СОФИЙСКИ УНИВЕРСИТЕТ „СВ. КЛИМЕНТ ОХРИДСКИ“
ФАКУЛТЕТ ПО МАТЕМАТИКА И ИНФОРМАТИКА

Задачи на оптималното управление в условие на неопределеност

Автор:
Боян Колев Стефанов

Научен ръководител:
Проф. д.н. Михаил Иванов Кръстанов

АВТОРЕФЕРАТ

на дисертация за придобиване на образователната и научна степен
Доктор в професионално направление 4.5 Математика,
докторска програма ”Изследване на операциите”

София, 2024

Абстракт

Целта на тази дисертация е да развие предходни резултати отнасящи се до необходими и достатъчни условия за оптималност за задачата на оптималното управление на безкраен времеви хоризонт, както с непрекъснато, така и с дискретно време. Изследването акцентира върху управляеми системи, които се сблъскват с предизвикателството на несигурността, като се обръща внимание на сценарии от реалния свят с ограничени функции на управление.

В контекста на линейно-квадратичен модел е приложена детерминистична рамка за разглеждане на несигурността, което предоставя структуриран метод за подход към нея.

В условията на непрекъснато време, са изведени достатъчни условия за оптималност за линейно-квадратични игри, които са важни за определянето на управленски стратегии в условия на несигурност.

За линейно-квадратични игри с дискретно време, посредством принципът за оптималност на Белман, са получени достатъчни условия за оптималност.

В допълнение за дискретни игри от общ вид са формулирани необходими условия за оптималност от тип принцип на максимума на Понтрягин. Изследван е още и опростен сценарий в дискретно време без смущения, като е формулирано ново достатъчно оптимално условие, базирано на подходящи предположения.

Общият принос на настоящата работа към обширната област на теорията на управлението е разработването на рамка за разглеждане на практически проблеми в реални условия. Вниманието е насочено към важните ограничения, свойствени за функциите на управление, с фокус върху справянето с несигурността, често срещана в динамичните системи.

Съдържание

Благодарности

Бих искал да изразя най-дълбоката си благодарност на всички онези, които ме подкрепяха по време на това обучение.

Първо и най-вече, сърдечно благодаря на моя научен ръководител, професор Михаил Кръстанов, за неговите насоки, търпение и критики по време на това изследване. Неговият опит и обмислено наставничество изиграха основна роля в оформянето на този труд.

Изразявам своята благодарност към Факултета по математика и информатика на Софийския университет и колектива на катетдра ВОИС за предоставената академичната среда за моето изследване.

Искрено благодаря на Росен Розенов за отделеното време и значителния принос, който внесе в това изследователско усилие.

Специални благодарности към проф. Надя Златева и проф. Надежда Рибарска както и към моите колеги докторанти за приятната работна атмосфера, стимулиращите дискусии и за всички приятни моменти, които споделихме през последните няколко години.

Задължен съм на всички от моето семейството за тяхната обич и непоколебимата вяра в мен. Тяхната всеотдайност и насърчение бяха моят постоянен източник на воля и мотивация. Това постижение е колкото мое, толкова и тяхно.

И накрая, бих искал да изразя признателността си към всички автори и изследователи, чиято работа и приноси са цитирани в този проект.

Това пътешествие беше трансформиращо преживяване и приносът на всеки един от вас остави незаличим отпечатък върху личното ми израстване. Благодаря ви!

Боян Стефанов

Въведение

Теорията на оптималното управление е важна математическа област, която цели да намери най-ефикасната стратегия за управление на динамична система във времето с цел максимизиране или минимизиране на конкретна целева функция.Тя има приложения в различни области, включително икономика, инженерство, медицина, информационните системи, науки за околната среда и управление и други. Основните понятия в теорията на оптималното управление обхващат: *променливи на състоянието*, които отразяват текущата ситуация на системата; *променливи на управлението*, които представляват параметри за влияние или управление на системата; *функции на разходите*, които измерват ефективността на различните управленски стратегии; *диференциални или диференчни уравнения*, използвани за описване на динамиката на системите и *оптимизация*, процесът на идентифициране на най-добрата възможна стратегия за управление в рамките на поставените ограничения и условия. Тези понятия са основополагащи за разбирането и прилагането на оптималното управление в практиката.

Корените на теорията на оптималното управление могат да бъдат проследени до вариационното смятане, област на математиката, датираща от ерата на Нютон и Лайбниц през 17 век. Тя се фокусира върху определянето на екстремуми (максимални или минимални стойности) на функционали, които представляват изображения от пространството на функциите към множеството на реалните числа. Счита се, че то води началото си от 1696 г. с работата на Йохан Бернули и задачата за брахистохроната, една от най-ранните задачи във вариационното смятане. Задачата се състои в намирането на кривата на най-бързото спускане, пътя, по който една частица ще се движи под силата на гравитацията от една точка до друга за най-кратко време. Решението на задачата с брахистохрона е важен момент в историята на математиката, тъй като това е един от първите случаи, когато е намерено математическо решение на оптимизационна задача, свързана с динамичен процес.

Задачата на вариационно смятане, може да бъде формулирана по следния начин: Като се има предвид функционала

$$J(x) = \int_a^b L(x(t), \dot{x}(t), t)\, dt,$$

където $x(t)$ е функция, която изобразява $[a, b] \subset \mathbb{R}$ в $\mathbb{R}$, а $\dot{x}(t)$ е производната на x по отношение на t. Задачата е да се намери функцията $x(t)$, при която се достига минимум или максимум на функционала $J(x)$.

В следващите години вариационното смятане е значително повлияно от математици като Ойлер и Лагранж. Един от приносите на Ойлер е извеждането на уравнението на Ойлер-Лагранж, което е основно уравнение за намиране на функция, която минимизира или максимизира функционал. Уравнението се изразява като:

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = 0.$$

Лагранж въвежда концепцията за обобщени координати и принципа на най-малкото действие, който гласи, че пътят, изминат от физическа система между две състояния е този, за който функционалът ”действие” се минимизира.

През 18-ти и 19-ти век ориентацията е до голяма степен върху разработването на математически техники за оптимизиране на функционали, поставяйки началото на формалното установяване на теорията на оптималното управление през 20-ти век. Тази рамка се разширява за да обхване функционали подложени на ограничения, което увеличава спектъра на проблемите, които могат да бъдат адресирани в контекста на оптимизацията. Задачата на оптималнато управление има следната формулировка:

$$\min J(u(t)) = \int_{t_0}^{t_f} L(x(t), u(t), t)\, dt$$
$$\text{Предмет на: } \begin{cases} \dot{x}(t) = f(x(t), u(t), t) \quad \text{(динамика)} \\ u(t) \in U \quad \text{(ограничения на управлението)} \\ x(t_0) = x_0 \quad \text{(начално условие)} \\ x(t_f) = x_f \quad \text{(терминално условие (не задължително))}, \end{cases} \tag{P}$$

където $x(t)$ представлява състоянието на системата в момент t, $u(t)$ представлява управлението в момент t, U е затворено множество от всички допустими управления, $f(x, u, t)$ е функцията определяща динамиката на системата, а $L(x, u, t)$ е функцията на разходите. Целта е да се намери функция на управление $u^*(t)$, която минимизира $J(u)$ в интервала $[t_0, t_f]$, придържайки се към определената динамика и ограничения.

Принципът на максимума на Понтрягин, разработен от руския математик Лев Понтрягин и неговите сътрудници в края на 50-те години на миналия век, представлява важен пробив в теорията на управлението (виж [45]). Той осигурява необходими условия за оптималност чрез въвеждане на спомагателна функция наречена *Хамилтониан* (функция на Хамилтон), която комбинира динамиката на системата, целевата функция и така наречената *спрегната* променлива. Формално принципът гласи:

Ако $u^*(t)$ е оптимална функция на управление в интервала $[t_0, t_f]$ за задачата на оптималното управление (P), със съответстваща оптимална траектория $x^*(t)$, тогава съществува нетривиална, абсолютно непрекъсната функция $\psi(t)$, наречена спрегната променлива, така че:

$\psi(t)$ удовлетворява спрегнатото уравнение:

$$\dot{\psi}(t) = -\frac{\partial \mathcal{H}}{\partial x}(x^*(t), u^*(t), \psi(t), T),$$

където $\mathcal{H}$ е Хамилтониана, дефиниран като:

$$\mathcal{H}(x, u, \psi, t) = L(x, u, t) + \psi \cdot f(x, u, t)$$

За почти всяко t, Хамилтониана е максимизиран (или минимизиран за задачи за минимизиране) по отношение на u по оптималната траектория:

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

където U е множеството от всички допустими управления.

