma Lagrange-feladatként a következő alakban fogalmazható meg.

$$\int_{-1}^{1} x(t)\,dt \to \max$$
$$\int_{-1}^{1} \sqrt{1+x'(t)^2}\,dt = c$$

Mivel a korábbi terminológiánknak megfelelően itt $Y=\mathbb{R}$, ezért a q funkcionál egy valós számmal történő szorzás lesz. Ezért a feltételes szélsőérték-feladat Lagrange-függvénye az alábbi alakban írható föl:

$$L(x,q) = \int_{-1}^{1} (x(t) + q\sqrt{1+x'(t)^2})\,dt\,.$$

Innen a $\partial_1 L(x,q)=0$ egyenlet az Euler-Lagrange-egyenlet formájában a következő alakot ölti:

$$1 - q\frac{d}{dt}\frac{x'(t)}{\sqrt{1+x'(t)^2}} = 0\,,$$

vagy ekvivalens megfogalmazásban

$$\frac{x'(t)}{\sqrt{1+x'(t)^2}} = \frac{1}{q}t + \mathrm{const}\,.$$

Könnyen ellenőrizhető, hogy a fenti differenciálegyenlet megoldása a

$$(t-a)^2 + (x-b)^2 = r^2$$

alakhoz vezet, ahol az a, b és r paramétereket úgy határozzuk meg, hogy azok kielégítsék a peremfeltételeket és az ívhosszra vonatkozó feltételt. Tehát a maximális területet biztosító grafikon egy körvonal.

7.6. Gyakorlatok

1. Mutassuk meg, hogy a faktortér-műveletek kielégítik a vektortér axiómáit.

2. Igazoljuk, hogy a (7.1) egyenlőség normát értelmez az X/L téren.

3. Legyen $k : [0,1]\times[0,1] \to \mathbb{R}$ olyan négyzetesen Lebesgue-integrálható függvény, tehát
$$\int_0^1 \int_0^1 |k(t,s)|^2\,ds\,dt < \infty\,,$$

és tekintsük azt az $A \in L(L^2[0,1])$ leképezést, amelyre minden $x \in L^2[0,1]$ és minden $t \in [0,1]$ mellett

$$(Ax)(t) = \int_0^1 k(t,s)x(s)\,ds\,.$$

Mutassuk meg, hogy A folytonos lineáris leképezés, továbbá ellenőrizzük, hogy bármely $y \in L^2[0,1]$ esetén

$$(A^*y)(t) = \int_0^1 k(s,t)y(s)\,ds\,,$$

ahol $t \in [0,1]$.

4. Legyen k újra az előző gyakorlatban vizsgált függvény, és értelmezzük most az A leképezést az $L^2[0,1]$ téren az

$$(Ax)(t) = \int_0^t k(t,s)x(s)\,ds$$

egyenlőséggel, ahol $t \in [0,1]$. Igazoljuk itt is, hogy A folytonos lineáris leképezés, valamint

$$(A^*y)(t) = \int_t^1 k(s,t)y(s)\,ds$$

minden $y \in L^2[0,1]$ és $t \in [0,1]$ esetén.

5. Legyenek X és Y valós Banach-terek, és tekintsünk egy $A \in L(X,Y)$ folytonos lineáris leképezést. Ellenőrizzük, hogy a

$$\ker A = (\operatorname{im} A^*)^{\perp}$$

egyenlőség a ráképezési feltétel nélkül is érvényes. Ilyenkor persze $\operatorname{im} A^*$ nem feltétlenül zárt altér!

8. fejezet

Optimális irányítás

Ebben a fejezetben megmutatjuk, hogy az optimális irányítási feladat tekinthető úgy, mint egy alkalmasan választott függvénytéren megfogalmazott Lagrange-feladat.

8.1. Az irányítási feladat

Legyenek X és Y n illetve m-dimenziós valós euklideszi terek, és tekintsük az

$$f : X \times Y \to \mathbb{R} \quad \text{illetve} \quad g : X \times Y \to X$$

folytonosan differenciálható függvényeket. Az X teret állapottérnek, az Y teret pedig irányítási térnek nevezzük.

Legyen adott egy $[0, T]$ időintervallum, továbbá értelmezzük a megengedett irányítások halmazát az

$$\hat{U} = \{u : [0, T] \to Y, u \in L^2[0, T]\}$$

egyenlőséggel. Tegyük fel, hogy adottak az x_0 és x_T vektorok az X térben, és vizsgáljuk az alábbi dinamikai rendszert $[0, T]$ intervallumon

$$\begin{aligned} x'(t) &= g(x(t), u(t)), \quad u \in \hat{U} \\ x(0) &= x_0, \quad x(T) = x_T \end{aligned} \tag{8.1}$$

A feltételeink alapján nyilvánvaló, hogy az $u \in \hat{U}$ irányítás megválasztásával a (8.1) rendszernek legfeljebb egy megoldása van. Nem biztos azonban, hogy a kezdetiérték-feladat megoldása kielégíti a végpontban megadott feltételt is.

Keresendő olyan $u \in \hat{U}$ irányítás, amelyre (x, u) megoldása a (8.1) nemlineáris rendszernek, továbbá

$$\int_0^T f(x(t), u(t))\, dt \to \min . \tag{8.2}$$

8.1 Definíció. A (8.1), (8.2) relációkkal értelmezett feladatot *optimális irányítási feladatnak* nevezzük.

Az (x, u) függvénypárt *megengedett folyamatnak* nevezzük, ha $u \in \hat{U}$ megengedett irányítás, és (x, u) kielégíti a (8.1) nemlineáris rendszert.

Azt mondjuk továbbá, hogy (x, u) *optimális folyamat*, ha (x, u) olyan megengedett folyamat, amelyre a (8.2) integrál felveszi a minimális értékét.

Az alábbiakban az optimális irányítási feladatot átfogalmazzuk függvénytéren értelmezett feltételes szélsőérték-feladattá. Vezessük be a következő jelöléseket. Legyen F az a $C[0,T] \times L^2[0,T]$ téren értelmezett valós értékű függvény, amelyre

$$F(x, u) = \int_0^T f(x(t), u(t))\, dt\,,$$

továbbá legyen $G : C[0,T] \times L^2[0,T] \to X \times C[0,T]$,

$$G(x, u) = \left(x(T) - x_T, x(t) - x_0 - \int_0^t g(x(s), u(s))\, ds \right)$$

Ezekkel a jelölésekkel a (8.1), (8.2) optimális irányítási feladat ekvivalens az

$$\begin{cases} F(x, u) \to \min \\ G(x, u) = 0 \end{cases} \tag{8.3}$$

feltételes szélsőérték-feladattal. A feltételeink szerint itt az 5.14 Példa alapján F és G egyaránt folytonosan differenciálhatóak.

8.2. Az irányíthatósági feltétel

A (8.3) feladat megoldásához szükségünk van a ráképezési feltételre. Nevezetesen, ha (x_0, u_0) a feladat megoldása, akkor

$$\operatorname{im} G'(x_0, u_0) = X \times C[0,T]\,. \tag{8.4}$$

Ennek vizsgálatát végezzük el ebben a szakaszban.

Az 5.14 Példa szerint G folytonosan differenciálható, és könnyen ellenőrizhető, hogy minden $v \in C[0,T]$, valamint $w \in L^2[0,T]$ mellett

$$G'(x_0, u_0)(v, w) = \left(v(T), v(t) - \int_0^t A(s)v(s)\, ds - \int_0^t B(s)w(s)\, ds \right)$$

ahol

$$A(s) = \partial_1 g(x_0(s), u_0(s)) \quad \text{és} \quad B(s) = \partial_2 g(x_0(s), u_0(s))$$

$n \times n$, illetve $n \times m$ méretű mátrixok. A továbbiakban ebben a szakaszban feltesszük, hogy ezekre a mátrixokra fennáll a következő feltétel.

Bármely $a \in X$ vektorhoz találhatók olyan $v \in C[0,T]$ és $w \in L^2[0,T]$ függvények, amelyekre

$$\begin{aligned} v(t) &= \int_0^t A(s)v(s)\,ds + \int_0^t B(s)w(s)\,ds \\ v(T) &= a \end{aligned}$$

minden $t \in [0,T]$ pontban. Vegyük észre, hogy ez a reláció azzal ekvivalens, hogy az

$$x'(t) = A(t)x(t) + B(t)u(t)$$

lineáris rendszer irányítható a $[0,T]$ intervallumon. Felhasználva a 2. fejezetben bevezetett

$$\Lambda u = \int_0^T \Phi(0,s)B(s)u(s)\,ds$$

folytonos lineáris leképezést az $L^2[0,T]$ téren, a 2.4 Tétel szerint a fenti lineáris rendszer pontosan akkor irányítható, ha

$$\operatorname{rang} \Lambda\Lambda^* = n \tag{8.5}$$

Ezt a feltételt a (8.1), (8.2) optimális irányítási feladatra vonatkozó *irányíthatósági feltételnek* nevezzük.

8.2 Állítás. *Ha fennáll a (8.5) irányíthatósági feltétel, akkor teljesül a (8.4) ráképezési tulajdonság.*

Bizonyítás. A (8.4) tulajdonság azt jelenti, hogy bármely $b \in X$ vektorhoz és $y \in C[0,T]$ függvényhez találhatók olyan $v \in C[0,T]$ és $w \in L^2[0,T]$ függények, amelyekre $G'(x_0,u_0)(v,w) = (b,y)$, azaz

$$\begin{aligned} y(t) &= v(t) - \int_0^t A(s)v(s)\,ds + \int_0^t B(s)w(s)\,ds \\ v(T) &= b \end{aligned}$$

minden $t \in [0,T]$ mellett. Ennek igazolásához tekintsük azt a φ függvényt, amely megoldása a

$$\begin{aligned} \varphi'(t) &= A(t)\varphi(t) + A(t)y(t) \\ \varphi(0) &= 0 \end{aligned}$$

kezdetiérték-feladatnak a $[0,T]$ intervallumon. Vezessük be a $\psi = \varphi + y$ jelölést, akkor $A(t)\psi(t) = A(t)\varphi(t) + A(t)y(t) = \varphi'(t)$, és ezért

$$\varphi(t) = \int_0^t A(s)\psi(s)\,ds = \psi(t) - y(t)\,.$$

Az egyenlőséget átrendezve azt kapjuk, hogy

$$\psi(t) = \int_0^t A(s)\psi(s)\,ds + y(t)$$

a $[0,T]$ intervallumon.

A (8.5) irányíthatósági feltétel alapján az $a = b - \psi(T)$ vektorhoz találhatók olyan $\bar{v} \in C[0,T]$ és $w \in L^2[0,T]$ függvények, amelyekre

$$\begin{aligned} \bar{v}(t) &= \int_0^t A(s)\bar{v}(s)\,ds + \int_0^t B(s)w(s)\,ds \\ \bar{v}(T) &= b - \psi(T)\,. \end{aligned}$$

minden $t \in [0,T]$ pontban. Ekkor a $v = \bar{v} + \psi$ és w függényekre egyszerű számolással ellenőrizhető, hogy

$$G'(x_0,u_0)(v,w) = (b,y)\,,$$

amit igazolnunk kellett. □

8.3. A Pontrjagin-féle maximumelv

Tekintsük újra a (8.1), (8.2) optimális irányítási feladatot, továbbá a vele ekvivalens (8.3) feltételes szélsőérték feladatot. A feladat Lagrange-függvénye

$$L(x,u,p) = F(x,u) + \langle p, G(x,u)\rangle\,,$$

ahol p folytonos lineáris funkcionál az $X \times C[0,T]$ téren. Minden ilyen funkcionál a

$$p(x,y) = q(x) + r(y)$$

alakban állítható elő, ahol $q \in X^*$, $r \in C[0,T]^*$, illetve $x \in X$, $y \in C[0,T]$.

Az alábbi tételben a Lagrange-multiplikátor elvét (7.7 Tétel) megfogalmazzuk az optimális irányítási feladatra.

8.3 Tétel. (Pontrjagin-féle maximumelv) *Tegyük fel, hogy teljesül a (8.5) irányíthatósági feltétel, és legyen (x_0,u_0) a (8.1), (8.2) feladat megoldása. Akkor található az*

$$y'(t) = -A(t)^* y(t) - \partial_1 f(x_0(t),u_0(t))$$

adjungált rendszernek olyan ψ megoldása, amelyre

$$\partial_2 f(x_0(t),u_0(t)) + B(t)^*\psi(t) = 0 \tag{8.6}$$

majdnem mindenütt a $[0,T]$ intervallumon.

Bizonyítás. Tegyük fel, hogy (x_0, u_0) a feladat megoldása. Az irányíthatósági feltétel alapján a (8.3) feladatra alkalmazható a 7.7 Tétel, azaz létezik olyan p folytonos lineáris funkcionál az $X \times C[0,T]$ téren, amelyre

$$F'(x_0, u_0) + pG'(x_0, u_0) = 0\,. \tag{8.7}$$

Az első változó szerinti parciális deriváltra térve ez azt jelenti, hogy

$$\partial_1 F(x_0, u_0) + p\partial_1 G(x_0, u_0) = 0\,.$$

Ez a Riesz-féle reprezentációs tétel[1] szerint úgy írható, hogy található olyan $q \in X$ vektor és olyan μ vektormérték a $[0,T]$ intervallum Borel-halmazain, amelyekre bármely $v \in C[0,T]$ mellett

$$\int_0^T \langle \partial_1 f(x_0(t), u_0(t)), v(t)\rangle\,dt + \langle q, v(T)\rangle + \int_0^T \langle v(t) - \int_0^t A(s)v(s)\,ds, d\mu(t)\rangle = 0.$$

Alkalmazzuk ezt az egyenlőséget speciálisan $v \in W_0^2[0,T]$ függvényekre, akkor egyrészt $v(0) = v(T) = 0$, másrészt a

$$\psi(t) = \int_t^T d\mu(s)\,, \quad t \in [0,T]$$

függvényt bevezetve parciális integrálás után azt nyerjük, hogy

$$\begin{aligned} 0 &= \int_0^T \langle \partial_1 f(x_0(t), u_0(t)), v(t)\rangle\,dt - \int_0^T \langle v'(t) - A(t)v(t), \psi(t)\,dt \\ &= \int_0^T \langle \partial_1 f(x_0(t), u_0(t)) + A(t)^*\psi(t), v(t)\rangle\,dt - \int_0^T \langle \psi(t), v'(t)\rangle\,dt\,, \end{aligned}$$

hiszen a kiintegrált tag nulla lesz. A második sor első integráljában újra parciálisan integrálva

$$\begin{aligned} 0 &= -\int_0^T \langle \int_0^t \partial_1 f(x_0(s), u_0(s)) + A(s)^*\psi(s)\,ds, v'(t)\rangle\,dt - \int_0^T \langle \psi(t), v'(t)\rangle\,dt \\ &= -\int_0^T \langle \psi(t) + \int_0^t \partial_1 f(x_0(s), u_0(s)) + A(s)^*\psi(s)\,ds, v'(t)\rangle\,dt\,, \end{aligned}$$

ugyanis a v peremfeltételei miatt a kiintegrált rész újra nullával egyenlő.

Innen a Du Bois-Reymond-lemma (lásd a 6.1 Lemmát) folytán a

$$\psi(t) + \int_0^t \partial_1 f(x_0(s), u_0(s)) + A(s)^*\psi(s)\,ds$$

konstans a $[0,T]$ intervallumon, azaz itt a ψ függvény kielégíti a

$$\psi'(t) = -A(t)^*\psi(t) - \partial_1 f(x_0(t), u_0(t))$$

[1] Kánnai Zoltán: *Analitikus módszerek a pénzügyben és a közgazdaságtanban, 2013.*

adjungált rendszert.

Tekintsük ezután újra a (8.7) egyenlőséget, és írjuk fel a második változó szerinti parciális deriváltakat. Ekkor ismét parciális integrálással bármely $w \in L^2[0,T]$ mellett

$$\begin{aligned} 0 &= \int_0^T \langle \partial_2 f(x_0(t), u_0(t)), w(t)\rangle\, dt - \int_0^T \langle \int_0^t B(s)w(s)\, ds, d\mu(t)\rangle \\ &= \int_0^T \langle \partial_2 f(x_0(t), u_0(t)), w(t)\rangle\, dt + \int_0^T \langle B(t)w(t), \psi(t)\rangle\, dt \\ &= \int_0^T \langle \partial_2 f(x_0(t), u_0(t)) + B(t)^*\psi(t), w(t)\rangle\, dt\,, \end{aligned}$$

ugyanis a kiintegrált rész nulla. Mivel ez az egyenlőség minden $w \in L^2[0,T]$ mellett teljesül, innen következik, hogy

$$\partial_2 f(x_0(t), u_0(t)) + B(t)^*\psi(t) = 0$$

majdnem mindenütt a $[0,T]$ intervallumon, és éppen ezt kellett igazolnunk. □

8.4 Definíció. A (8.1), (8.2) alatti irányítási feladat *Hamilton-függvényén* a $H : X \times Y \times X \to \mathbb{R}$,

$$H(x, u, p) = f(x, u) + \langle p, g(x, u)\rangle$$

függvényt értjük.

A Hamilton-függvény használatával a Pontrjagin-féle maximumelv az alábbi formában fogalmazható meg.

8.5 Tétel. (Pontrjagin-féle maximumelv, Hamilton-formalizmus) *Tegyük fel, hogy fennáll a (8.5) irányíthatósági feltétel. Ha (x_0, u_0) optimális folyamat, akkor az*

$$y'(t) = -\partial_1 H(x_0(t), u_0(t), y(t))$$

adjungált rendszernek található olyan ψ megoldása, amelyre

$$\partial_2 H(x_0(t), u_0(t), \psi(t)) = 0$$

majdnem mindenütt a $[0,T]$ intervallumon.

8.4. A transzverzalitási feltétel

Tekintsük újra a (8.1), (8.2) irányítási feladatot azzal a kiegészítéssel, hogy egy korlátozás érvényes a megengedett irányításokra. Nevezetesen, tegyük fel, hogy adott egy nem üres $U \subset Y$ zárt halmaz, és értelmezzük a megengedett irányítások halmazát az alábbi módon:

$$\hat{U} = \{u \in L^2[0,T] : u(t) \in U \text{ m.m. } \} .$$

Ebben az esetben nem teszünk megkötést a trajektóriák x_T végpontjára, $x(T)$ az X állapottér tetszőleges eleme

A megengedett folyamat, illetve optimális folyamat fogalmát analóg módon értelmezzük. A szabad végpontú feladatra érvényes Pontrjagin-féle maximumelv megfogalmazása előtt egy lemmát bocsátunk előre.

8.6 Lemma. *Legyenek E_1 és E_2 valós normált terek, $h : E_1 \times E_2 \to \mathbb{R}$ valamely folytonosan differenciálható függvény, valamint $x : E_2 \to E_1$ olyan Lipschitz-folytonos függvény, amelyre valamely $u_0 \in E_2$ vektorra*

$$\partial_1 h\left(x\left(u_0\right), u_0\right) = 0 .$$

Akkor van olyan u_0-ban kisrendű $r : E_2 \to \mathbb{R}$ függvény, hogy minden $u \in E_2$ esetén

$$h\left(x\left(u\right), u\right) - h\left(x\left(u_0\right), u_0\right) = h\left(x\left(u_0\right), u\right) - h\left(, x\left(u_0\right), u_0\right) + r\left(u\right) \ .$$

Bizonyítás. Megmutatjuk, hogy az

$$r\left(u\right) := h\left(x\left(u\right), u\right) - h\left(x\left(u_0\right), u\right)$$

választással teljesül a kívánt egyenlőség. Jelölje $\lambda > 0$ az x függvény Lipschitz-együtthatóját. A feltettek miatt minden $\varepsilon > 0$ számhoz létezik $\delta > 0$ szám, hogy $\|u - u_0\| + \|z - x\left(u_0\right)\| \leq \delta$ esetén

$$\|\partial_1 h\left(z, u\right)\| \leq \frac{\varepsilon}{\lambda} .$$

Ekkor minden $u \in E_2$,

$$\|u - u_0\| \leq \frac{\delta}{2\lambda + 2}$$

vektorra egyrészt $\|x\left(u\right) - x\left(u_0\right)\| \leq \delta/2$, másrészt a Lagrange-féle középérték-tétel (lásd az 5.11 Tételt) alapján

$$\begin{aligned}
\|r\left(u\right)\| &= \|h\left(x\left(u\right), u\right) - h\left(x\left(u_0\right), u\right)\| \leq \\
&\leq \sup_{z \in [x(u_0), x(u)]} \|\partial_1 h\left(z, u\right)\| \cdot \|x\left(u\right) - x\left(u_0\right)\| \leq \\
&\leq \sup_{\|z - x(u_0)\| \leq \delta/2} \|\partial_1 h\left(z, u\right)\| \cdot \|x\left(u\right) - x\left(u_0\right)\| \leq \\
&\leq \frac{\varepsilon}{\lambda} \cdot \|x\left(u\right) - x\left(u_0\right)\| \leq \varepsilon \cdot \|u - u_0\| \ ,
\end{aligned}$$

ahonnan azonnal adódik, hogy r kisrendű az u_0 helyen. □

A továbbiakban tegyük fel, hogy $\partial_1 F$ egyenletesen korlátos, továbbá a g függvényre fennáll a

$$\|g(x,u) - g(z,v)\| \leq \lambda(\|x-z\| + \|u-v\|)$$

Lipschitz-tulajdonság valamely $\lambda > 0$ mellett.

8.7 Tétel. **(Maximumelv, szabadvégpontú feladat)** *Ha* (x_0, u_0) *optimális folyamat, akkor az*

$$\begin{aligned} y'(t) &= -\partial_1 H(x_0(t), u_0(t), y(t)) \\ y(T) &= 0 \end{aligned}$$

adjungált rendszernek található olyan ψ *megoldása, amelyre*

$$H\left(x_0\left(t\right), u_0\left(t\right), \psi(t)\right) = \min_{u \in U} H\left(x_0\left(t\right), u, \psi(t)\right)$$

majdnem mindenütt a $[0,T]$ *intervallumon.*

Bizonyítás. Adott $u \in \hat{U}$ megengedett irányítás mellett jelentse $x(u)$ az

$$\begin{aligned} x'(t) &= g(x(t), u(t)) \\ x(0) &= x_0 \end{aligned}$$

kezdeti érték feladat egyetlen megoldását. A Peano-egyenlőtlenség alapján tetszőleges $u, v \in \hat{U}$ irányítások mellett

$$\|x(u) - x(v)\| \leq \lambda e^{\lambda T} \|u - v\| .$$

Az $E_1 = C[0,T]$ és $E_2 = L^2[0,T]$ választás mellett alkalmazzuk a fenti lemmát a

$$h(x,u) = \int_0^T \left(H(x(t), u(t), \psi(t) + \langle \psi'(t), x(t)\rangle\right) dt$$

függényre. Az 5.13 Példa értelmében h differenciálható, nevezetesen az adjungált egyenletet figyelembe véve

$$\partial_1 h(x(u_0), u_0)v = \int_0^T \langle A(t)^* \psi(t) + \partial_1 f(x_0(t), u_0(t)) + \psi'(t), v(t)\rangle \, dt = 0$$

minden $v \in C[0,T]$ mellett. Ez azt jelenti, hogy

$$\partial_1 h(x(u_0), u_0) = 0 \, .$$

Tehát a fenti lemma alapján van olyan $r : \hat{U} \to \mathbb{R}$ az u_0-ban kisrendű függvény, hogy tetszőleges $u \in \hat{U}$ esetén

$$h\left(x\left(u\right), u\right) - h\left(x\left(u_0\right), u_0\right) = h\left(x\left(u_0\right), u\right) - h\left(x\left(u_0\right), u_0\right) + r\left(u\right) =$$

$$= \int_0^T H(x(u_0)(t), u(t), \psi(t))\, dt - \int_0^T H(x(u_0)(t), u_0(t), \psi(t))\, dt + r(u) \ .$$

Ugyanakkor minden $u \in \hat{U}$-ra $x(u)$ választása miatt

$$\begin{aligned}
& h(x(u), u) = \\
= \ & \int_0^T (f(x(u)(t), u(t)) + \langle \psi(t), g(x(u)(t), u(t)) \rangle + \langle \psi'(t), x(u)(t) \rangle)\, dt \\
= \ & \int_0^T f(x(u)(t), u(t)) + \left\langle \psi(t), [x(u)]'(t) \right\rangle + \langle \psi'(t), x(u)(t) \rangle\, dt = \\
= \ & \int_0^T f(x(u)(t), u(t))\, dt + [\langle \psi(t), x(u)(t) \rangle]_0^T = \\
= \ & \int_0^T f(x(u)(t), u(t))\, dt + 0 - \langle \psi(0), x_0 \rangle \ ,
\end{aligned}$$

ezért

$$\int_0^T (f(x(u)(t), u(t)) - f(x(u_0)(t), u_0(t)))\, dt = h(x(u), u) - h(x(u_0), u_0) =$$

$$= \int_0^T H(x(u_0)(t), u(t), \psi(t))\, dt - \int_0^T H(x(u_0)(t), u_0(t), \psi(t))\, dt + r(u) \ .$$

Emiatt tetszőleges megengedett u irányításra u_0 optimalitása alapján

$$\begin{aligned}
0 \ \leq \ & \int_0^T f(x(u)(t), u(t))\, dt - \int_0^T f(x(u_0)(t), u_0(t))\, dt = \\
= \ & \int_0^T (H(x(u_0)(t), u(t), \psi(t)) - H(x(u_0)(t), u_0(t), \psi(t)))\, dt + r(u) \ .
\end{aligned}$$

Indirekt módon tegyük föl, hogy van olyan $v \in U$ és olyan pozitív mértékű $A_0 \subset [0, T]$ halmaz, hogy minden $t \in A_0$ esetén

$$H(x_0(t), u_0(t), \psi(t)) > H(x_0(t), v, \psi(t)) \ .$$

Ekkor olyan pozitív mértékű mérhető $A_1 \subset A_0$ halmaz is van, hogy alkalmas $\varepsilon > 0$ számmal minden $t \in A_1$-re

$$H(x_0(t), u_0(t), \psi(t)) > H(x_0(t), v, \psi(t)) + \varepsilon \ ,$$

továbbá $\|u_0(\cdot) - \overline{u}\|$ is korlátos az A_1 halmazon valamely $\alpha > 0$ korláttal. Tetszőleges pozitív mértékű Lebesgue-mérhető $A \subset A_1$ halmazra jelölje

$$u_A(t) = \begin{cases} u_0(t) & \text{ha } t \notin A \\ v & \text{ha } t \in A \end{cases} .$$