Спрегнатата променлива трябва да отговаря на определени гранични условия, известни като *условия за трансверзалност*, които зависят от конкретната задача.

Принципът може да бъде формулиран както за непрекъснати, така и за дискретни управляеми системи.

По същото време Ричард Белман разработва принципа за оптималност на Белман, който представлява рекурсивна стратегия за оптимизация (виж [8]). За динамична система със състояние $x(t)$ във време t и управление $u(t)$ за времеви хоризонт $[t_0, T]$, *функция цена* $V^*(x(t), t)$ се дефинира като:

$$V^*(x(t), t) = \min_{u(\tau)} \left\{ \int_t^T g(x(\tau), u(\tau), \tau) d\tau + h(x(T)) \right\},$$

където g е функцията на текущите разходи, h е функцията на терминалните разходи и τ варира от t до T. Принципът гласи, че ако $u^*(t)$ е оптимална стратегия за управление от t_0 до T, тогава за всяко междинно време t_1, $u^*(t)$ от t_1 до T също трябва да бъде оптимално за системата, стартираща в състояние $x(t_1)$. Това означава, че решението по всяко време t зависи само от състоянието в този момент, а не от предходния път, предприет за достигане на това състояние.

Този метод разделя сложните задачи на дългосрочното планиране на серия от подзадачи в различни времеви моменти, които могат да бъдат решени независимо и след това комбинирани за постигане на окончателното решение на първоначалната задача. Подходът е изключително ефективен при системи с дискретно време, като води до рекурсивна алгоритмична структура, често описвана като "динамично програмиране".

Несигурността е основен аспект на различни явления от реалния свят, оформяйки и влияейки върху множество процеси. Тя присъства в природните процеси, инженерството, производството, икономиката, здравеопазването, информационните системи, социалните и поведенчески процеси и други. Условията на околната среда, свойствата на материалите, производствените процеси и търсенето на пазари поставят предизвикателства за прогнозиране и управление. Икономическите и финансовите системи по своята същност са несигурни поради пазарни колебания, геополитически събития и непредвидени кризи. Здравеопазването е изправено пред несигурност в биологичните системи, ефективността на лечението и разпространението на болести. Информационните системи и технологии са изправени пред претоварване на мрежи, софтуерни грешки и заплахи за киберсигурността. Социалните и поведенческите процеси се влияят от индивидуалните предпочитания, обществената динамика и непредвидените събития. Политическите, културните и икономическите промени допринасят за сложния и непредвидим характер на социалните системи.

Важността на задачата, при наличие на несигурност се подчертава от съществуването на множество различни математически концепции за нейното решаване. Въпреки, че изследването не се задълбочава в тях, някои примери за такива са:

Теорията на вероятностите е ключова за оптималното управление и вземане на решения (виж [38], [34] и [32]); Стохастичните процеси, изучават динамични системи под произволни влияния ([19], [50]). Стохастичните диференциални уравнения (SDE) са диференциални уравнения, в които един или повече от членовете са стохастичен процес, често представляващ шум или други произволни влияния. SDE се използват за моделиране на динамиката на системите, засегнати от случайни флуктуации, както е изследвано от [59]); В статистиката работата на Фишър [24] и Нойман [41], включително въвеждането на Фишър за оценка на максималната вероятност и лемата на Нойман-Пиърсън, за анализиране на несигурни данни и систематично тестване на хипотези; Размитата (fuzzy) логика, разработена в [57], разширява класическата логика, като включва степени на истинност, подпомагайки вземането на решения при двусмислени или несигурни условия; Информационната теория, въведена от Шанън [49], използва ентропията за измерване на несигурността в комуникационните системи и анализа на данни, като ентропията определя количествено несигурността във вероятностното разпределение; Методите на Монте Карло, създадени от пионери като Николай Метрополис и Станислав Улам, са изчислителни техники за симулиране на случайни процеси, инструмент за оценка на поведението на сложни системи при несигурност; Бен-Тал и Немировски [9] и Берцимас и Сим [11] играят важна роля във формирането и развитието на робастната оптимизация.

Проектирането и прилагането на робастни управления за системи при наличие на смущения има предимството да гарантира приемливи резултати при редица обстоятелства, докато решенията, изцяло базирани на очакваната полезност, обикновено биха довели до по-ниска производителност, ако реализацията на несигурните количества се отклонява значително от очакванията. Освен това, понякога не е възможно да се присвоят вероятности на различните сценарии (Knightian несигурност), което прави подхода на очакваната полезност неприложим. Алтернатива е да се търси минимизиране на загубата, свързана с най-лошия резултат. Робастните методи за управление са намерили широко приложение в инженерството (виж [15], [22], [58], [53]) и промишления сектор (напр. [2], [20] и [44]), но също и към различни икономически проблеми, особено в областта на паричната теория и политика. Примерите включват [46],[42], [27] и [18]. Подходът в [30] към устойчивостта се мотивира от факта, че често вземащите решения работят с модели, които са само приближения на истинския модел, който генерира данните. Това ограничение на ентропията обикновено се преобразува в наказателен член в целевата функция, за да се улови предпочитанието за устойчивост. Интерпретира се като игра против хипотетичен враждебен агент (природа), който избира смущенията така, че да максимизира загубата, която водещият политиката се опитва да минимизира.

Централно място в този контекст заема теорията на игрите, която осигурява рамка за подход към сценарии с множество взаимодействащи участници, чиито решения са взаимозависими. Нейните основи са положени от Джон фон Нойман [40] и Джон Неш [39]. В центъра на анализа на тези игри е концепцията за равновесие, при което стратегиите на играчите достигат такова състояние, че никой от тях няма интерес едностранно да се отклони от избраната от него стратегия. В контекста на безкраен времеви интервал има различни понятия за равновесие, като равновесие на Неш, равновесие на Щакелберг, слабо изпреварващо равновесие и други. Динамичните игри могат да се провеждат в непрекъснато или в дискретно време, както и да бъдат определени върху крайни или безкрайни времеви хоризонти. Те са широко използвани за моделиране на взаимодействия, при които играчите преследват различни цели.

Глава 1 подхожда към въпроса за устойчивостта и преодоляването на неблагоприятни условия през призмата на динамичните игри. Разглеждат се некооперативни линейно-квадратични игри между двама души с непрекъснато време върху безкраен времеви хоризонт. Този подход е в съответствие с рамката предложена в [6], която използва теорията на игрите, за да формализира най-лошия възможен случай. Въпреки, че линейно-квадратичните игри са добре проучени в литературата (напр. [7], [21] и [23]), случаят върху безкраен времеви хоризонт обикновено се представя в контекст, в който не се налагат ограничения върху променливите на управлението. Този недостатък е особено съществен в сценарии, в които управленията по своята същност са обвързани с определени ограничения. Пример за такова ограничение е нулевата долна граница на номиналния лихвен процент. Глава 1 има за цел да преодолее тази празнина. Основният принос се състои в установяването на достатъчно условие за седлова точка (съответства на равновесието на Неш) за линейно-квадратична игра върху безкраен времеви хоризонт, при наличие на ограничения върху управлението на минимизиращия играч. Демонстрирано е съществуването на компактна околност на началото във фазовото пространство, където тези ограничения не са активни. Съществуването на такава околност е подробно описано в [29] за специфична линейно-квадратична задача на оптималното управление. Използвайки я, можем да трансформираме задачата на безкраен времеви хоризонт в задача върху краен интервал, която е разрешима с помощта на числени методи. Въпреки, че резултатите са оформени в контекста на седлова точка, те са еднакво приложими към формулировките на игри на Щакелберг, които често са по-подходящи за определени икономически модели. Тази приложимост произтича от съвпадането на стационарните обратни връзки определящи равновесията на Неш и Щакелберг за игри, характеризиращи се с ортогонални реакционни функции, както е посочено в [47].[1]