Ezzel (ahol μ a Lebesgue-mértéket jelenti)

$$\|u_A - u_0\|_{L^2} = \int_A \|u_0(t) - v\| \, dt \le \alpha \cdot \mu(A) \ .$$

Így az A_1-nek valamely elegendően kicsi pozitív mértékű A részhalmazára r kisrendűsége miatt

$$|r(u_A)| \le \frac{\varepsilon}{2\alpha} \cdot \|u_A - u_0\|_{L^1} \le \frac{\varepsilon}{2} \cdot \mu(A) \ .$$

Ugyanakkor u_A megengedett irányítás, így a fenti A halmazra

$$\begin{aligned} 0 &\le \int_0^T f(x(u_A)(t), u_A(t)) \, dt - \int_0^T f(x(u_0)(t), u_0(t)) \, dt \\ &= \int_0^T (H(x(u_0)(t), u_A(t), \psi(t)) - H(x(u_0)(t), u_0(t), \psi(t))) \, dt + r(u_A) \\ &= \int_A (H(x(u_0)(t), v, \psi(t)) - H(x(u_0)(t), u_0(t), \psi(t))) \, dt + r(u_A) \le \\ &\le -\varepsilon \cdot \mu(A) + r(u_A) \le -\varepsilon \cdot \mu(A) + \frac{\varepsilon}{2} \cdot \mu(A) < 0 \ , \end{aligned}$$

ami ellentmondás (hiszen A_1-nek létezik tetszőlegesen kicsi mértékű mérhető részhalmaza). Ezért minden $v \in U$ esetén m.m. $t \in [0, T]$-re

$$H(x_0(t), u_0(t), \psi(t)) \le H(x_0(t), v, \psi(t)) \ .$$

Legyen most $U_0 \subseteq U$ megszámlálható sűrű részhalmaz. A most kapottak alapján még mindig m.m. $t \in [0, T]$-re

$$H(x_0(t), u_0(t), \psi(t)) \le \inf_{v \in U_0} H(x_0(t), v, \psi(t)) \ ,$$

ami a $H(x_0(t), \cdot, \psi(t))$ függvények folytonossága miatt azt jelenti, hogy m.m. $t \in [0, T]$-re

$$H(x_0(t), u_0(t), \psi(t)) = \min_{u \in U} H(t, x_0(t), u, \psi(t)) \ .$$

amit igazolnunk kellett. $\square$

8.8 Definíció. Szabad végpontú feladatokban az adjungált rendszerre vonatkozó

$$\psi(T) = 0$$

feltételt *transzverzalitási feltételnek* nevezzük.

8.9 Példa. Oldjuk meg a $[0, T]$ intervallumon az

$$x'(t) = u(t)\,, \quad x(0) = x_0 \quad x(T) \text{ szabad}$$

irányítási rendszerre az

$$F(x, u) = \int_0^T (1 - tx(t) - u(t)^2)\, dt \to \max$$

optimális irányítási feladatot. A probléma Hamilton-függvénye:

$$H(x, u, p) = 1 - tx - u^2 + pu\,.$$

Ha (x_0, u_0) optimális folyamat, akkor a Pontrjagin-féle maximumelv szerint u_0 maximalizálja a Hamilton-függvényt, azaz

$$0 = \partial_2 H(x_0(t), u_0(t), \psi(t)) = -2u_0(t) + \psi(t)\,,$$

ezért $2u_0(t) = \psi(t)$ a $[0, T]$ intervallumon, ahol ψ az adjungált egyenlet megoldása. Az adjungált egyenlet a következő alakú:

$$\psi'(t) = -\partial_1 H(x_0(t), u_0(t), \psi(t)) = t\,,$$

és a transzverzalitási feltétel alapján $\psi(T) = 0$, ezért

$$\psi(t) = \frac{1}{2}(t^2 - T^2)\,,$$

és így az optimális irányításra

$$u_0(t) = \frac{1}{4}(t^2 - T^2)$$

adódik. Innen a differenciálegyenletet integrálva azt kapjuk, hogy

$$x_0(t) = x_0 + \frac{1}{12}t^3 - \frac{T^2}{4}t\,,$$

és a maximumelvet ez az egyetlen függvény elégíti ki.

8.5. A jövedelemallokációs probléma

Tekintsünk egy termelőt, aki egyetlen homogén terméket állít elő a $[0, T]$ időintervallumon. Jelentse $x(t)$ a t időpillanatban előállított termékmennyiséget.

Ennek egy részét a termelő értékesíti, amelyből származó jövedelmet beruházásra fordítja. Ezzel a beruházással növelni képes a termelési kapacitását. Mekkora részt fordítson beruházásra, hogy a legnagyobb készletet érje el?

Jelentse $u(t)$ az előállított termék azon hányadát, amelyet a termelő beruházásra fordít a t időpontban. Ekkor a termelés növekedését leíró egyenlet

$$\begin{aligned} x'(t) &= u(t)x(t) \\ x(0) &= x_0 \,. \end{aligned} \tag{8.8}$$

Világos, hogy ekkor $0 \leq u(t) \leq 1$. A termelő úgy kívánja megválasztani az u allokációs irányítást, hogy a $[0, T]$ időintervallumon a felhalmozott készlete maximális legyen, azaz

$$F(x, u) = \int_0^T (1 - u(t))x(t)\, dt \to \max \,. \tag{8.9}$$

Ebben a (8.8), (8.9) szabad végpontú optimális irányítási feladatban $U = [0, 1]$.

A fenti feladat Hamilton-függvénye a következő alakban írható fel

$$H(x, u, p) = (1 - u)x + pux \,,$$

és így az adjungált rendszer az alábbi alakot ölti

$$\psi'(t) = -\partial_1 H(x_0(t), u_0(t), \psi(t)) = -u_0(t)\psi(t) - (1 - u_0(t)) \,,$$

ahol (x_0, u_0) optimális folyamat, továbbá a transzverzalitási feltételből

$$\psi(T) = 0 \,.$$

A Hamilton-függvény maximális az u_0 optimális irányítás mentén, azaz

$$H(x_0(t), u_0(t), \psi(t)) = \max_{u \in [0,1]} H(x_0(t), u, \psi(t))$$

a $[0, T]$ intervallumon majdnem mindenütt. Mivel $x(t) \geq 0$ ezen az intervallumon, innen azonnal adódik, hogy

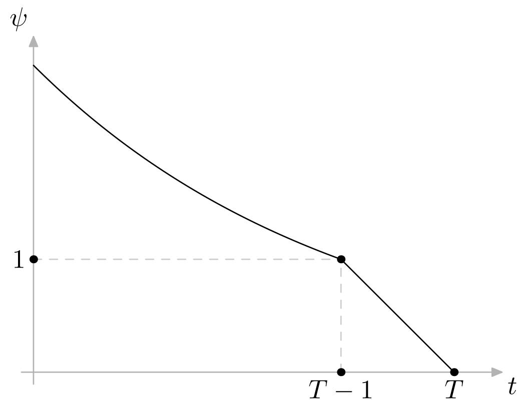
$$u_0(t) = \begin{cases} 1 & \text{ha } \psi(t) > 1 \\ 0 & \text{ha } \psi(t) < 1 \,. \end{cases}$$

Mivel $\psi(T) = 0$, innen azt kapjuk, hogy $u_0(T) = 0$. Ezután az adjungált egyenletet a $t = T$ végponttól a $t = 0$ pontig fordított irányban integrálva az optimális irányításra adódik, hogy

$$u_0(t) = \begin{cases} 1 & \text{ha } 0 \leq t \leq T - 1 \\ 0 & \text{ha } T - 1 < t \leq T \,, \end{cases}$$

míg az adjungált rendszer ψ megoldása exponenciálisan fogyó a $[0, T - 1]$ intervallumon, majd pedig $\psi(t) = T - t$ a $[T - 1, T]$ intervallumon. Lásd a 8.1 ábrát!

Ez azt jelenti, hogy a maximális készlet elérése érdekében a termelő a $[0, T-1]$ intervallumon mindent beruházásra fordít, majd a $[T - 1, T]$ időintervallumon teljes kapacitását a készletek növelésére fordítja. Felhívjuk a figyelmet arra, hogy itt az optimális irányítás bang-bang, azaz extremális irányítás.



8.1. ábra. Az adjungált egyenlet megoldása

8.6. Gyakorlatok

1. Oldjuk meg az alábbi szabad végpontú feladatot:

$$x'(t) = x(t) + u(t) \quad x(0) = 1\,, \quad x(1) \text{ szabad}$$
$$\int_0^1 (1 - u(t)^2)\, dt \to \max$$

2. Keressük meg az alábbi feladat megoldását:

$$x'(t) = u(t)\,, \quad x(0) = 0\,, \quad x(1) = 0\,, \quad u(t) \in [-1, 1]$$
$$\int_0^1 (x(t)^2 - 2x(t))\, dt \to \min$$

3. Állítsuk elő a maximumelv segítségével az

$$x'(t) = u(t)\,, \quad x(0) = 1\,, \quad x(2) \text{ szabad} \quad u(t) \in [0, 1]$$
$$\int_0^2 (x(t)^2 - 2u(t))\, dt \to \max$$

feladat egyetlen lehetséges megoldását. (Útmutatás: igazoljuk, hogy $\psi(t)$ szigorúan monoton fogyó a $[0, 2]$ intervallumon.)

9. fejezet

A maximumelv elégségessége

Ebben a fejezetben megmutatjuk, hogy bizonyos konvexitási feltételek mellett a Pontrjagin-féle maximumelv elégséges is az optimalitásra.

9.1. A maximumelv egységes végpontfeltétellel

9.1 Definíció. Legyen $[0,T]$ adott intervallum, $U \subset Y$ adott halmaz, legyenek továbbá $f : X \times Y \to \mathbb{R}$ és $g : X \times Y \to X$ folytonosan differenciálható függvények. Legyenek $x_0,\ x_T \in X$ adott pontok.

Tekintsünk a következő *optimális irányítási feladatot*:

$$\begin{cases} \int_0^T f(x(t), u(t))\, dt \to \min \\ x'(t) = g(x(t), u(t)) \ \ \text{m.m. } t \in [0,T] \\ u(t) \in U \ \ \text{m.m. } t \in [0,T] \\ x(0) = x_0, \\ (a)\ x(T) = x_T, \quad (b)\ x(T) \text{ szabad} \end{cases} . \tag{9.1}$$

Jelölje $\hat{U} = \{u \in L^2[0,T] :\ u(t) \in U \text{ m.m.}\}$, és tegyük fel, hogy, $\forall u \in \hat{U}$ esetén létezik a fenti differenciaegyenletnek a $[0,T]$-n megoldása.

Adott (x,u) megengedett folyamat mellett jelölje

$$F(x,u) = \int_0^T f(x(t), u(t))\, dt.$$

A feladat *Hamilton-függvénye*:

$$H : [0,T] \times X \times Y \times X \to \mathbb{R}, \quad H(x,u,p) = f(x,u) + \langle p, g(x,u) \rangle,$$

illetve a Hamilton-függvény *értékfüggvénye*:

$$H^\wedge : [0,T] \times X \times X \to \mathbb{R}, \quad H^\wedge(x,p) = \min_{u \in U} H(x,u,p).$$

Ha az optimális irányítási feladatot maximum feladatként fogalmazzuk meg, akkor a Hamilton-függvény értékfüggvénye:

$$H^\vee : [0,T] \times X \times X \to \mathbb{R}, \quad H^\vee(x,p) = \max_{u \in U} H(x,u,p).$$

9.2 Tétel. (Pontrjagin-féle maximumelv) *Ha* (x_0, u_0) *optimális folyamat a (9.1) feladatban, akkor az*

$$y'(t) = -\partial_2 H(x_0(t), u_0(t), y(t)),$$

adjungált rendszernek létezik olyan ψ *megoldása, amelyre*

$$H(x_0(t), u_0(t), \psi(t)) = \min_{u \in U} H(x_0(t), u, \psi(t)),$$

továbbá teljesül az alábbi transzverzalitási feltétel:

$$\langle \psi(T), x(T) - x_0(T) \rangle = 0.$$

9.3 Megjegyzés. A Hamilton-függvényhez hasonló módon az (9.1) feladatnak értelmezhető a *Lagrange-függvénye*:

$$L : [0,T] \times X \times Y \times X \times X \to \mathbb{R}, \quad L(x,u,p,x') = f(x,u) + \langle p, g(x,u) - x' \rangle.$$

Ennek a segítségével is megfogalmazható a maximumelv:

(1) A minimumfeltétel gyakorlatilag azonos:

$$L(x_0(t), u_0(t), p(t), x_0'(t)) = \min_{u \in U} L(x_0(t), u, p(t), x_0'(t)).$$

(2) Az adjungált feltétel viszont a

$$\frac{d}{dt} \partial_3 L(x_0(t), x_0'(t), u_0(t), p(t)) = \partial_2 L(x_0(t), x_0'(t), u_0(t), p(t))$$

alakot ölti, ami analóg a variációs feladatra vonatkozó Euler-Lagrange-differenciálegyenlettel, csak az f függvény helyett a feladat Lagrange-függvénye szerepel benne.

(3) A transzverzalitási feltétel ugyanaz.

9.2. A Mangasarian-féle elégséges feltétel

9.4 Tétel. **(Mangasarian-tétel)** *Legyen (x_0, u_0) megengedett folyamata a (9.1) feladatnak. Tegyük fel, hogy létezik olyan p, amelyre teljesül, hogy*

(1) $H(x_0(t), u_0(t), p(t)) = \min_{u \in U} H(x_0(t), u, p(t))$,

(2) $p'(t) = -\partial_2 H(x_0(t), u_0(t), p(t))$,

(3) a peremértékekre vonatkozó transzverzalitási feltétel:

$$\langle p(T), x(T) - x_0(T) \rangle = 0.$$

Ha az U halmaz konvex, és a Hamilton-függvény $(x, u) \mapsto H(x, u, p)$ szeletei konvexek, akkor az (x_0, u_0) folyamat megoldása a (9.1) feladatnak.

Bizonyítás: Legyen (x, u) egy tetszőleges másik megengedett folyamat. Be kell látni, hogy

$$F(x, u) - F(x_0, u_0) \quad = \quad \int_0^T f(x(t), u(t))\ dt - \int_0^T f(x_0(t), u_0(t))\ dt \geq 0.$$

Az (x, u) megengedett folyamatra az (9.1) feladat $x'(t) = g(x(t), u(t))$ feltétele alapján a Hamilton-függvény definíciója szerint

$$\begin{aligned} f(x(t), u(t)) &= H(x(t), u(t), p(t)) - \langle p(t), g(x(t), u(t)) \rangle \\ &= H(x(t), u(t), p(t)) - \langle p(t), x'(t) \rangle , \end{aligned}$$

hasonlóan az (x_0, u_0) megengedett folyamatra is, amiből

$$\begin{aligned} F(x, u) - F(x_0, u_0) &= \int_0^T H(x(t), u(t), p(t)) - H(x_0(t), u_0(t), p(t))\ dt \\ &+ \int_0^T \langle -p(t), x'(t) - x_0'(t) \rangle \ dt. \end{aligned} \tag{9.2}$$

Tekintsük először a (9.2) második tagját. Parciális integrálással kapjuk a következőt:

$$\begin{aligned} \int_0^T \langle -p(t), (x - x_0)'(t) \rangle \ dt &= -\left[\langle p(t), x(t) - x_0(t) \rangle \right]_0^T \\ &+ \int_0^T \langle p'(t), x(t) - x_0(t) \rangle \ dt. \end{aligned} \tag{9.3}$$

Ennek első tagja

$$\left[\langle p(t), x(t) - x_0(t) \rangle \right]_0^T = \langle p(T), x(T) - x_0(T) \rangle - \langle p(0), x(0) - x_0(0) \rangle ,$$

a kezdeti feltétel szerint

$$\langle p(0), x(0) - x_0(0)\rangle = \langle p(0), x_0 - x_0\rangle = 0,$$

a transzverzalitási feltétel szerint pedig

$$\langle p(T), x(T) - x_0(T)\rangle = 0.$$

Ezek alapján a (9.3) egyenlőség a következő alakú

$$\int_0^T \langle -p(t), (x'(t) - x_0'(t))\rangle \ dt = \int_0^T \langle p'(t), (x(t) - x_0(t))\rangle \ dt. \tag{9.4}$$

Tekintsük most a (9.2) első tagját. Mivel az $(x, u) \mapsto H(x, u, p)$ függvény konvex, és a deriválható konvex függvények jellemzése szerint a függvény gráfja az érintője felett van, ezért

$$\begin{aligned}
& H(x(t), u(t), p(t)) - H(x_0(t), u_0(t), p(t)) \\
\geq\ & [\partial_{2,3} H(x_0(t), u_0(t), p(t))] \cdot \begin{bmatrix} x(t) - x_0(t) \\ u(t) - u_0(t) \end{bmatrix} \\
=\ & \langle \partial_2 H(x_0(t), u_0(t), p(t)), x(t) - x_0(t)\rangle \\
+\ & \langle \partial_3 H(x_0(t), u_0(t), p(t)), u(t) - u_0(t)\rangle \\
=\ & \langle -p'(t), x(t) - x_0(t)\rangle \\
+\ & \langle \partial_3 H(x_0(t), u_0(t), p(t)), u(t) - u_0(t)\rangle,
\end{aligned}$$

ahol az utolsó egyenlőséghez felhasználtuk a tétel (2) feltételét, az adjungált feltételt. A tétel (1) feltétele, a minimumfeltétel alapján az 5.16. Tétel szerint

$$\langle \partial_3 H(x_0(t), u_0(t), p(t)), u(t) - u_0(t)\rangle \geq 0.$$

A fentieket összefoglalva kapjuk, hogy

$$\begin{aligned}
& \int_0^T H(x(t), u(t), p(t)) - H(x_0(t), u_0(t), p(t)) \ dt \qquad (9.5) \\
\geq\ & \int_0^T \langle -p'(t), x(t) - x_0(t)\rangle \ dt.
\end{aligned}$$

A (9.4) és (9.5) egyenlőtlenségeket a (9.2) kifejezésbe helyettesítve kapjuk, a kívánt $F(x, u) - F(x_0, u_0) \geq 0$ egyenlőtlenséget.

Ha az $(x, u) \mapsto H(x, u, p)$ függvény szigorúan konvex, akkor a (9.5) egyenlőtlenség szigorúan teljesül, ezért ekkor a megoldás egyértelmű. □

9.3. Az Arrow-féle elégséges feltétel

9.5 Tétel. (Arrow-tétel) *Legyen* (x_0, u_0) *megengedett folyamata a (9.1) feladatnak. Tegyük fel, hogy létezik olyan* p*, amelyre teljesül a 9.4.tétel (1), (2), (3) feltétele. Ha az* U *halmaz konvex, és a Hamilton-függvény értékfüggvényének* $x \mapsto H^\wedge(x,p)$ *szeletei konvexek, akkor az* (x_0, u_0) *folyamat megoldása a (9.1) feladatnak.*

Bizonyítás: Legyen (x,u) egy tetszőleges másik megengedett folyamat. Be kell látni, hogy

$$F(x,u) - F(x_0,u_0) \quad = \quad \int_0^T f(x(t),u(t))\,dt - \int_0^T f(x_0(t),u_0(t))\,dt \geq 0.$$

A bizonyítás megegyezik a Mangasarian-tétel bizonyításával a (9.5) egyenlőtlenség bizonyításának a kivételével. Nem tudjuk, ugyanis, hogy a Hamilton-függvény értékfüggvényének szeletei differenciálhatók-e. Azt fogjuk belátni, hogy Gâteaux-differenciálható. Legyen $v \in \mathbb{R}^n$ tetszőleges.

Mivel a Hamilton-függvény a második változójában differenciálható, ezért ebben a változójában minden irány mentén differenciálható, ami azt jelenti, hogy

$$\begin{aligned} & \lim_{\lambda\to 0} \left(H(x_0(t)+\lambda v, u_0(t), p(t)) - H(x_0(t), u_0(t), p(t))\right) \qquad (9.6)\\ = \quad & \langle \partial_2 H(x_0(t),u_0(t),p(t)), v\rangle \\ = \quad & \langle -p(t), v\rangle, \end{aligned}$$

ahol felhasználtuk, hogy a tétel (2) feltétele, az adjungált feltétel szerint

$$\partial_2 H(x_0(t),u_0(t),p(t)) = -p'(t).$$

Mivel a Hamilton-függvény értékfüggvényének $x \mapsto H^\wedge(x,p)$ szelete konvex, ezért léteznek az alábbi határértékek, valamint a következő nagyságrendi viszony teljesül rájuk:

$$\begin{aligned} & \lim_{\lambda\to 0-} \frac{1}{\lambda} \left(H^\wedge(x_0(t)+\lambda v, p(t)) - H^\wedge(x_0(t),p(t))\right) \qquad (9.7)\\ \leq \quad & \lim_{\lambda\to 0+} \frac{1}{\lambda} \left(H^\wedge(x_0(t)+\lambda v, p(t)) - H^\wedge(x_0(t),p(t))\right). \end{aligned}$$

A Hamilton-függvény értékfüggvényének a definíciója szerint, továbbá a tétel (1) feltétele, a minimumfeltétel alapján

$$\begin{aligned} H(x,u(t),p(t)) \quad &\geq \quad H^\wedge(x,p(t)) \\ H(x_0(t),u_0(t),p(t)) \quad &= \quad H^\wedge(x_0(t),p(t)), \end{aligned}$$

ezért

$$
\begin{aligned}
& H^\wedge(x, p(t)) - H^\wedge(x_0(t), p(t)) \\
\leq \quad & H(x, u_0(t), p(t)) - H(x_0(t), u_0(t), p(t)),
\end{aligned} \tag{9.8}
$$

amiből (9.7) egyenlőtlenség felhasználásával kapjuk, hogy

$$
\begin{aligned}
& \lim_{\lambda \to 0-} \frac{1}{\lambda} \left(H(x_0(t) + \lambda v, u_0(t), p(t)) - H(x_0(t), u_0(t), p(t)) \right) \\
\leq \quad & \lim_{\lambda \to 0-} \frac{1}{\lambda} \left(H^\wedge(x_0(t) + \lambda v, p(t)) - H^\wedge(x_0(t), p(t)) \right) \\
\leq \quad & \lim_{\lambda \to 0+} \frac{1}{\lambda} \left(H^\wedge(x_0(t) + \lambda v, p(t)) - H^\wedge(x_0(t), p(t)) \right) \\
\leq \quad & \lim_{\lambda \to 0+} \frac{1}{\lambda} \left(H(x_0(t) + \lambda v, u_0(t), p(t)) - H(x_0(t), u_0(t), p(t)) \right).
\end{aligned}
$$

A (9.6) egyenlőséget is felhasználva látjuk, hogy végig egyenlőség van, továbbá

$$
\lim_{\lambda \to 0} \frac{1}{\lambda} \left(H^\wedge(x_0(t) + \lambda v, p(t)) - H^\wedge(x_0(t), p(t)) \right) = \langle -p(t), v \rangle .
$$

Mivel ez igaz tetszőleges v esetén, ezért ez azt jelenti, hogy a Hamilton-függvény $x \mapsto H^\wedge(x, p)$ értékfüggvénye Gâteaux-differenciálható a $(x_0(t), p(t))$ pontban, és

$$
\partial_2 H^\wedge(x_0(t), p(t)) = -p(t).
$$

Ismét a Hamilton-függvény $x \mapsto H^\wedge(x, p)$ értékfüggvénye konvex volta miatt

$$
\langle -p'(t), x(t) - x_0(t) \rangle \leq H^\wedge(x(t), p(t)) - H^\wedge(x_0(t), p(t)),
$$

amiből újra a (9.8) egyenlőtlenséget felhasználva adódik, hogy

$$
\langle -p'(t), x(t) - x_0(t) \rangle \leq H(x(t), u(t), p(t)) - H(x_0(t), u_0(t), p(t)).
$$

Ebből az egyenlőtlenségből pedig már nyilvánvalóan következik a Mangasarian-tétel bizonyításának az (9.5) egyenlőtlensége:

$$
\begin{aligned}
& \int_0^T H(x(t), u(t), p(t)) - H(x_0(t), u_0(t), p(t)) \; dt \\
\geq \quad & \int_0^T \langle -p'(t), x(t) - x_0(t) \rangle \; dt.
\end{aligned}
$$

□

9.4. Gyakorlatok

1. Keressük meg az alábbi irányítási feladat megoldását:

$$
\int_0^1 (1 - 4x(t) - 2u(t)^2) \, dt \to \max
$$

ahol
$$x'(t) = u(t)\,, \qquad x(0) = 0\,, \qquad x(1) \text{ szabad}$$
és $u(t) \in \mathbb{R}$. Használjuk a Mangasarian-tételt.

2. Oldjuk meg a következő irányítási feladatot:
$$\int_0^2 (u(t)^2 - x(t))\,dt \to \max\,, \qquad u(t) \in [0,1]\,,$$
ahol
$$x'(t) = u(t)\,, \qquad x(0) = 0\,, \qquad x(2) \text{ szabad.}$$
Használjuk az Arrow-féle elégséges feltételt.

10. fejezet

Optimális irányítási modellek

10.1. Az Atkinson-féle haszonmaximalizálási feladat

10.1 Példa. (Haszonmaximalizálás élethossziglan, Atkinson, 1971) Tekintsünk egy fogyasztót, aki a jelenlegi $t = 0$ időponttól a $t = T$ időpontig, amit a várható élettartamának gondol, maximális hasznosságot szándékozik elérni.

A fogyasztó úgy véli, hogy a $t \in [0, T]$ időpontban $y(t)$ jövedelemmel fog rendelkezni, a kamatláb pedig $r(t)$ lesz. Feltételezzük, hogy a $t \in [0, T]$ időpontbeli $c(t)$ fogyasztás mellett a $w(t)$ vagyoni helyzetét az ilyen modellekben szokásos differenciálegyenlet írja le:

$$w'(t) = r(t) \cdot w(t) + y(t) - c(t). \tag{10.1}$$

Tegyük fel, hogy a vagyona kezdetben w_0, a T időpontban pedig nullára csökken.

A fogyasztót egy a fogyasztásától függő $u : (0, \infty) \to \mathbb{R}$ hasznossági függvény jellemez, amelyről feltesszük, hogy kétszer deriválható, $u' > 0$ és $u'' < 0$. A $t \in [0, T]$ időpontbeli $c(t)$ fogyasztását pedig úgy választja meg, hogy a $[0, T]$ időszakbeli

$$\int_0^T e^{-\rho t} \cdot u(c(t))\ dt$$

összhasznossága maximális legyen, ahol ρ a diszkontkamatláb.