Определение 0.1 (Равновесие на Щакелберг [16])**.** *Дадена е игра с двама души, където първия играч иска да минимизира целева функция $J_1(u_1, u_2)$, а втория играч иска да минимизира целева функция $J_2(u_1, u_2)$. Бележим допустимите множества от управления на първия и втория играч съответно с U_1 и U_2. Двойката управления (u_1^*, u_2^*) се нарича равновесие на Щакелберг с втория играч като лидер и първия като последовател, ако за всяко $u_2 \in U_2$ и всяко $u_1 \in U_1$ е изпълнено*

$$J_2(u_1^*, u_2^*) \leq J_2(u_1'(u_2), u_2),$$

където

$$J_1(u_1'(u_2), u_2) = \min_{u_1 \in U_1} J_1(u_1, u_2) \quad \textit{и} \quad u_1^* = u_1'(u_2^*).$$

В Глава 2 вниманието е върху линейно-квадратични игри с дискретно време, отново над безкраен времеви хоризонт с акцент върху сценарии, при които се прилагат ограничения към управлението на агента, търсещ минимизиране на целевата функция. Използването на системи с дискретно време става особено уместно, когато естеството на разглеждания проблем е по своята същност дискретно, съобразно дискретния характер на управленията. Въпреки, че споделя прилики с непрекъснатия си аналог, случаят с дискретно време предоставя предизвикателства, които изискват отделен анализ. Централен за този подход е принципът за оптималност на Белман, който, както е демонстрирано, е приложим за широкия спектър от задачи с ”мин-макс” управление.

Приносите на тази глава включват ”мин-макс” теорема за неограничената линейно-квадратична игра и изследване на асимптотиката на траекториите на управляемата система. Също така е установена еквивалентност между задачата върху безкраен хоризонт и съответна на нея задача върху краен интервал. Показано е отново, че в случаите, когато системата е устойчива, съществува околност на началото във фазовото пространство, където ограниченията наложени върху управлението стават необвързващи. Това прави възможно конвертирането на играта върху безкраен хоризонт във игра върху краен интервал, разрешима чрез подходящи методи. Основният принос в Глава 2 представлява достатъчни условия за оптималност за линейно-квадратична игра с дискретно време.

[1] Диференциалните игри с ортогонални моментни реакционни функции се характеризират с независимостта на първите производни на целевите функциите и диференциалните уравнения по отношение на управлението на даден играч от управлението на другия играч (вижте [47], Определение 2.4).

Въпреки, че някои от резултатите получени в тази глава са известни, доказателствата, които са представени са нови. Практическото приложение на подхода е илюстриран чрез модел, описващ краткосрочната динамика на самолет F-16.

Глава 3 се задълбочава в намирането на условия за оптималност от тип принцип на максимума на Понтрягин за дискретни управляеми системи (разработен първоначално от Болтянски в [12]), повлияни от смущения или несигурност. Разглежданата оптимизационна задача попада в класа задачи от тип "мин-макс".

Примери за автори, допринесли в тази област са Фуденберг и Тирол [25], които представят резултати главно върху равновесията в игрите с дискретно време. Подхода, основан на динамичното програмиране и обратна индукция, се използва за формулиране и решаване на тези задачи. Осбърн и Рубинщайн [43] се задълбочават в равновесията на Неш за игри с дискретно време, като използват най-добрата динамика на реакцията и математически инструменти като теореми за фиксирана точка.

Основен резултат в тази глава е необходимо условие за оптималност, структурирано по линия на принципа на максимума на Понтрягин. Доказателството е получено с помоща на Теорема 2.2 в [3], представляваща локално условие за максимум на Хамилтониана. Допълнително, Глава 3 се занимава със специфично подмножество на игри с дискретно време, по-специално задачи на оптималното управление без смущения.

Основният резултат във този раздел, Теорема 3.3, осигурява при подходящи предположения ново достатъчно условие за оптималност. Доказателството използва установена връзка между целевата функция на задачатата и съответната функция на Хамилтон, както е обяснено в Твърдение 3.1. Тази взаимовръзка е изведена чрез използване на дефиниция в [3], където спрегнатата променлива – критична за условието за максимум на Хамилтониана в принципа на Понтрягин – е явно дефинирана за всеки даден оптимален процес.

1 Непрекъсната линейно-квадратична игра

1.1 Постановка на задачата

Нека разгледаме конкретен сценарий, при който е фиксирано реално число $t_0 \geq 0$ и начален вектор $x_0 \in \mathbb{R}^n$. Разглеждаме класа от непрекъснати некооперативни линейно-квадратични диференциални игри. Динамиката на играта се задава от следното линейно диференциално уравнение:

$$\dot{x}(t) = Ax(t) + B_u u(t) + B_v v(t), \quad x(t_0) = x_0, \tag{1}$$

където A, B_u и B_v представляват матрици с размерности съответно $n \times n$, $n \times m_u$ и $n \times m_v$, а n, m_u и m_v са размерностите на съответните векторни пространства. Векторите $x(t)$, където $t \in [t_0, +\infty)$ и x_0 означават съответно състоянието на системата в момент t и началното състояние в момента t_0.

Играчите определят своите действия (или управления) чрез функциите **u** и **v**, съответно за първия и втория играч. Управленията могат да бъдат отворени управление (open-loop) или обратни връзки (closed-loop).

Отворени управление **u** на първия играч е измерима функция $u : [t_0, +\infty) \to \mathbf{U} \subset \mathbb{R}^{m_u}$ с $\int_{t_0}^{\infty} \|u(t)\|^2 \, dt < \infty$. Множеството **U** представлява затворена и изпъкнала околност на началото в $\mathbb{R}^{m_u}$. Множеството от всички отворени управления на първия играч се означава с $\mathcal{U}_{\mathbf{U}}$. В неограничения случай ($\mathbf{U} = \mathbb{R}^{m_u}$) означаваме това множество с $\mathcal{U}$. Обратната връзка $u : \mathbb{R}^n \to \mathbb{R}^{m_u}$ е функция $u(x)$, когато тя е линейна, $u(x) = K_u x$, където K_u е $m_u \times n$ матрица.

Предполага се, че вторият играч, представляващ смущенията, не е изправен пред ограничения за своето управление. Отворено управление **v** на втория играч е измерима функция $v : [t_0, +\infty) \to \mathbb{R}^{m_v}$ с $\int_{t_0}^{\infty} \|v(t)\|^2 \, dt < \infty$. Множеството от всички такива управления означаваме с $\mathcal{V}$. Обратната връзка $v : \mathbb{R}^n \to \mathbb{R}^{m_v}$ е функция $v(x)$, когато тя е линейна, $v(x) = K_v x$, където K_v е $m_v \times n$ матрица.

Предвид динамиката (1), разглеждаме следната линейно-квадратична игра върху безкраен хоризонт:

$$\inf_{\mathbf{u} \in \mathcal{U}_{\mathbf{U}}} \sup_{\mathbf{v} \in \mathcal{V}} J(x_0, t_0, \mathbf{u}, \mathbf{v}), \tag{2}$$

където целевата функция се дефинира като:

$$J(x_0, t_0, \mathbf{u}, \mathbf{v}) := \int_{t_0}^{\infty} \left(x_{\mathbf{u},\mathbf{v}}^{\intercal}(t) Q x_{\mathbf{u},\mathbf{v}}(t) + u^{\intercal}(t) R_u u(t) - v^{\intercal}(t) R_v v(t) \right) dt.$$

В тази формулировка Q е симетрична, положително полу-определена матрица с размерност $n \times n$, докато R_u и R_v са симетрични, положително определени матрици, съответно с размерности $m_u \times m_u$ и $m_v \times m_v$. Целта на първия играч е да минимизира целевата функция $J(x_0, t_0, \mathbf{u}, \mathbf{v})$, докато вторият играч се стреми да я максимизира. Означението $\mathbf{x}_{\mathbf{u},\mathbf{v}}(\cdot)$ бележи *фазовата траектория*, съответстваща на функциите на управление **u** и **v**. Функцията *цена* $V_{\mathbf{U}} : \mathbb{R}^n \times [t_0, \infty) \to \mathbb{R}$ на играта се определя като:

$$V_{\mathbf{U}}(x_0, t_0) := \inf_{\mathbf{u} \in \mathcal{U}_{\mathbf{U}}} \sup_{\mathbf{v} \in \mathcal{V}} J(x_0, t_0, \mathbf{u}, \mathbf{v}).$$

Допустимата двойката управления $(\mathbf{u}, \mathbf{v})$ се счита за *осъществима*, ако целевата функция J, както е определена в (2), приема крайна стойност за нея. Полагаме

$$V(x_0, t_0) := V_{\mathbb{R}^{m_u}}(x_0, t_0).$$

Следното предположение се прилага за останалата част от тази глава. То поставя основата за анализиране и решаване на диференциалната игра, като гарантира, че системата, описана от линейно-квадратичната рамка, е устойчива.