A fentiek a következő optimális irányítási feladatban foglalhatók össze:

$$\left\{ \begin{array}{l} \int_0^T e^{-\rho t} \cdot u(c(t))\ dt \to \max \\ w'(t) = r(t) \cdot w(t) + y(t) - c(t) \quad \text{m.m. } t \in [0, T] \\ c(t) \in (0, \infty) \\ w(0) = w_0,\ w(T) = 0 \end{array} \right. , \tag{10.2}$$

ahol a $w \in W^2[0,T]$ vagyon az állapotváltozó, a $c \in L^2[0,T]$ fogyasztás az irányítási változó, a $(0,\infty)$ nyílt intervallum pedig az irányítási tartomány, továbbá $r, y \in C[0,T]$.

A feladat megoldása:

A feladat Hamilton-függvénye:

$$H(t,w,c,p) = e^{-\rho t} \cdot u(c) + p \cdot (rw + y - c).$$

A maximumelv szerint, ha a (w,c) függvénypár megoldása a feladatnak, akkor van olyan p függvény, amelyre teljesülnek a következők:

(1) $\partial_3 H(t,w(t),c(t),p(t)) = 0$ ($(0,\infty)$ nyílt volta miatt), ami azt jelenti, hogy $e^{-\rho t} \cdot u'(c(t)) - p(t) = 0$, azaz

$$p(t) = e^{-\rho t} \cdot u'(c(t)).$$

(2) $p'(t) = -\partial_2 H(t,w(t),c(t),p(t))$, tehát

$$p'(t) = -p(t) \cdot r(t),$$

ez egy homogén lineáris differenciálegyenlet, aminek a megoldása:

$$p(t) = p(0) \cdot e^{-\int_0^t r(\tau)d\tau},$$

emiatt (1) szerint

$$e^{-\rho t} \cdot u'(c(t)) = p(0) \cdot e^{-\int_0^t r(\tau)d\tau}. \tag{10.3}$$

Ebből az is látszik, hogy $u' > 0$ volta miatt minden $t \in [0,T]$ esetén $p(t) > 0$.

Végül a (10.1) inhomogén lineáris differenciálegyenlet megoldása, mint ismert:

$$w(t) = e^{\int_0^t r(\tau)d\tau} \cdot \left(w_0 + \int_0^\tau e^{-\int_0^t r(s)ds} \cdot (y(\tau) - c(\tau))\ d\tau \right).$$

Mivel az u hasznossági függvény konkáv volta miatt a

$$(w,c) \mapsto H(t,w,c,p)$$

függvény konkáv, ezért a Mangasarian-tétel értelmében (w,c) valóban optimális folyamat.

Speciális eset 1.: A (10.3) egyenlet ilyen általánosan nem oldható meg, ezért tegyük fel, hogy az r kamatláb konstans, és megegyezik a diszkontkamatlábbal:

$$\forall\ t \in [0,T] \quad \text{esetén} \quad r(t) = \rho.$$

Ekkor a (10.3) egyenlet a következő alakú:

$$e^{-\rho t} \cdot u'(c(t)) = p(0) \cdot e^{-\rho t}, \quad \text{azaz} \quad u'(c(t)) = p(0).$$

Mivel $u'' < 0$, azaz u' szigorúan monoton fogyó, ezért c konstans függvény, azaz minden $t \in [0, T]$ esetén $c(t) = c_0$.

Ebben az esetben a (10.1) differenciálegyenlet

$$w'(t) = \rho \cdot w(t) + y(t) - c_0,$$

alakú, amiből a vagyon alakulása:

$$\begin{aligned} w(t) &= e^{\int_0^t \rho d\tau} \cdot \left(w_0 + \int_0^\tau e^{-\int_0^t \rho ds} \cdot (y(\tau) - c_0)\, d\tau \right) \\ &= e^{\rho t} \cdot \left(w_0 + \frac{c_0}{\rho}(e^{-\rho t} - 1) + \int_0^t e^{-\rho\tau} \cdot y(\tau) d\tau \right). \end{aligned}$$

Mivel $w(T) = 0$, ezért

$$e^{\rho T} \cdot \left(w_0 + \frac{c_0}{\rho}(e^{-\rho t} - 1) + \int_0^T e^{-\rho\tau} \cdot y(\tau) d\tau \right) = 0,$$

amiből kiszámolható a fogyasztás:

$$c(t) = c_0 = \frac{\rho}{1 - e^{-\rho T}} \cdot \left(w_0 + \int_0^T e^{-\rho\tau} \cdot y(\tau) d\tau \right).$$

Speciális eset 2.: A (10.3) egyenlet megoldása érdekében a fentiek helyett azt tegyük fel, hogy az r kamatláb konstans:

$$\forall\, t \in [0, T] \quad \text{esetén} \quad r(t) = r_0,$$

de ez nem feltétlenül egyezik meg a diszkontkamatlábbal. Továbbá tegyük még azt is fel, hogy a hasznossági függvény speciálisan a következő alakú:

$$u(c) = \begin{cases} \frac{(c-c_1)^{1-\varepsilon}}{1-\varepsilon} & \text{ha } \varepsilon > 0,\ \varepsilon \neq 1 \\ \ln(c - c_1) & \text{ha } \varepsilon = 1 \end{cases},$$

ahol $c_1 \in \mathbb{R}$ adott paraméter. Ekkor $u'(c) = (c - c_1)^{-\varepsilon}$. Megjegyezzük, hogy $c_1 = 0$ esetén az u' függvény rugalmassága állandó, méghozzá ε, ugyanis

$$u''(c) \cdot \frac{c}{u'(c)} = -\varepsilon \cdot c^{-\varepsilon-1} \cdot \frac{c}{c^{-\varepsilon}} = -\varepsilon.$$

Az (1) feltétel szerint

$$p(t) = e^{-\rho t} \cdot u'(c(t)) = e^{-\rho t} \cdot (c(t) - c_1)^{-\varepsilon},$$

a (2) feltétel szerint pedig

$$p(t) = p(0) \cdot e^{-\int_0^t r_0 d\tau} = p(0) \cdot e^{-r_0 t},$$

így

$$e^{-\rho t} \cdot (c(t) - c_1)^{-\varepsilon} = p(0) \cdot e^{-r_0 t},$$

amiből a fogyasztás:

$$c(t) = c_1 + p^{-\frac{1}{\varepsilon}}(0) \cdot e^{\frac{r_0 - \rho}{\varepsilon} t}. \tag{10.4}$$

Ebben az esetben a (10.1) differenciálegyenlet

$$\begin{aligned} w'(t) &= r_0 \cdot w(t) + y(t) - c_1 - p^{-\frac{1}{\varepsilon}}(0) \cdot e^{\frac{r_0 - \rho}{\varepsilon} t} \\ &= r_0 \cdot w(t) + y(t) - c_1 - A \cdot e^{\gamma t}, \end{aligned}$$

alakú, ahol $A = p^{-\frac{1}{\varepsilon}}(0)$ és $\gamma = \frac{r_0 - \rho}{\varepsilon}$. Ennek a megoldása:

$$\begin{aligned} w(t) &= e^{\int_0^t r_0 d\tau} \cdot \left(w_0 + \int_0^t e^{-\int_0^t r_0 ds} \cdot \left(y(\tau) - c_1 - A \cdot e^{\gamma t}\right) d\tau \right) \\ &= e^{r_0 t} \left(w_0 + \int_0^t e^{-r_0 \tau} y(\tau) d\tau + \frac{c_1}{r_0}(e^{-r_0 t} - 1) - \frac{A}{\gamma - r_0}(e^{(\gamma - r_0)t} - 1) \right) \end{aligned}$$

a t időpontbeli vagyon.

Mivel $w(T) = 0$, ezért a fenti összefüggés T-beli értékéből A kiszámolható:

$$A = \frac{\gamma - r_0}{e^{(\gamma - r_0)T} - 1} \cdot \left(w_0 + \int_0^T e^{-r_0 \tau} y(\tau) d\tau + \frac{c_1}{r_0}(e^{-r_0 T} - 1) \right).$$

A modellben kapott eredmények érdekes összefüggésekre mutatnak rá. A (10.4) képlet alapján meghatározott fogyasztás növekszik, ha a kamat nagyobb mint a diszkontráta, azaz $r_0 > \rho$, viszont csökken, ha kisebb, azaz $r_0 < \rho$, valamint stagnál, ha egyenlők, azaz $r_0 = \rho$. Továbbá, ha a c_1 paraméter pozitív, akkor ez tekinthető úgy, mint a fogyasztás minimális szintje.

10.2. Monopólium befektetési problémája

10.2 Példa. (Monopólium befektetési problémája) Egy monopólium a $[0, T]$ tervezési időszakban egyféle árut termel, és maximális profitra törekszik.

A monopólium a $t \in [0, T]$ időpontban az áru $K(t)$ mennyiségű tőkével rendelkezik, valamint $I(t)$ beruházást hajt végre. A tőke megváltozását a következő, a beruházástól és egy $\delta > 0$ paramétertől függő differenciálegyenlet írja le:

$$K'(t) = -\delta K(t) + I(t). \tag{10.5}$$

A gazdaság kibocsátását egy a tőkétől függő $F(K)$ termelési függvény, a beruházás költségét pedig egy $C(I)$ költségfüggvény határozza meg. Feltéve, hogy a termék ára P, a monopólium profitja egy $t \in [0,T]$ időpontban

$$P \cdot f(K(t)) - C(I(t)).$$

A monopólium feladata olyan tőke-beruházás függvénypár kialakítása, amely mellett a $[0,T]$ időszakbeli

$$\int_0^T P \cdot f(K(t)) - C(I(t))dt$$

profitja maximális. Tegyük fel, hogy a monopólium kezdeti tőkéje K_0.

A fentiek a következő optimális irányítási feladatra vezetnek:

$$\begin{cases} \int_0^T P \cdot f(K(t)) - C(I(t))dt \to \max \\ K'(t) = -\delta K(t) + I(t) \\ I(t) \in (0,\infty) \\ K(0) = K_0,\ K(T) \text{ szabad} \end{cases}, \tag{10.6}$$

ahol a $K \in W^2[0,T]$ tőke az állapotváltozó, az $I \in L^2[0,T]$ beruházás az irányítási változó, a $(0,\infty)$ nyílt intervallum pedig az irányítási tartomány, továbbá f és C folytonos függvények.

A feladat megoldása:

A feladat Hamilton-függvénye:

$$H(t,K,I,p) = P \cdot f(K) - C(I) + p \cdot (-\delta K + I).$$

A maximumelv szerint, ha a (K,I) függvénypár megoldása a feladatnak, akkor van olyan p függvény, amelyre teljesülnek a következők:

(1) $\partial_3 H(t,K(t),I(t),p(t)) = 0$ (a $(0,\infty)$ nyílt volta miatt), ami azt jelenti, hogy
$$p(t) = C'(I(t)).$$

(2) $p'(t) = -\partial_2 H(t,K(t),I(t),p(t))$, azaz
$$p'(t) = -P \cdot f'(K(t)) + p(t) \cdot \delta.$$

(3) $p(T) = 0$.

Ha feltesszük azt, hogy az f termelési függvény konkáv, és a c költségfüggvény konvex, akkor a
$$(K,I) \mapsto H(t,K,I,p)$$
függvény konkáv, ezért a Mangasarian-tétel értelmében (K,I) valóban optimális folyamat.

Speciális eset: Legyen a termelési függvény egy $\alpha > 0$ paraméterű

$$f(K) = -\alpha K^2 + K$$

másodfokú polinom, a költségfüggvény pedig

$$C(I) = I^2.$$

Ekkor a fenti feltételek:

(1) $p(t) = 2I(t)$, így $p'(t) = 2I'(t)$.

(2) $p'(t) = 2\alpha P \cdot K(t) - P + \delta \cdot p(t)$.

(3) $p(T) = 0$.

Ezek szerint

$$2I'(t) = 2\alpha P \cdot K(t) - P + 2\delta \cdot I(t).$$

A tőke megváltozását leíró (10.5) differenciálegyenlet szerint

$$I(t) = K'(t) + \delta K(t), \quad \text{így} \quad I'(t) = K''(t) + \delta K'(t),$$

ezeket behelyettesítve a fenti egyenletbe a $K(T) = K_0$ kezdeti feltétellel együtt a következő inhomogén másodrendű differenciálegyenlethez jutunk:

$$\begin{cases} K''(t) = (\alpha P + \delta^2)K(t) - \frac{P}{2} \\ K(0) = K_0 \end{cases}.$$

10.3. A Shell-féle növekedési modell

10.3 Példa. (A gazdasági növekedés Shell-féle modellje, 1967) Tekintsünk egy egyszektoros gazdaságot a $[0, T]$ időintervallumon.

A gazdaságban a $t \in [0, T]$ időpontban $k(t)$ tőkemennyiség áll rendelkezésre. A tőke kezdeti értéke $k_0 > 0$, valamint T-beli értéke az elvárások szerint legalább k_T, ahol $k_T > k_0$. Az amortizációs normát jelölje a $\lambda > 0$ paraméter.

A gazdaság kibocsátását egy a tőkétől függő $f(k)$ termelési függvény határozza meg, amelyről feltesszük, hogy deriválható és $f'(k) > 0$. Tegyük még azt is fel, hogy amennyiben $k > k_0 e^{-\lambda T}$, úgy $f(k) > 0$.

A tőke növekedését a kibocsátás $s(t) \in [0, 1]$ hányadának a megtakarításával érik el. A tőke növekedését ezért a következő differenciálegyenlet írja le:

$$k'(t) = e^{rt} s(t) f(k(t)) - \lambda k(t),$$

ahol az $r > 0$ paraméter a kamatlábat jelöli.

A gazdaság szereplői az $s(t)$ megtakarítási hányadot úgy választják meg, hogy a $[0, T]$ időszakbeli

$$\int_0^T e^{-\rho t} e^{rt} (1 - s(t)) f(k(t))\, dt$$

összfogyasztás maximális legyen, ahol a $\rho > 0$ paraméter a kamatlábat jelöli.

A fentiek a következő optimális irányítási feladatban foglalhatók össze:

$$\left\{ \begin{array}{l} \int_0^T e^{(r-\rho)t} (1 - s(t)) f(k(t))\, dt \to \max \\ k'(t) = e^{rt} s(t) f(k(t)) - \lambda k(t) \quad \text{m.m. } t \in [0, T] \\ s(t) \in [0, 1] \\ k(0) = k_0, \ k(T) = k_T \end{array} \right. , \tag{10.7}$$

ahol a $k \in W^2[0, T]$ tőke az állapotváltozó, az $s \in L^2[0, T]$ megtakarítási hányad az irányítási változó, a $[0, 1]$ zárt intervallum pedig az irányítási tartomány, továbbá f deriválható függvény.

ahol az állapotváltozó a k tőke, az irányítási változó az s megtakarítási hányad, az irányítási tartomány pedig a $[0, 1]$ intervallum.

A feladat megoldása:

A feladat Hamilton-függvénye:

$$\begin{array}{rcl} H(t, k, s, p) & = & e^{(r-\rho)t} (1 - s) f(k) + p \cdot (e^{rt} s f(k) - \lambda k) \\ & = & e^{(r-\rho)t} f(k) + (p - e^{-\rho t}) e^{rt} f(k) s - p \lambda k. \end{array}$$

A maximumelv szerint, ha a (k, s) függvénypár megoldása a feladatnak, akkor van olyan p függvény, amelyre teljesülnek a következők:

(1) Maximumfeltétel:

$$\begin{array}{rl} \max\limits_{s \in [0,1]} & e^{(r-\rho)t} f(k(t)) + (p(t) - e^{-\rho t}) e^{rt} f(k(t)) s - \lambda p(t) k(t) \\ = & e^{(r-\rho)t} f(k(t)) + (p(t) - e^{-\rho t}) e^{rt} f(k(t)) s(t) - \lambda p(t) k(t), \end{array}$$

ezek szerint

$$s(t) = \left\{ \begin{array}{ll} 1 & \text{ha } p(t) > e^{-\rho t} \\ 0 & \text{ha } p(t) < e^{-\rho t} \end{array} \right. .$$

(2) Adjungált egyenlet:

$$\begin{array}{rcl} p'(t) & = & -\partial_2 H(t, k(t), s(t), p(t)) \\ & = & -e^{(r-\rho)t} (1 - s(t)) f'(k(t)) - e^{rt} p(t) s(t) f'(k(t)) + \lambda p(t). \end{array}$$

Speciális eset: A (10.7) egyenlet ilyen általánosan nem oldható meg, ezért tegyük fel, hogy az r kamatláb, a ρ diszkontkamatláb és a λ amortizációs norma nulla, a termelési függvény pedig

$$f(k) = ak$$

alakú, ahol $a > 0$ paraméterre teljesül, hogy

$$k_T < k_0 e^{-aT}. \tag{10.8}$$

A (10.7) feladat a következő feladatra egyszerűsödik:

$$\left\{ \begin{array}{l} \int_0^T (1 - s(t))ak(t)\ dt \to \max \\ k'(t) = s(t)ak(t) \ \ \text{m.m.}\ \ t \in [0, T] \\ s(t) \in [0, 1] \\ k(0) = k_0,\ \ k(T) = k_T \end{array} \right. .$$

A feladat megoldása:

A feladat Hamilton-függvénye:

$$H(t, k, s, p) \quad = \quad (1 - s)ak + p \cdot sak = ak + (p - 1)aks.$$

A maximumelv szerint, ha a (k, s) függvénypár megoldása a feladatnak, akkor van olyan p függvény, amelyre teljesülnek a következők:

(1) Maximumfeltétel:

$$\max_{s \in [0,1]} ak(t) + (p(t) - 1)ak(t)s = ak(t) + (p(t) - 1)ak(t)s(t),$$

ezek szerint

$$s(t) = \left\{ \begin{array}{ll} 1 & \text{ha } p(t) > 1 \\ 0 & \text{ha } p(t) < 1 \end{array} \right. . \tag{10.9}$$

(2) Adjungált egyenlet:

$$\begin{aligned} p'(t) &= -\partial_2 H(t, k(t), s(t), p(t)) \\ &= -a - (p(t) - 1)as(t) \\ &= \left\{ \begin{array}{ll} -ap(t) & \text{ha } p(t) > 1 \\ -a & \text{ha } p(t) < 1 \end{array} \right. . \end{aligned} \tag{10.10}$$

Ebből következik, hogy $p'(t) < 0$, így p szigorúan monoton fogyó.

Látható, hogy $p(0) > 1$. Ugyanis, indirekt módon tegyük fel, hogy $p(0) \leq 1$, ekkor p szigorúan monoton fogyó volta miatt minden $t \in (0, T]$ esetén $p(t) < 1$, így $s(t) = 0$, ezért viszont $k'(t) = 0$, amiből

$$k(t) = k_0 + \int_0^t k'(\tau) d\tau = k_0,$$

speciálisan $k(T) = k_0$, de a feltétel szerint $k(T) = k_T > k_0$, ami ellentmondás.

Látható az is, hogy $p(T) < 1$. Indirekt módon tegyük fel, hogy $p(T) \geq 1$. Mivel p szigorúan monoton fogyó, ezért minden $t \in [0, T)$ esetén $p(t) > 1$, így a (10.9) szerint $s(t) = 1$, ezek szerint

$$k'(t) = s(t)ak(t) = ak(t),$$

amiből $k(t) = k_0 e^{at}$, speciálisan $k(T) = k_0 e^{aT}$. Az a paraméterre a (10.8)-ben feltettük, hogy $k_T < k_0 e^{-aT}$, emiatt $k_T < k(T)$. A feltétel szerint viszont $k(T) = k_T$, ami ellentmondás.

Mivel p szigorúan monoton fogyó, folytonos függvény, valamint $p(0) > 1$, $p(T) < 1$, ezért létezik olyan $t_0 \in (0,T)$, hogy $p(t_0) = 1$. Ezek szerint

$$p|_{[0,t_0)} > 1 \quad \text{és} \quad p|_{(t_0,T]} < 1.$$

A (10.10) szerint minden $t \in (t_0, T]$ esetén $p'(t) = -a$, így

$$p(t) = p(t_0) + \int_{t_0}^{t} p'(\tau) d\tau = a(t_0 - t) + 1,$$

továbbá minden $t \in [0, t_0)$ esetén $p'(t) = -ap(t)$, így

$$p(t) = p(t_0) e^{a(t_0 - t)} = e^{a(t_0 - t)}.$$

A (10.9) szerint

$$s(t) = \begin{cases} 1 & \text{ha } t \in [0, t_0) \\ 0 & \text{ha } t \in (t_0, T] \end{cases}.$$

Végül mivel

$$k'(t) = s(t) a k(t) = \begin{cases} ak(t) & \text{ha } t \in [0, t_0) \\ 0 & \text{ha } t \in (t_0, T] \end{cases},$$

ezért

$$k(t) \quad = \quad k_0 + \int_0^t k'(\tau) d\tau = \begin{cases} k_0 e^{at} & \text{ha } t \in [0, t_0) \\ k(t_0) = k_0 e^{at_0} & \text{ha } t \in (t_0, T] \end{cases}.$$

A feladat értékfüggvénye

$$V(k_0, k_T, T) = \int_0^T (1 - s(t)) a k(t) \, dt = \int_{t_0}^T a k_0 e^{a t_0} = a k_0 e^{a t_0} (T - t_0).$$

A feltétel szerint $k(T) = k_T$, azaz $k_T = k_0 e^{a t_0}$ ami azt jelenti, hogy

$$t_0 = \frac{1}{a} \ln \frac{k_T}{k_0}.$$

Vegyük észre, hogy az a paraméterre (10.8)-ben tett $k_T < k_0 e^{aT}$ feltevés biztosítja, hogy $t_0 < T$, azaz $t_0 \in (0, T)$. Az értékfüggvény

$$V(k_0, k_T, T) = a k_T \left(T - \frac{1}{a} \ln \frac{k_T}{k_0} \right),$$

amiből látható, hogy

$$\partial V / \partial k_0 = \tfrac{k_T}{k_0} = p(t),$$

$$\partial V/\partial k_T = a(t - t_0) - 1 = -p(T),$$
$$\partial V/\partial T = ak_T = ak(T) = H(T).$$

A fenti (k,s) függvénypár valóban megoldása is a feladatnak. Bár a feladat Hamilton-függvényének a $(k,s) \mapsto H(t,k,s,p) = ak + (p-1)aks$ szelete nem konkáv (a ks szorzat miatt), ezért a Mangasarian-tétel nem alkalmazható ennek az igazolásához. A Hamilton-függvény értékfüggvénye

$$H^\vee(t,k,p) = \max_{s\in[0,1]} H(t,k,s,p) = \begin{cases} pak & \text{ha } p > 1 \\ ak & \text{ha } p < 1 \end{cases},$$

ezért a $k \mapsto H^\vee(t,k,p)$ függvény lineáris, így konkáv, ezek szerint az Arrow-tétel alapján (k,s) optimális folyamat.

10.4. A Hotelling-szabály

10.4 Példa. (Olajkitermelés, Hotelling-szabály) Tegyük fel, hogy egy gazdaságnak K mennyiségű olajvagyona van a felmérések szerint. Ez szándékoznak kitermelni valamely $[0,T]$ időintervallumban, ahol az időintervallum hossza is szabadon megválasztható. Egy $t \in [0,T]$ időpontban $u(t)$ mennyiséget termelnének ki, a még rendelkezésre álló olajvagyon pedig $x(t)$. Látható, hogy $x'(t) = -u(t)$, ugyanis

$$x(t) = K - \int_0^t u(\tau)d\tau.$$

A kitermelés költsége a t időponttól és az u kitermelt mennyiségtől függő $C(t,u)$ költségfüggvény írja le, amelyről tegyük fel, hogy pozitív, továbbá $\partial_2\partial_2 C(t,u) > 0$, azaz $u \mapsto C(t,u)$ konvex függvény. Feltéve, hogy az olaj világpiaci ára $P(t)$, az olajból elért profit:

$$P(t)u(t) - C(t,u(t)).$$

Az $u(t)$ kitermelési mennyiséget úgy kell megválasztani, hogy a $[0,T]$ időszakbeli

$$\int_0^T e^{-\rho t}(P(t)u(t) - C(t,u(t)))\,dt$$

összprofit maximális legyen, ahol a $\rho > 0$ paraméter a diszkontkamatlábat jelöli.

A fentiek a következő optimális irányítási feladatban foglalhatók össze:

$$\begin{cases} \int_0^T e^{-\rho t}(P(t)u(t) - C(t,u(t)))\,dt \to \max \\ x'(t) = -u(t) \quad \text{m.m. } t \in [0,T] \\ u(t) \in \mathbb{R}_+ \\ x(0) = K,\ x(T) = 0 \\ T \in \mathbb{R}_+ \end{cases}, \tag{10.11}$$

ahol az $x \in W^2[0,T]$ olajvagyon az állapotváltozó, az $u \in L^2[0,T]$ kitermelés az irányítási változó, az $\mathbb{R}_+$ intervallum pedig az irányítási tartomány.