Постоянно предположение (*)

- *Матричното алгебрично уравнение на Рикати (3)*

$$Q + XB_vR_v^{-1}B_v^{\intercal}X - XB_uR_u^{-1}B_u^{\intercal}X + A^{\intercal}X + XA = 0, \tag{3}$$

  относно $n \times n$ матрица X, притежава симетрично положително определено решение, означено с P.

- *Всички собствени стойности на матриците*

$$A,\ [A - B_uK_u],\ [A + B_vK_v] \text{ и } [A - B_uK_u + B_vK_v]$$

  имат отрицателни реални части, където

$$K_u := R_u^{-1}B_u^{\intercal}P \quad \text{и} \quad K_v := R_v^{-1}B_v^{\intercal}P.$$

По отношение на информационния модел предполагаме, че и двамата играчи имат пълно познание за всички параметри, влияещи на динамиката, матриците A, B_u и B_v, както и тези на целевата функция, матриците Q, R_u и R_v. Освен това приемаме, че през цялата продължителност на играта и двамата играчи имат достъп до текущото състояние на системата, означено като $x(t)$ за $t \in [t_0, \infty)$. Подобен модел се разглежда в [23].

1.2 Приближение на ограничената игра

За да се справим със сценария, включващ ограничения върху управлението идеята е да се установи връзка между задачата на безкраен времеви хоризонт и нейн еквивалент върху краен интервал. Този подход се оказва удачен, когато системата е устойчива. В такива случаи се демонстрира, че съществува околност на началото във фазовото пространството, където ограниченията върху управленията стават неактивни. Когато фазовата траектория принадлежи на тази околност, решението на ограничената задача съвпада с решението на неограничената. Следователно, това редуцира разглежданата задачата до задача върху краен интервал.

Твърдението по-долу установява, че за произволна двойка L^2-управления $(\mathbf{u}, \mathbf{v})$, фазовата траектория клони към началото, с течение на времето. Тук $L^2([t_0, \infty), \mathbb{R}^m)$ е определено като множеството $\{w : [t_0, \infty) \to \mathbb{R}^m : \int_{t_0}^{\infty} |w(t)|^2 dt < \infty\}$, което образува банахово пространство с норма $\|\mathbf{w}\|_{L^2} = \big(\int_{t_0}^{\infty} |w(t)|^2 dt\big)^{1/2}$.

Твърдение 1.1 (Сходяща фазова траектория)**.** *Нека Постоянното предположение (*) е изпълнено, нека $\mathbf{u} \in L^2([t_0, \infty), \mathbb{R}^{m_u})$ и $\mathbf{v} \in L^2([t_0, \infty), \mathbb{R}^{m_v})$ са произволни измерими функции и нека $\mathbf{x}_{\mathbf{u},\mathbf{v}}$ е съответната фазова траектория. Тогава,*

$$\lim_{t \to +\infty} x_{\mathbf{u},\mathbf{v}}(t) = \mathbf{0}.$$

1.2.1 Решение на играта без ограничения

Първоначално нека разгледаме сценария, при който управлението на първия играч е неограничено. При подходящи технически предположения последващото твърдение предоставя обратни връзки, които определят *равновесие на Неш* за диференциалната игра (2). С други думи, профилът на стратегиите се счита за оптимален, ако като се имат предвид стратегиите избрани от играчите, нито един отделен играч не може да подобри собственият си резултат чрез независима промяна на стратегията си.

Твърдение 1.2 (Решение на неограничената игра). *Нека Постоянното предположение (*) е изпълнено и нека дефинираме обратните връзки*

$$\bar{u}(x) := -K_u x \in \mathbb{R}^{m_u},\ \bar{v}(x) := K_v x \in \mathbb{R}^{m_v},\ x \in \mathbb{R}^n. \tag{4}$$

Тогава,

$$J(x_0, t_0, \bar{\mathbf{u}}, \mathbf{v}) \leq V(x_0, t_0) = J(x_0, t_0, \bar{\mathbf{u}}, \bar{\mathbf{v}}) \leq J(x_0, t_0, \mathbf{u}, \bar{\mathbf{v}}) \tag{5}$$

за всяко $\mathbf{u} \in \mathcal{U}$ и всяко $\mathbf{v} \in \mathcal{V}$, т.е. $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ е седлова точка (равновесие на Неш) за диференциална игра (2) в неограничения случай, когато $\mathbf{U} = \mathbb{R}^{m_u}$ и

$$J(x_0, t_0, \bar{\mathbf{u}}, \bar{\mathbf{v}}) = x_0^\intercal P x_0.$$

1.2.2 δ-околност с неактивни ограничения

Продължавайки с нашия анализ, нека сега проучим сценария, при който се налагат ограничения върху функцията на управление на минимизиращия играч. Въвеждаме следните означения: $\bar{\mathbf{B}}^n$ обозначава затвореното единично кълбо с център началото на $\mathbb{R}^n$, а $\bar{x}_{\bar{\mathbf{u}},\bar{\mathbf{v}}}(\cdot, y, \tau)$ траекторията, породена от (1), съответстваща на управленията $\bar{\mathbf{u}}$ и $\bar{\mathbf{v}}$ и започвайки от точката y в момента τ ($\bar{\mathbf{u}}$ и $\bar{\mathbf{v}}$ се определят от (4)). Следващото твърдение установява съществуването на δ-околност на началото в $\mathbb{R}^n$, където наложените ограничения върху управлението на първия играч са удовлетворени във всичките ѝ точки. Освен това е възможно тази околност да се избере по такъв начин, че всяка траектория, започваща от точка в нея, не само да не поражда нарушаване на ограниченията наложени на управлението ($-K_u x \notin \mathbf{U}$), но също така и да клони към началото с течение на времето.

Твърдение 1.3 (δ-околност). *Нека Постоянното предположение (*) е изпълнено и нека матрицата*

$$\bar{Q} := Q + PB_u K_u - PB_v K_v$$

е положително определена. Тогава съществуват положителни реални числа δ_0 и δ удовлетворяващи неравенството $\delta_0 > \delta$, такива че:

(i) включването $-K_u x \in \mathbf{U}$ е изпълнено за всяка точка $x \in \delta_0 \bar{\mathbf{B}}^n$;

(ii) за всяка точка $y \in \delta \bar{\mathbf{B}}^n$ и за всяко $\tau \geq 0$, включването $\bar{x}_{\bar{\mathbf{u}},\bar{\mathbf{v}}}(t, y, \tau) \in \delta_0 \bar{\mathbf{B}}^n$ е изпълнено за всяко $t \geq \tau$ и траекторията $\bar{x}_{\bar{\mathbf{u}},\bar{\mathbf{v}}}(t, y, \tau)$ клони към началото при $t \to +\infty$.

Бележка 1.1. *За остналата част на тази глава ще фиксираме параметъра δ, въведен в горното твърдение, дефиниращ δ-околността.*

Бележка 1.2. *Твърдението 1.1 налага съществуването на момент във времето $T \geq t_0$, когато фазовата траектория $x_{\mathbf{u},\mathbf{v}}(T) \in \delta\bar{\mathbf{B}}^n$, съответстваща на произволна допустима двойка управления $(\mathbf{u}, \mathbf{v})$, навлиза във δ-околността.*

Нека фиксираме реално число $T > 0$ и разгледаме следната диференциална игра на краен времеви интервал $[t_0, T]$

$$V_{\mathrm{U},T}(x_0, t_0) := \inf_{\mathbf{u}\in\mathcal{U}_{\mathrm{U},T}} \sup_{\mathbf{v}\in\mathcal{V}_T} J_T(x_0, t_0, \mathbf{u}, \mathbf{v}), \tag{6}$$

предмет на динамиката (1), където

$$\begin{aligned} J_T(x_0, t_0, \mathbf{u}, \mathbf{v}) :=& x^{\intercal}_{\mathbf{u},\mathbf{v}}(T) P x_{\mathbf{u},\mathbf{v}}(T) \\ &+ \int_{t_0}^{T} \left(x^{\intercal}_{\mathbf{u},\mathbf{v}}(t) Q x_{\mathbf{u},\mathbf{v}}(t) + u^{\intercal}(t) R_u u(t) - v^{\intercal}(t) R_v v(t) \right) dt. \end{aligned}$$

В този контекст ние използваме дефиниции на допустими управления, подобни на тези в случая върху безкраен хоризонт, като въвеждаме индекса ”T”, за да разграничим двата случая. Следователно, използваме означението $\mathcal{U}_{\mathrm{U},T}$, за да представим множеството от отворени управления на първия играч и $\mathcal{V}_T$ за множеството от отворени управления на втория играч.

Бележка 1.3. *Налични са достатъчни условия, които гарантират съществуването на равновесие на Неш за диференциалната игра на краен хоризонт (6) (например [31] и [54]).*

1.3 Достатъчни условия за оптималност

Последното твърдение в тази глава предоставя достатъчни условия за решението на диференциалната игра (2), когато двамата играчи не избират задължително своите оптимални стратегии едновременно. Такъв е например случаят с игрите на Щакелберг, където един от играчите има предимството да избере първи оптималната си стратегия (лидерът), а другият играч (последователят) приема тази стратегия като дадена в стремежа си да минимизира загубата си (виж [21]). В някои приложения Щакелберг игрите са удобно представяне на ситуации, при които, вместо да се изправи пред интелигентен противник, минимизиращият агент играе срещу злонамерена природа, която действа като лидер и избира смущения така че да увеличите максимално загубата на последователя. Решения на такива игри водят до оптимални стратегии, които са устойчиви на обща несигурност.

Твърдение 1.4 (Решение на играта с ограничения)**.** *Нека Постоянното предположение (*) е изпълнено и нека $(\hat{\mathbf{u}}, \hat{\mathbf{v}}) \in \mathcal{U}_{\mathrm{U},T} \times \mathcal{V}_T$ е решение на диференциалната игра върху краен времеви интервал (6), т.е.*

$$J_T(x_0, t_0, \hat{\mathbf{u}}, \hat{\mathbf{v}}) = \inf_{\mathbf{u}\in\mathcal{U}_{\mathrm{U}}} \sup_{\mathbf{v}\in\mathcal{V}} J_T(x_0, t_0, \mathbf{u}, \mathbf{v}).$$

Предполагаме, че $\|x_{\hat{\mathbf{u}},\hat{\mathbf{v}}}(T)\| \leq \delta$ (където δ е определено в Твърдение 1.3) и дефинираме $(\hat{\mathbf{u}}_\infty, \hat{\mathbf{v}}_\infty)$ както следва:

$$(\hat{\mathbf{u}}_\infty, \hat{\mathbf{v}}_\infty) := \begin{cases} (\hat{\mathbf{u}}, \hat{\mathbf{v}}) & \text{за } t \in [t_0, T], \\ (\bar{\mathbf{u}}, \bar{\mathbf{v}}) & \text{за } t \in (T, +\infty) \end{cases}$$

($\bar{\mathbf{u}}$ и $\bar{\mathbf{v}}$ са обратните връзки, дефинирани от (4)). Тогава $(\hat{\mathbf{u}}_\infty, \hat{\mathbf{v}}_\infty)$ решава диференциалната игра върху безкраен хоризонт (2), т.е.

$$J(x_0, t_0, \hat{\mathbf{u}}_\infty, \hat{\mathbf{v}}_\infty) = \inf_{\mathbf{u}\in\mathcal{U}_{\mathrm{U}}} \sup_{\mathbf{v}\in\mathcal{V}} J(x_0, t_0, \mathbf{u}, \mathbf{v}).$$

Бележка 1.4. *В този момент възниква интригуващ въпрос. Да предположим, че $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ е решение на диференциалната игра върху краен времеви интервал при предположенията на Твърдение 1.4. Какво би било естеството на $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ в интервала $[0, T]$? Нашето наблюдение предполага, че съществува момент във времето $\hat{T} \in [0, T]$, такъв че $\hat{u}(t)$ принадлежи на границата на множеството $\mathbf{U}$ за всяко $t \in [0, \hat{T}]$. Въпреки това, изричното определяне на това остава отворен въпрос за нас.*

Ако диференциалната игра върху краен времеви интервал (6) притежава седлова точка (равновесие на Неш), тогава следствието по-долу ни позволява да "удължим" това равновесие за диференциалната игра на безкраен времеви хоризонт (2).

Следствие 1.1 (Равновесие на Неш)**.** *Нека Постоянното предположение (*) е изпълнено и нека $(\hat{\mathbf{u}}, \hat{\mathbf{v}}) \in \mathcal{U}_{\mathrm{U},T} \times \mathcal{V}_T$ е седлова точка за диференциалната игра върху краен интервал (6), т.е.*

$$J_T(x_0, t_0, \hat{\mathbf{u}}, \mathbf{v}) \leq J_T(x_0, t_0, \hat{\mathbf{u}}, \hat{\mathbf{v}}) \leq J_T(x_0, t_0, \mathbf{u}, \hat{\mathbf{v}})$$

за всички $\mathbf{u} \in \mathcal{U}_{\mathrm{U},T}$ и $\mathbf{v} \in \mathcal{V}_T$. Предполагаме, че $\|x_{\hat{\mathbf{u}},\hat{\mathbf{v}}}(T)\| \leq \delta$ (където δ е определено в Твърдение 1.3) и дефинираме $(\hat{\mathbf{u}}_\infty, \hat{\mathbf{v}}_\infty)$ както следва:

$$(\hat{\mathbf{u}}_\infty, \hat{\mathbf{v}}_\infty) := \begin{cases} (\hat{\mathbf{u}}, \hat{\mathbf{v}}) & \text{за } t \in [t_0, T], \\ (\bar{\mathbf{u}}, \bar{\mathbf{v}}) & \text{за } t \in (T, +\infty) \end{cases}$$

($\bar{\mathbf{u}}$ и $\bar{\mathbf{v}}$ са обратните връзки, дефинирани от (4)). Тогава $(\hat{\mathbf{u}}_\infty, \hat{\mathbf{v}}_\infty)$ осигурява седлова точка (равновесие на Неш) за диференциалната игра върху безкраен хоризонт (2), т.е.

$$J(x_0, t_0, \hat{\mathbf{u}}_\infty, \mathbf{v}) \leq J(x_0, t_0, \hat{\mathbf{u}}_\infty, \hat{\mathbf{v}}_\infty) \leq J(x_0, t_0, \mathbf{u}, \hat{\mathbf{v}}_\infty)$$

за всяко $\mathbf{u} \in \mathcal{U}_{\mathrm{U}}$ и всяко $\mathbf{v} \in \mathcal{V}$.

Бележка 1.5. *Следствие 1.1 всъщност ни позволява да намерим решението на диференциалната игра на безкраен времеви хоризонт, като я трансформира в еквивалентна такава върху краен времеви интервал, която може да бъде адресирана с помощта на подходящи методи.*

2 Дискретна линейно-квадратична игра

2.1 Постановка на задачата

Нека k е произволно неотрицателно цяло число. Означаваме с $\mathbb{N}_k$ множеството от всички неотрицателни цели числа, по-големи или равни на k. Нека фиксираме произволно неотрицателно цяло число $k_0 \in \mathbb{N}_0$ и вектор $x_0 \in \mathbb{R}^n$. Разглеждаме класа от дискретни некооперативни линейно-квадратични игри. Управленията на играчите се определят от техния избор на функции: **u** за първия играч и **v** за втория играч.

За всяко положително цяло число m дефинираме множеството

$$\ell^2(\mathbb{N}_{k_0}, \mathbb{R}^m) := \left\{ (w_{k_0}, w_{k_0+1}, w_{k_0+2}, ...) : \sum_{k=k_0}^{\infty} \|w_k\|^2 < \infty \right\},$$

което формира банахово пространство с норма $\|\mathbf{w}\|_2 = \left(\sum_{k=k_0}^{\infty} \|w_k\|^2 \right)^{1/2}$.