A feladat megoldása:

A feladat Hamilton-függvénye:

$$H(t,x,u,p) = e^{-\rho t}(Pu - C(t,u)) - pu.$$

A maximumelv szerint, ha az (x,u) függvénypár megoldása a feladatnak, akkor van olyan p függvény, amelyre teljesülnek a következők:

(1) Maximumfeltétel:

$$\begin{aligned} \max_{u \geq 0} \quad & e^{-\rho t}(P(t)u - C(t,u)) - p(t)u \\ = \quad & e^{-\rho t}(P(t)u(t) - C(t,u(t))) - p(t)u(t). \end{aligned}$$

(2) Adjungált egyenlet:

$$p'(t) = -\partial_2 H(t,x(t),u(t),p(t)) = 0.$$

(3) Az intervallum végpontjára vonatkozó feltétel:

$$H(T,x(T),u(T),p(T)) = (P(T)u(T) - C(T,u(T)))e^{-\rho T} - p(T)u(T) = 0.$$

Mivel feltétel szerint az $u \mapsto C(t,u)$ költségfüggvény konvex, emiatt az $u \mapsto H(t,x,u,p)$ függvény konkáv, ezért a maximumfeltétel ekvivalens azzal, hogy

$$0 \leq \partial_3 H(t,x(t),u(t),p()t) = e^{-\rho t}(P(t) - \partial_2 C(t,u(t))) - p(t),$$

azaz

$$P(t) - \partial_2 C(t,u(t)) \geq p(t)e^{\rho t},$$

továbbá $u(t) > 0$ esetén

$$P(t) - \partial_2 C(t,u(t)) = p(t)e^{\rho t}. \tag{10.12}$$

Az adjungált egyenlet szerint p konstans függvény, például $p = p_0$, ahol $p_0 \in \mathbb{R}$. A T-re vonatkozó (4) feltétel pedig azzal ekvivalens, hogy

$$P(T)u(T) - C(T,u(T)) = e^{\rho T} p_0 u(T).$$

Ebből látszik, hogy $u(T) > 0$, különben a baloldal negatív, a jobboldal nulla volna. Osztva $u(T)$-vel azt kapjuk, hogy

$$P(T) - \frac{C(T,u(T))}{u(T)} = e^{\rho T} p_0.$$

Mivel $u(T) > 0$, ezért (10.12) szerint

$$P(T) - \partial_2 C(T,u(T)) = p_0 e^{\rho T}.$$

E két egyenlőségből következik az úgynevezett Hotelling-szabály, ami szerint az olajkitermelés határköltsége megegyezik az átlagköltséggel:

$$\partial_2 C(T, u(T)) = \frac{C(T, u(T))}{u(T)}.$$

Mivel a c költségfüggvény konvex, ezért

$$(x, u) \mapsto H(t, x, u, p)$$

függvény konkáv, ezért ha az (x, u) pár kielégíti a fenti feltételeket, akkor a Mangasarian-tétel értelmében optimális folyamat.

Speciális eset:

Legyen $P(t) = ae^{\alpha t}$, $C(t, u(t)) = u^2(t)e^{\beta t} - c$, ahol a, α, β, c pozitív paraméterek.

10.5. Egy makroökonómiai beruházási modell

10.5 Példa. (Egyszektoros makroökonómiai modell, beruházási probléma) Tekintsünk egy egyszektoros gazdaságot a $[0, T]$ időszakban.

A gazdaság kibocsátása a $t \in [0, T]$ időpontban $y(t)$, ami kezdetben y_0.

A termelésének a növekedését a megtermelt javak $i(t) \in [0, 1]$ hányadának a ráfordításával lehet elérni:

$$y'(t) = ai(t)y(t),$$

ahol a pozitív állandó. A megtermelt javak $1 - i(t) \in [0, 1]$ hányadát fogyasztásra lehet használni:

$$y'(t) = (1 - i(t))y(t).$$

A gazdaságban az $i(t)$ beruházási hányadot úgy kell megválasztani, hogy a $[0, T]$ időszak

$$\int_0^T (1 - i(t))y(t)\ dt$$

összfogyasztása maximális legyen.

A fentiek a következő optimális irányítási feladatban foglalhatók össze:

$$\begin{cases} \int_0^T (1 - i(t))y(t)\ dt \to \max \\ y'(t) = ai(t)y(t) \quad \text{m.m. } t \in [0, T] \\ i(t) \in [0, 1] \\ y(0) = y_0, \ y(T) \text{ szabad} \end{cases}, \tag{10.13}$$

ahol az $y \in W^2[0, T]$ kibocsátás (termelés) az állapotváltozó, az $i \in L^2[0, T]$ beruházási hányad az irányítási változó, a $[0, 1]$ zárt intervallum pedig az irányítási tartomány.

A feladat megoldása:

A feladat Hamilton-függvénye:

$$H(t, y, i, p) = (1 - i)y + piy = y + (p - 1)iy.$$

A maximumelv szerint, ha az (y, i) függvénypár megoldása a feladatnak, akkor van olyan p függvény, amelyre teljesülnek a következők:

(1) Maximumfeltétel:

$$\max_{i \in [0,1]} y(t) + (p(t) - 1)iy(t) = y(t) + (p(t) - 1)i(t)y(t).$$

(2) Adjungált egyenlet:

$$p'(t) = -1 - (p(t) - 1)i(t) = -i(t)p(t) + i(t) - 1.$$

(3) Transzverzalitási feltétel: $p(T) = 0$.

Az y-ra vonatkozó differenciálegyenletből:

$$y(t) = y_0 e^{\int_0^t ai(\tau)d\tau},$$

ami pozitív, ezért a maximumfeltétel szerint

$$i(t) = \begin{cases} 0 & \text{ha } p(t) < 1 \\ 1 & \text{ha } p(t) > 1 \end{cases}.$$

Az adjungált egyenlet szerint

$$p_1'(t) = -i(t)p(t) + i(t) - 1 = \begin{cases} -1 & \text{ha } p(t) < 1 \\ -p(t) & \text{ha } p(t) > 1 \end{cases}.$$

Mivel a transzverzalitási feltétel szerint $p(T) = 0 < 1$, ezért $i(T) = 0$, így $p'(T) = -1$. Emiatt a T egy baloldali környezetében p szigorúan monoton fogy, sőt $p(t) = -t + c$, attól a t_0 ponttól fogva, amelyre $p(t_0) = 1$. Mivel $0 = p(T) = -T + c$, ezért $c = T$, így $p(t) = T - t$. Mivel $1 = p(t_0) = T - t_0$, ezért $t_0 = T - 1$. Összefoglalva ez azt jelenti, hogy

$$\forall\ t \in [T - 1, T] \quad \text{esetén} \quad p(t) = T - t.$$

Legyen $t \in [0, T - 1)$ tetszőleges pont. Ha $p'(t) = -1$, akkor $p(t) = T - t > 1$, ami ami ellentmond az adjungált egyenletnek, ezért

$$p'(t) = -p(t).$$

Mivel $p(T - 1) = 1$, ezért

$$\forall\ t \in [0, T - 1] \quad \text{esetén} \quad p(t) = e^{-t+T-1}.$$

Ebből a beruházási hányad

$$i(t) = \begin{cases} 1 & \text{ha } t \in [0, T-1) \\ 0 & \text{ha } t \in (T-1, T] \end{cases} .$$

Ez azt jelenti, hogy a $T-1$ időpontig minden megtermelt terméket beruházásra kell fordítani, majd ettől az időponttól érdemes minden terméket a fogyasztásra fordítani.

A fentiekből meghatározható a termelés a $[0, T]$ intervallumon. Mivel

$$x'(t) = ai(t)x(t) = \begin{cases} ax(t) & \text{ha } t \in [0, T-1) \\ 0 & \text{ha } t \in (T-1, T] \end{cases} ,$$

ezért

$$x(t) = \begin{cases} e^{at} & \text{ha } t \in [0, T-1] \\ e^{a(T-1)} & \text{ha } t \in (T-1, T] \end{cases} .$$

A fenti (y, i) függvénypár nem biztos, hogy valóban megoldása is a feladatnak. Mivel az $(y, i) \mapsto H(t, y, i, p) = y+(p-1)iy$ Hamilton-függvény nem konkáv (az iy szorzat miatt), ezért a Mangasarian-tétel nem alkalmazható ennek az eldöntésére. Azonban az Arrow-tétel alapján látható, hogy valóban megoldás. A Hamilton-függvény értékfüggvénye

$$H^{\vee}(t, y, p) = \max_{i \in [0,1]} H(t, y, i, p) = \begin{cases} 2y & \text{ha } p_1 \geq 1 \\ y & \text{ha } p_1 < 1 \end{cases} ,$$

ezért az $y \mapsto H^{\vee}(t, y, p)$ függvény lineáris, így konkáv.

11. fejezet

Többszektoros irányítási modellek

11.1. Kétszektoros makroökonómiai modell

11.1 Példa. (Kétszektoros makroökonómiai modell, Mahalanobis) (Ez a példa a 10.5. példa kétszektoros változata.) Tekintsünk egy kétszektoros gazdaságot a $[0, T]$ tervezési időszakban. Az első szektor beruházási javakat, a második szektor fogyasztási javakat termel.

A gazdaság kibocsátása a $t \in [0, T]$ időpontban a beruházási javakból $y_1(t)$, ami jelenleg y_1^0, a fogyasztási javakból $y_2(t)$, ami jelenleg y_2^0.

A beruházási javak termelésének a növekedését a megtermelt beruházási javak $i(t) \in [0, 1]$ hányadának a ráfordításával érik el:

$$y_1'(t) = ai(t)y_1(t),$$

ahol a pozitív állandó. A fogyasztási javak termelésének a növekedését a megtermelt beruházási javak $1 - i(t) \in [0, 1]$ másik hányadának a ráfordításával érik el:

$$y_2'(t) = a(1 - i(t))y_1(t).$$

A gazdaságban az $i(t)$ beruházási hányadot úgy kell megválasztani, hogy a fogyasztási javak $[0, T]$ időszakban termelt

$$\int_0^T y_2 \, dt$$

mennyisége maximális legyen.

A fentiek a következő optimális irányítási feladatban foglalhatók össze:

$$\left\{\begin{array}{l} \int_0^T y_2\ dt \to \max \\ \left\{\begin{array}{l} y_1'(t) = ai(t)y_1(t) \ \text{ m.m. } t \in [0,T] \\ y_2'(t) = a(1-i(t))y_1(t) \ \text{ m.m. } t \in [0,T] \end{array}\right. \\ i(t) \in [0,1] \\ \left\{\begin{array}{l} y_1(0) = y_1^0,\ y_1(T) \text{ szabad} \\ y_2(0) = y_2^0,\ y_2(T) \text{ szabad} \end{array}\right. \end{array}\right. , \tag{11.1}$$

ahol az $(y_1, y_2) \in W^2_{\mathbb{R}^2}[0,T]$ kibocsátásvektor az állapotváltozó, az $i \in L^2[0,T]$ beruházási hányad az irányítási változó, a $[0,1]$ zárt intervallum pedig az irányítási tartomány.

A feladat megoldása:

A feladat Hamilton-függvénye:

$$\begin{aligned} H(t,(y_1,y_2),i,(p_1,p_2)) &= y_2 + \langle (p_1,p_2),(aiy_1, a(1-i)y_2)\rangle \\ &= y_2 + (p_1 - p_2)aiy_1 + p_2 a y_1. \end{aligned}$$

A maximumelv szerint, ha az $((y_1,y_2),i)$ függvénypár megoldása a feladatnak, akkor van olyan (p_1,p_2) függvény, amelyre teljesülnek a következők:

(1) Maximumfeltétel:

$$\begin{aligned} \max_{i\in[0,1]} \quad & y_2(t) + (p_1(t) - p_2(t))aiy_1(t) + p_2(t)ay_1(t) \\ = \quad & y_2(t) + (p_1(t) - p_2(t))ai(t)y_1(t) + p_2(t)ay_1(t). \end{aligned}$$

(2) Adjungált egyenlet:

$$\begin{bmatrix} p_1'(t) \\ p_2'(t) \end{bmatrix} = -\begin{bmatrix} a(p_1(t)i(t) + p_2(t)(1-i(t)) \\ 1 \end{bmatrix}$$

(3) Transzverzalitási feltétel: $p_1(T) = 0$ és $p_2(T) = 0$.

Az y_1-re vonatkozó differenciálegyenletből:

$$y_1(t) = y_1^0 e^{\int_0^t ai(\tau)d\tau},$$

amely pozitív, ezért az (1) a maximumfeltétel szerint

$$i(t) = \left\{\begin{array}{ll} 0 & \text{ha } p_1(t) < p_2(t) \\ 1 & \text{ha } p_1(t) > p_2(t) \end{array}\right. . \tag{11.2}$$

A (2) adjungált egyenlet szerint $p_2'(t) = -1$, ezért

$$p_2(t) = p_2(T) + \int_T^t p_2'(\tau)d\tau = T - t.$$

Továbbá

$$p_1'(t) = -a(p_1(t)i(t) + p_2(t)(1 - i(t)), \tag{11.3}$$

ebből a (3) transzverzalitási feltételt is felhasználva következik, hogy

$$p_1'(T) = -a(p_1(T)i(T) + p_2(T)(1 - i(T)) = 0,$$

A fentieket összefoglalva:

$$p_1(T) = p_2(T) = 0, \quad \text{illetve} \quad p_1'(T) = 0,\ p_2'(T) = -1,$$

ezek szerint T-nek van olyan baloldali környezete, amelyen $p_1(t) < p_2(t)$. Jelölje $t_0 \in [0, T)$ azon $t \in [0, T)$pontok szuprémumát, amelyre $p_1(t) \geq p_2(t)$, lehet, hogy $t_0 = 0$. Ekkor

$$\forall\ t \in (t_0, T) \quad \text{esetén} \quad p_1(t) < p_2(t). \tag{11.4}$$

Emiatt a (11.2) szerint $\forall\ t \in (t_0, T)$ esetén $i(t) = 0$, ezért

$$p_1'(t) = -p_2(t)a = a(t - T),$$

amiből

$$p_1(t) = p_1(T) + \int_T^t p_1'(\tau)d\tau = 0 + \int_T^t a(t - T)d\tau = \frac{a}{2}(T - t)^2.$$

Ebből meghatározható t_0 értéke: $p_1(t) = p_2(t)$, ha $T - t = \frac{a}{2}(T - t)^2$, azaz $t = T - \frac{2}{a}$, ezek szerint

$$t_0 = \begin{cases} T - \frac{2}{a} & \text{ha } T > \frac{2}{a} \\ 0 & \text{ha } T \leq \frac{2}{a} \end{cases}.$$

Nyilván $p_1(t_0) = p_2(t_0) = T - t_0 = T - (T - \frac{2}{a}) = \frac{2}{a}$.