Играчите могат да използват отворени управления или обратни връзки. Отворено управление $\mathbf{u} := \{u_k\}_{k \in \mathbb{N}_{k_0}}$ на първия играч е редица, принадлежаща на пространството $\ell^2(\mathbb{N}_{k_0}, \mathbb{R}^{m_u})$, елементите, на която удовлетворяват включването $u_k \in U_k \subset \mathbb{R}^{m_u}, k \in \mathbb{N}_{k_0}$. Означаваме с $\mathcal{U}_{\mathbf{U}}$ множеството от всички отворени управления на първия играч

$$\mathcal{U}_{\mathbf{U}} := \{\mathbf{u} = (u_{k_0}, u_{k_0+1}, ..., u_{k_0+k}, ...) : u_i \in U_i, i = k_0, k_0+1, ..., k_0+k, ...\},$$

където

$$\mathbf{U} := U_{k_0} \times U_{k_0+1} \times ... \times U_{k_0+k} \times ....$$

В това, което следва, понякога ще се нуждаем от версията на играта върху ограничен хоризонт. Поради тази причина, за някое $\kappa \in \mathbb{N}_{k_0}$, ще означим с $\mathcal{U}_{\mathbf{U}}^{\kappa}$ множеството от допустими управления, започващи в момента k_0 и завършващи в момента κ, т.е.

$$\mathcal{U}_{\mathbf{U}}^{\kappa} := \{\mathbf{u}^{\kappa} = (u_{k_0}, u_{k_0+1}, ..., u_{\kappa}) : \ u_i \in U_i, i = k_0, k_0+1, ..., \kappa\}. \tag{7}$$

Когато разглеждаме по-кратък интервал от време, т.е. $k > k_0$, това ще бъде означено с долен индекс в нотацията, както следва: $\mathcal{U}_{k,\mathbf{U}}^{\kappa}$. Тази конвенция ще се прилага и в случай, когато няма ограничения наложени на управлението ($U_k = \mathbb{R}^{m_u}$), но в този случай индексът "**U**" ще да бъде пропуснат за простота. Така например, ще пишем $\mathcal{U}$ за множеството от допустими (неограничени) управления на безкраен времеви хоризонт, започващ от k_0, и $\mathcal{U}_k^{\kappa}$ за множеството от допустими управления, започващи от k и завършва в момента κ.

Приемаме, че няма ограничения за управлението на втория играч. Отворено управление $\mathbf{v} := \{v_k\}_{k \in \mathcal{N}_{k_0}}$ на втория играч е редица принадлежаща на $\ell^2(\mathbb{N}_{k_0}, \mathbb{R}^{m_v})$, елементите, на която удовлетворяват включването $v_k \in \mathbb{R}^{m_v}, k \in \mathbb{N}_{k_0}$. Означаваме с $\mathcal{V}$ множеството от всички отворени управления на втория играч

$$\mathcal{V} := \{\mathbf{v} = (v_{k_0}, v_{k_0+1}, ..., v_{k_0+k}, ...) : \ v_i \in \mathbb{R}^{m_v}, i = k_0, k_0+1, ..., k_0+\kappa, ...\}.$$

Както по-рано, за всяко $\kappa \in \mathbb{N}_{k_0}$ ще означим с $\mathcal{V}^\kappa$ крайномерната проекция на $\mathcal{V}$

$$\mathcal{V}^\kappa := \{\mathbf{v}^\kappa = (v_{k_0}, v_{k_0+1}, ..., v_k) : \ v_i \in \mathbb{R}^{m_v}, i = k_0, k_0+1, ..., \kappa\}. \tag{8}$$

Множествата $\mathcal{V}_k$ и $\mathcal{V}_k^\kappa$ се дефинират аналогично на тези за първия играч.

Нека $\mathbf{u} \in \mathcal{U}_{\mathbf{U}}$ и $\mathbf{v} \in \mathcal{V}$ са допустими управления съответно на първия и втория играч. *Фазовата траектория*

$$\mathbf{x} = (x_{k_0}, x_{k_0+1}, x_{k_0+2}, ...),$$

съответстваща на тази двойка управления се определя рекурсивно от динамиката на играта, която е описана по следния начин:

$$x_{k+1} = Ax_k + B_u u_k + B_v v_k, \quad x_{k_0} = x_0, \quad k = k_0, k_0+1, ..., \tag{9}$$

където A, B_u и B_v са матрици, съответно от размерности $n \times n$, $n \times m_u$ и $n \times m_v$ и $n \times m_v$, а n, m_u и m_v са размерностите на съответните векторни пространства. Тук x_k, $k \in \mathbb{N}_{k_0}$ и x_0 означават съответно състоянието на системата в момента k и началното състояние в момента k_0.

Обратната връзка

$$\mathbf{u}^c = (u^c_{k_0}, u^c_{k_0+1}, u^c_{k_0+2}, ...)$$

на първия играч е функция от вида $\mathbf{x} \to \mathbf{u}^c(\mathbf{x}) \in \mathbf{U}$, където

$$\mathbf{u}^c(\mathbf{x}) := (u^c_{k_0}(x_{k_0}), u^c_{k_0+1}(x_{k_0+1}), ..., u^c_k(x_k), ...)$$

с $u^c_k(x_k) \in U_k$ за всяко $k = k_0, k_0+1, k_0+2, ...$

Обратната връзка

$$\mathbf{v}^c = (v^c_{k_0}, v^c_{k_0+1}, v^c_{k_0+2}, ...)$$

на втория играч се определя аналогично. Отбелязваме, че обратната връзка зависи само от текущото състояние, а не от цялата история на процеса (марковско свойство).

Нека $\mathbf{u}^c$ и $\mathbf{v}^c$ са произволни допустими обратни връзки, съответно на първия и втория играч. Фазовата траектория

$$\mathbf{x} = (x_{k_0}, x_{k_0+1}, x_{k_0+2}, ...),$$

съответстваща на тази двойка управления се определя рекурсивно, както следва:

$$x_{k+1} = Ax_k + B_u u^c_k(x_k) + B_v v^c_k(x_k), \quad x_{k_0} = x_0, \quad k = k_0, k_0+1, k_0+2, ... \tag{10}$$

Ако функцията $\mathbf{u}^c$ е линейна, тогава $u^c_k(x_k) = K_{u_k} x_k$, където K_{u_k}, $k \in \mathbb{N}_{k_0}$ е матрицата е с размерност $m_u \times n$. Аналогично, ако $\mathbf{v}^c$ е линейна, тогава $v^c_k(x_k) = K_{v_k} x_k$, където K_{v_k}, $k \in \mathbb{N}_{k_0}$ е матрица с размерност $m_v \times n$.

Ясно е, че могат да се дефинират подобни допустими управления от смесен тип (част от компонентите им да са отворени управление а останалите да са обратни връзки в зависимост от текущото състояние на системата). Съответната траектория се определя по подобен начин.

Допустима двойка управления $(\mathbf{u}, \mathbf{v})$ се счита за *осъществима*, ако критерият J, определен в (11) приема крайна стойност за тази двойка.

Върху множеството от допустими двойки управления, предвид динамиката (9) разглеждаме следната дискретна линейно-квадратична игра върху безкраен хоризонт:

$$\inf_{\mathbf{u} \in \mathcal{U}_{\mathbf{U}}} \sup_{\mathbf{v} \in \mathcal{V}} J(x_0, k_0, \mathbf{u}, \mathbf{v}), \tag{11}$$

където целевата функция е дефинира като

$$J(x_0, k_0, \mathbf{u}, \mathbf{v}) := \sum_{k=k_0}^{\infty} \left(x_k^\intercal Q x_k + u_k^\intercal R_u u_k - v_k^\intercal R_v v_k \right).$$

Тук Q е $n \times n$ симетрична положителна полу-определена матрица, а R_u и R_v са симетрични, положително определени матрици от размерност, съответно $m_u \times m_u$ и $m_v \times m_v$.

Що се отнася до информационния модел, ние приемаме, че и двамата играчи познават всички параметри на динамиката (матриците A, B_u и B_v) и на целевата функция (матриците Q, R_u и R_v). Освен това предполагаме, че и двамата играчи имат достъп до текущото състояние на системата x_k, $k \in \mathbb{N}_{k_0}$.

Нека k е неотрицателно цяло число, удовлетворяващо неравенството $k \geq k_0$ и нека $x \in \mathbb{R}^n$. Определяме функцията *цена* $V_{\mathbf{U}} : \mathbb{R}^n \times \mathbb{N}_{k_0} \to \mathbb{R}$ на играта (9) $\div$ (11) като

$$V_{\mathbf{U}}(x, k) := \inf_{\mathbf{u} \in \mathcal{U}_{k,\mathbf{U}}} \sup_{\mathbf{v} \in \mathcal{V}_k} J(x, k, \mathbf{u}, \mathbf{v}).$$

За неограничения случай, т.е. $\mathbf{U} = \mathbb{R}^{\mathbf{m_u}} \times \mathbb{R}^{\mathbf{m_u}} \times \cdots \times \mathbb{R}^{\mathbf{m_u}} \times \cdots$, пропускаме долния индекс "**U**", така че

$$V(x, k) := \inf_{\mathbf{u} \in \mathcal{U}_k} \sup_{\mathbf{v} \in \mathcal{V}_k} J(x, k, \mathbf{u}, \mathbf{v}).$$

2.2 Предварителни резултати

Класическите задачи за оптимизация често се решават чрез използване на Принципа за оптималност на Белман. Припомняме, че според него оптималната стратегия има свойството, че независимо от първоначалното състояние и първоначалното решение, останалите решения трябва да представляват оптимална стратегия по отношение на състоянието, произтичащо от първото решение. Т.е. оптималното решение на даден етап зависи само от състоянието, достигнато до момента, и не зависи от предходната история на вземане на решения. Принципът предоставя рекурсивен израз за функцията цена, изразяваща оптималните разходи за движение по отношение на непосредствените разходи и оптималните разходи за движение на следващата времева стъпка.

По-долу показваме, че този принцип се прилага и за "мин-макс" задачи на оптималното управление с дискретно време и следователно, към разглежданата линейно-квадратична игра. Нека разгледаме по-общата версия на играта. Предвид следната динамиката:

$$x_{k+1} = f_k(x_k, u_k, v_k), \quad x_{k_0} = x_0, \quad k = k_0, k_0 + 1, k_0 + 2, ...,$$

търсим двойка допустими управления $(\mathbf{u}, \mathbf{v})$, която да е решание на следната задача:

$$\inf_{\mathbf{u} \in \mathcal{U}_{\mathbf{U}}} \sup_{\mathbf{v} \in \mathcal{V}} J_g(x_0, k_0, \mathbf{u}, \mathbf{v}),$$

където целевата функция и съответната цена на играта са дефинирани съответно като

$$J_g(x_0, k_0, \mathbf{u}, \mathbf{v}) := \sum_{k=k_0}^{\infty} g_k \left(x_k, u_k, v_k \right)$$

и

$$V_{\mathbf{U}}^g(x_0, k_0) := \inf_{\mathbf{u} \in \mathcal{U}_{\mathbf{U}}} \sup_{\mathbf{v} \in \mathcal{V}} J_g(x_0, k_0, \mathbf{u}, \mathbf{v}).$$

Теорема 2.1. *(Принцип за оптималност на Белман) Нека $k \geq k_0$, $x \in \mathbb{R}^n$ и нека цената $V^g_{\mathbf{U}}(x, k)$ приема реална стойност. Тогава е изпълнено следното равенство:*

$$V^g_{\mathbf{U}}(x, k) = \inf_{u_k \in U_k} \sup_{v_k \in \mathbb{R}^{m_v}} \{g_k(x_k, u_k, v_k) + V^g_{\mathbf{U}}(f_k(x_k, u_k, v_k), k+1)\}.$$

Бележка 2.1. *Същият резултат може да се изведе за случая когато са наложени ограничения върху управлението на втория играч. Освен това, резултатът е в сила, ако разменим местата на инфимума и супремума. По този начин, ако изберем неотрицателно цяло число $k \geq k_0$, точка $x \in \mathbb{R}^n$, положим*

$$\overline{V}^g_{\mathbf{U}}(x, k) := \sup_{\mathbf{v} \in \mathcal{V}_k} \inf_{\mathbf{u} \in \mathcal{U}_{k,\mathbf{U}}} J_g(x, k, \mathbf{u}, \mathbf{v}),$$

ако $\overline{V}^g_{\mathbf{U}}(x, k)$ е реално число, тогава принципът за оптималност на Белман налага, че

$$\overline{V}^g_{\mathbf{U}}(x, k) = \sup_{v_k \in \mathbb{R}^{m_v}} \inf_{u_k \in U_k} \left\{g_k(x_k, u_k, v_k) + \overline{V}^g_{\mathbf{U}}(f_k(x_k, u_k, v_k), k+1)\right\}.$$

Постоянното предположение, формулирано по-долу, дефинира алгебрично уравнение на Рикати за играта и налага някои условия върху участващите матрици. Тези условия гарантират свойства като симетрия, положителна определеност и специфични неравенства за P и свързаните матрици в уравнението.

Постоянно предположение (**)

Матрицата P е симетрично положително определено решение на следното алгебрично уравнение на Рикати:

$$P = Q + A^\intercal (L^{-1})^\intercal P A, \tag{12}$$

където $L := I + (B_u R_u^{-1} B_u^\intercal - B_v R_v^{-1} B_v^\intercal) P$ е обратима матрица. Освен това приемаме, че $\|A\| < 1$, а матриците $[P - \tilde{A}^\intercal P \tilde{A}]$ и $[R_v - B_v^\intercal P B_v]$ са положително определени, където

$$\tilde{A} := A - (B_u R_u^{-1} B_u^\intercal - B_v R_v^{-1} B_v^\intercal) P L^{-1} A.$$

Тук нормата на матрицата се дефинира като: $\|A\| = \max_{\|x\|=1} \|Ax\|$.

Нека разгледаме сега неограничения вариант на оригиналната линейно-квадратична игра 9 $\div$ 11. Следващото твърдение предоставя резултат от типа "мин-макс" за тази игра. Получената еквивалентност има значение в теорията на игрите и оптимизацията, особено в сценарии, характеризиращи се с конкурентни или състезателни отношения. Докато съществуващата литература изобилства от "мин-макс" теореми, те често изискват компактност на поне едно от участващите множества. В нашето изследване демонстрираме, че за квадратичната целева функция такива условия за компактност не са необходими.

Твърдение 2.1. *Нека Постоянното предположение (**) е изпълнено. Тогава*

$$\min_{u \in \mathbb{R}^{m_u}} \max_{v \in \mathbb{R}^{m_v}} \{x^\intercal Q x + u^\intercal R_u u - v^\intercal R_v v + (Ax + B_u u + B_v v)^\intercal P (Ax + B_u u + B_v v)\}$$

$$= \max_{v \in \mathbb{R}^{m_v}} \min_{u \in \mathbb{R}^{m_u}} \{x^\intercal Q x + u^\intercal R_u u - v^\intercal R_v v + (Ax + B_u u + B_v v)^\intercal P (Ax + B_u u + B_v v)\}$$

за всяка точка $x \in \mathbb{R}^n$.

2.3 Приближение на ограничената игра

Както беше отбелязано по-рано, основната идея на нашия подход е да се апроксимира играта, описана в (9) $\div$ (11) с игра върху краен времеви интервал. В тази ситуация, разглеждаме произволни допустими управления $\mathbf{u}$ и $\mathbf{v}$, принадлежащи съответно на множествата $\mathcal{U}_{\mathrm{U}}$ и $\mathcal{V}$. За всяко $\kappa \in \mathcal{N}_{k_0}$, разглеждаме играта

$$\inf_{\mathbf{u}^\kappa \in \mathcal{U}_{\mathrm{U}}^\kappa} \sup_{\mathbf{v}^\kappa \in \mathcal{V}^\kappa} J^\kappa(x_0, k_0, \mathbf{u}^\kappa, \mathbf{v}^\kappa) \tag{13}$$

предмет на динамиката (9), където $\mathbf{u}^\kappa$ и $\mathbf{v}^\kappa$ са определени от (7) и (8). Целевата функция на играта се определя като

$$J^\kappa(x_0, k_0, \mathbf{u}^\kappa, \mathbf{v}^\kappa) := x_{\kappa+1}^\intercal P x_{\kappa+1} + \sum_{k=k_0}^{\kappa} \left(x_k^\intercal Q x_k + u_k^\intercal R_u u_k - v_k^\intercal R_v v_k \right),$$

където P е симетричното положително определено решение на алгебричното уравнение на Рикати (12). Функцията цена $V_{\mathrm{U}}^\kappa : \mathbb{R}^n \times \{k_0, k_0+1, k_0+2, ..., \kappa\} \to \mathbb{R}$ на тази игра се определя като

$$V_{\mathbf{U}}^\kappa(x_0, k_0) := \inf_{\mathbf{u}^\kappa \in \mathcal{U}_{\mathrm{U}}^\kappa} \sup_{\mathbf{v}^\kappa \in \mathcal{V}^\kappa} J^\kappa(x_0, k_0, \mathbf{u}^\kappa, \mathbf{v}^\kappa).$$