Tegyük fel. hogy $t_0 > 0$, azaz $T > \frac{2}{a}$, és számoljuk ki $p_1(t)$-t a $[0, t_0)$ intervallumon. Legyen $t \in [0, t_0)$ tetszőleges. A (11.3) kifejezésből a (11.2) összefüggést felhasználva az adódik, hogy

$$p_1'(t) = \begin{cases} -ap_1(t) & \text{ha } p_1(t) > p_2(t) \\ -ap_2(t) & \text{ha } p_1(t) \leq p_2(t) \end{cases}. \tag{11.5}$$

Ha $p_1(t) \leq p_2(t)$, akkor

$$p_1'(t) = -ap_2(t) = a(t - T),$$

ha $p_1(t) > p_2(t)$, akkor $-ap_1(t) < -ap_2(t)$, így

$$p_1'(t) = -ap_1(t) < -ap_2(t) = a(t - T).$$

Emiatt mindkét esetben

$$p_1'(t) \leq a(t - T) < a(t_0 - T) = a(T - \frac{2}{a} - T) = -2 < p_2'(t),$$

továbbá $p_1(t_0) = p_2(t_0)$, ezek szerint $p_1(t) > p_2(t)$. Összefoglalva ez azt jelenti, hogy

$$\forall\ t \in [0, t_0) \quad \text{esetén} \quad p_1(t) > p_2(t). \tag{11.6}$$

Eszerint a (11.5)-böl következik, hogy minden $t \in [0, t_0)$ esetén $p_1'(t) = -ap_1(t)$, felhasználva, hogy $p_1(t_0) = \frac{2}{a}$ adódik hogy

$$p_1(t) = \frac{2}{a} e^{-a(t-T+\frac{2}{a})}.$$

A (11.4) és (11.6) alapján

$$i(t) = \begin{cases} 0 & \text{ha } t \in [0, t_0) \\ 1 & \text{ha } t \in (t_0, T] \end{cases},$$

továbbá ezért

$$y_1(t) = \begin{cases} y_1^0 & \text{ha } t \in [0, t_0) \\ y_1^0 e^{a(t-t_0)} & \text{ha } t \in (t_0, T] \end{cases},$$

valamint

$$\begin{aligned} y_2(t) &= \begin{cases} y_2(0) + \int_0^t ay_1(\tau)\, d\tau & \text{ha } t \in [0, t_0) \\ y_2(0) + \int_0^{t_0} ay_1(\tau)\, d\tau & \text{ha } t \in (t_0, T] \end{cases} \\ &= \begin{cases} y_2^0 + ay_1^0 t & \text{ha } t \in [0, t_0) \\ y_2^0 + ay_1^0 t_0 & \text{ha } t \in (t_0, T] \end{cases}. \end{aligned}$$

A fenti $((y_1, y_2), i)$ függvénypár nem biztos, hogy valóban megoldása is a feladatnak. Mivel az $((y_1, y_2), i) \mapsto H(t, (y_1, y_2), i, (p_1, p_2))$ Hamilton-függvény nem konkáv (az iy_1 szorzat miatt), ezért a Mangasarian-tétel nem alkalmazható ennek az eldöntésére. Azonban az Arrow-tétel alapján látható, hogy valóban megoldás. A Hamilton-függvény értékfüggvénye

$$\begin{aligned} H^\vee(t, (y_1, y_2), (p_1, p_2)) &= \max_{i \in [0,1]} H(t, (y_1, y_2), i, (p_1, p_2)) \\ &= \begin{cases} y_2 + ap_2 y_1 & \text{ha } p_1 \le p_2 \\ y_2 + ap_1 y_1 & \text{ha } p_1 > p_2 \end{cases}, \end{aligned}$$

ezért az $(y_1, y_2) \mapsto H^\vee(t, (y_1, y_2), (p_1, p_2))$ függvény lineáris, így konkáv.

11.2. Egy környezetgazdaságtani probléma

11.2 Példa. (Fogyasztás, olajkitermelés, környezetszennyezés) Tekintsünk egy kétszektoros gazdaságot a $[0, T]$ tervezési időszakban. Az első szektor fogyasztási javakat termel, a második szektor a termelés okozta környezetszennyezést méri.

A gazdaság a $t \in [0,T]$ időpontban a fogyasztási javakból $x_1(t)$, mennyiséget termel, ami kezdetben x_1^0, a környezetszennyezés mértéke pedig $x_2(t)$, ami kezdetben x_2^0. A környezetszennyezés változása arányos a termeléssel, azaz

$$x_2'(t) = bx_1(t),$$

ahol b pozitív állandó.

A fogyasztási javak termelésének a növekedése a gazdaság olajvagyonának az időszakban rendelkezésre álló u_0 mennyiségének a kitermelésével lehet elérni. Ennek $t \in [0,T]$ időpontbeli értéke $u(t) \in [0, u_0]$, a fentiek szerint pedig

$$x_1'(t) = u(t).$$

Az $u(t)$ kitermelt olaj mennyiségét úgy kell megválasztani, hogy a $[0,T]$ időszakban termelt fogyasztási javak mennyisége minél több legyen, ez minél kevesebb környezetszennyezést okozzon, és ehhez minél kevesebb olajat használjanak fel, azaz maximalizáljuk a következő integrált:

$$\int_0^T x_1(t) - cx_2(t) + u_0 - u(t)\ dt,$$

ahol b pozitív állandó.

A fentiek a következő optimális irányítási feladatban foglalhatók össze:

$$\left\{ \begin{array}{l} \int_0^T x_1(t) - cx_2(t) + u_0 - u(t)\ dt \to \max \\ \left\{ \begin{array}{l} x_1'(t) = u(t) \quad \text{m.m. } t \in [0,T] \\ x_2'(t) = bx_1(t) \quad \text{m.m. } t \in [0,T] \end{array} \right. \\ u(t) \in [0, u_0] \\ \left\{ \begin{array}{l} x_1(0) = x_1^0,\ x_1(T) \text{ szabad} \\ x_2(0) = x_2^0,\ x_2(T) \text{ szabad} \end{array} \right. \end{array} \right. , \tag{11.7}$$

ahol az $(x_1, x_2) \in W^2_{\mathbb{R}^2}[0,T]$ termelés-környezetszennyezés vektor az állapotváltozó, az $u \in L^2[0,T]$ olajkitermelés az irányítási változó, a $[0, u_0]$ zárt intervallum pedig az irányítási tartomány.

A feladat megoldása:

A feladat Hamilton-függvénye:

$$\begin{aligned} H(t,(x_1,x_2),u,(p_1,p_2)) &= x_1 - cx_2 + u_0 - u + \langle (p_1,p_2),(u,bx_1)\rangle \\ &= (1+bp_2)x_1 - cx_2 + (-1+p_1)u + u_0. \end{aligned}$$

A maximumelv szerint, ha az $((x_1,x_2),u)$ függvénypár megoldása a feladatnak, akkor van olyan (p_1,p_2) függvény, amelyre teljesülnek a következők:

(1) Maximumfeltétel:

$$\begin{aligned} \max_{u\in[0,u_0]} &\quad (1+bp_2(t))x_1(t) - cx_2(t) + (-1+p_1(t))u + u_0 \\ = &\quad (1+bp_2(t))x_1(t) - cx_2(t) + (-1+p_1(t))u(t) + u_0. \end{aligned}$$

(2) Adjungált egyenlet:

$$\begin{bmatrix} p_1'(t) \\ p_2'(t) \end{bmatrix} = -\begin{bmatrix} 1 + bp_2(t) \\ -c \end{bmatrix}$$

(3) Transzverzalitási feltétel: $p_1(T) = 0, \quad$ és $\quad p_2(T) = 0.$

Az (1) a maximumfeltétel szerint

$$u(t) = \begin{cases} 0 & \text{ha } p_1(t) < 1 \\ u_0 & \text{ha } p_1(t) > 1 \end{cases}.$$

A (2) adjungált egyenlet szerint $p_2'(t) = c$, ezért

$$p_2(t) = p_2(T) + \int_T^t p_2'(\tau) d\tau = c(t - T),$$

továbbá $p_1'(t) = -1 - bp_2(t)$, ezért

$$\begin{aligned} p_1(t) &= p_1(T) + \int_T^t p_1'(\tau) d\tau \\ &= 0 + \int_T^t -1 - bc(t - T) d\tau \\ &= -\frac{bc}{2} t^2 + (bcT - 1)t + T - \frac{bc}{2} T^2, \end{aligned}$$

Ebből eldönthető, hogy mikor teljesül, hogy $p_1(t) \geq 1$, illetve $p_1(t) < 1$. A $p_1(t) - 1 = 0$ másodfokú egyenlet gyökei

$$t_{1,2} = T - \frac{1 \pm \sqrt{1 - 2bc}}{bc}.$$

Ezek szerint

$$u(t) = \begin{cases} 0 & \text{ha } t \in [0, t_1) \cup (t_2, T] \\ u_0 & \text{ha } t \in [t_1, t_2] \end{cases}.$$

Ennek az alapján az állapot változók kiszámolhatók a rájuk vonatkozó differenciálegyenletekből:

$$\begin{aligned} x_1(t) &= x_1(0) + \int_0^t x_1'(\tau)\, d\tau = x_1^0 + \int_0^t u(\tau)\, d\tau \\ &= \begin{cases} x_1^0 & \text{ha } t \in [0, t_1) \\ x_1^0 + u_0(t - t_1) & \text{ha } t \in [t_1, t_2] \\ x_1^0 + u_0(t_2 - t_1) & \text{ha } t \in (t_2, T] \end{cases}, \end{aligned}$$

valamint

$$\begin{aligned} x_2(t) &= x_2(0) + \int_0^t x_2'(\tau)\, d\tau = x_1^0 + \int_0^t bx_1(\tau)\, d\tau \\ &= \begin{cases} x_2^0 + bx_1^0 t & \text{ha } t \in [0, t_1) \\ x_2^0 + bx_1^0 t_1 + bx_1^0(t - t_1) + \frac{1}{2} bu_0 (t - t_1)^2 & \text{ha } t \in [t_1, t_2] \\ x_2^0 + bx_1^0 t_1 + bx_1^0(t_2 - t_1) + \frac{1}{2} bu_0 (t_2 - t_1)^2 & \\ + \left(x_1^0 + u_0(t_2 - t_1)\right)(t - t_2) & \text{ha } t \in (t_2, T] \end{cases}. \end{aligned}$$

Az $((x_1, x_2), u) \mapsto H(t, (x_1, x_2), u, (p_1, p_2))$ Hamilton-függvény lineáris, így konkáv, ezért a Mangasarian-tétel szerint $((x_1, x_2), u)$ valóban megoldása a feladatnak.

11.3. Gyakorlatok

1. Mutassuk meg, hogy a variációszámítás Lagrange-feladatában az Euler-Lagrange-egyenlet a Pontrjagin-féle maximumelv speciális esete. Útmutatás: tekintsük a Lagrange-feladatot, és vizsgáljuk az
$$x'(t) = u(t)\,, \qquad x(0) = x_0\,, \qquad x(T) = x_T$$
dinamikai rendszer mellett az
$$\int_0^T f(t, x(t), u(t))\,dt \to \min$$
irányítási feladatot. Ellenőrizzük, hogy erre az irányítási feladatra az adjungált egyenlet és a maximumelv éppen az Euler-Lagrange-egyenletet szolgáltatja.

2. Igazoljuk az előző gyakorlat alapján, hogy szabad végpontú feladatok esetében az irányítási feladatra vonatkozó transzverzalitási feltétel éppen a variációs feladatra vonatkozó transzverzalitási feltételt adja.

3. Vizsgáljuk meg, hogy az első gyakorlat alapján a Mangasarian-féle elégséges feltételből következik a Lagrange-feladat integrandusának konvexitása. Ez azt jelenti, hogy a Mangasarian-feltétel mellett a Lagrange-feladat stacionárius függvényei egyúttal a Lagrange-feladat megoldásai is.

Ajánlott irodalom

1. Aoki, M., *Optimal Control and Systems Theory in Dynamic Economic Analysis*, North-Holland, New York, 1980.

2. De la Fuente, A., *Mathematical Methods and Models for Economists*, Cambridge University Press, Cambridge, New York, Melbourne, Madrid, 2000.

3. Kamien, M., Schwartz, L., *Dynamic Optimization and Optimal Control in Economics and Management*, North-Holland, New York, 1981.

4. Kánnai Z., *Analitikus módszerek a pénzügyben és a közgazdaságtanban*, www.tankonytar.hu, 2013.

5. Kósa A., *Variációszámítás*, Tankönyvkiadó, Budapest, 1973.

6. Kósa A. (Szerk.), *Optimumszámítási módszerek*, Műszaki Könyvkiadó, Budapest, 1979.

7. Luenberger, D. G., *Optimization by vector space methods*, Wiley, New York, London, Sidney, Toronto, 1969.

8. Magyarkuti Gy., *Mértékelmélet és dinamikus programozás*, www.tankonyvtar.hu, 2013.

9. Sethi, S., Thompson, G., *Optimal Control Theory, Applications to Management Science and Economics*, Springer, Berlin, Heidelberg, New York, 2006.

10. Sydsæter, K., Hammond, P., Seierstad, A., *Further Mathematics for Economic Analysis*, Prentice Hall, New York, 2005.

11. Tallos P., *Dinamikai rendszerek alapjai*, Aula, Budapest, 1999.