Нека фиксираме произволна точка $x \in \mathbb{R}^n$ и нека k е произволен елемент от индексното множество

$$\mathcal{I}_{k_0}^\kappa := \{k_0, k_0+1, ..., \kappa\}.$$

Тогава е ясно, че

$$V_{\mathbf{U}}^\kappa(x, \kappa) = x^\intercal P x. \tag{14}$$

Допълнително от принципа за оптималност на Белман получаваме, че

$$V_{\mathbf{U}}^\kappa(x, k) = \inf_{u_k \in U_k} \sup_{v_k \in \mathbb{R}^{m_v}} \left\{ x^\intercal Q x + u_k^\intercal R_u u_k - v_k^\intercal R_v v_k + V_{\mathrm{U}}^\kappa(Ax + B_u u_k + B_v v_k, k+1) \right\}.$$

Отбелязваме, че (14) в комбинация с равенството по-горе ни позволява да изчислим цената $V_{\mathbf{U}}^\kappa(x, k)$ не само за матрицата P, но и за произволна положително определена матрица. Да предположим, че знаем цената $V_{\mathbf{U}}^\kappa(x, k)$. Следващото твърдение, което изследва асимптотичното поведение на фазовата траектория ще бъде използвано за установяване на връзката между неограничената игра върху краен времеви интервал (13) с такава върху безкраен хоризонт (11).

Твърдение 2.2. *Нека Постоянното предположение (**) е изпълнено, нека $k_0 \in \mathbb{N}_0$, нека $(\mathbf{u}, \mathbf{v}) \in \ell^2(\mathbb{N}_{k_0}, \mathbb{R}^{m_u}) \times \ell^2(\mathbb{N}_{k_0}, \mathbb{R}^{m_v})$ са произволни ℓ^2-редици и нека $\mathbf{x}$ е съответната фазова траектория. Тогава*

$$\lim_{k \uparrow +\infty} x_k = \mathbf{0}.$$

2.3.1 Решение на играта без ограничения

Следващата теорема установява условие за съществуването на равновесие на Неш за неограничената игра (9) $\div$ (11), осигурявайки рамка за идентифициране на оптималните стратегии.

Теорема 2.2. *Нека Постоянното предположение (**) е изпълнено. Дефинираме матрици K_u и K_v, както следва:*

$$K_u = R_u^{-1} B_u^{\intercal} P L^{-1} A \quad \text{и} \quad K_v = R_v^{-1} B_v^{\intercal} P L^{-1} A, \tag{15}$$

и нека $\|A - B_u K_u\| < 1$, $\|A + B_v K_v\| < 1$ и $\|A - B_u K_u + B_v K_v\| < 1$. Тогава

$$\begin{aligned} V(x_0, k_0) &= x_0^{\intercal} P x_0 = J(x_0, k_0, \tilde{\mathbf{u}}, \tilde{\mathbf{v}}) \\ &= \min_{\mathbf{u} \in \mathcal{U}} \max_{\mathbf{v} \in \mathcal{V}} J(x_0, k_0, \mathbf{u}, \mathbf{v}) = \max_{\mathbf{v} \in \mathcal{V}} \min_{\mathbf{u} \in \mathcal{U}} J(x_0, k_0, \mathbf{u}, \mathbf{v}), \end{aligned} \tag{16}$$

където $\tilde{\mathbf{u}} := (\tilde{u}_{k_0}, \tilde{u}_{k_0+1}, \tilde{u}_{k_0+2}, \dots)$ и $\tilde{\mathbf{v}} := (\tilde{v}_{k_0}, \tilde{v}_{k_0+1}, \tilde{v}_{k_0+2}, \dots)$ са линейни обратни връзки, дефинирани като $\tilde{u}_k(x) = -K_u x$ и $\tilde{v}_k(x) = K_v x$ за всяко $k \in \mathbb{N}_{k_0}$ и за всяка точка $x \in \mathbb{R}^n$.

2.3.2 δ-околност с неактивни ограничения

Сега се връщаме към оригиналната игра с наложени ограничения върху управлението на първия играч. За да установим следващия резултат се нуждаем от допълнително предположение относно множествата на управления на минимизиращия играч, а именно, че множествата U_k, $k \in \mathbb{N}_{k_0}$ съдържат изпъкнала околност U на началото в $\mathbb{R}^{m_u}$. При това предположение следващото твърдение установява съществуването на околност на началото в $\mathbb{R}^n$, където тези ограничения не са активни.

Нека означим с $\tilde{\mathbf{x}}_\kappa^y := (\tilde{x}_\kappa^y, \tilde{x}_{\kappa+1}^y, \dots)$ траекторията на дискретната система

$$\tilde{x}_{k+1} = \tilde{A} \tilde{x}_k, \quad \tilde{x}_\kappa = y, \quad k \in \mathbb{N}_\kappa,$$

която започва от точката y в момента κ, където матрицата $\tilde{A}$ е въведена в Постоянното предположение (**), т.е. $\tilde{\mathbf{x}}_\kappa^y$ е породена от линейните обратни връзки $-K_u x$ и $K_v x$, определени чрез (15). С $\bar{\mathbf{B}}^n$ означаваме затвореното единично кълбо в $\mathbb{R}^n$.

Твърдение 2.3. *Нека Постоянното предположение (**) и предположенията на Теорема 2.2 са изпълнени. Нека множествата U_k, $k = k_0, k_0 + 1, \dots$, съдържат изпъкнала околност U на началото в $\mathbb{R}^{m_u}$. Тогава съществуват положителни реални числа δ_0 и δ, удовлетворяващи неравенството $\delta_0 > \delta$, такива че:*

(i) включването $-K_u x \in U$ е изпълнено за всяка точка $x \in \delta_0 \bar{\mathbf{B}}^n$;

(ii) за всяка точка $y \in \delta \bar{\mathbf{B}}^n$ и за всяко $\kappa \geq k_0$ е изпълнено следното включване: $\tilde{x}_k^y \in \delta_0 \bar{\mathbf{B}}^n$ за всяко $k \in \mathbb{N}_\kappa$ и редицата $\{\tilde{x}_k^y\}_{k=\kappa}^{\infty}$ клони към началото при $k \to +\infty$.

Занапред ще фиксираме параметъра δ, въведен в Твърдение 2.3, дефиниращ околността на началото в $\mathbb{R}^n$.

Бележка 2.2. *Твърдение 2.2 налага съществуването на момент във времето $\tau \geq k_0$, такъв че фазовата траектория $\mathbf{x} = (x_{k_0}, x_{k_0+1}, x_{k_0+2}, \dots)$, породена от произволна двойка ℓ^2-редици $(\mathbf{u}, \mathbf{v})$, започваща от произволна точка $y \in \mathbb{R}^n$, влиза в δ-околност на началото, т.е. $x_\tau \in \delta \bar{\mathbf{B}}^n$.*

2.4 Достатъчни условия за оптималност

Следната Теорема 2.3 демонстрира, че при подходящ избор на терминален момент, цената на играта върху определения от него краен интервал е равна на цената на играта върху безкраен хоризонт.

Теорема 2.3. *Нека Постоянното предположение (**) е изпълнено, нека* $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ *е допустима двойка управления за играта* $(9) \div (11), \delta > 0$ *е реалното число въведено в Твърдение 2.3 и нека* $\bar{\mathbf{x}}$ *е съответната фазова траектория. Ако равенството*

$$V_{\mathbf{U}}(x_0, k_0) = J(x_0, k_0, \bar{\mathbf{u}}, \bar{\mathbf{v}})$$

е изпълнено, тогава съществува неотрицателно цяло число $\kappa \geq k_0$*, такова че* $\bar{x}_\kappa \in \delta\bar{\mathbf{B}}$ *и*

$$V_{\mathbf{U}}(x_0, k_0) = J(x_0, k_0, \bar{\mathbf{u}}^\kappa, \bar{\mathbf{v}}^\kappa) = V_{\mathbf{U}}^\kappa(x_0, k_0),$$

където $\bar{\mathbf{u}}^\kappa$ *и* $\bar{\mathbf{v}}^\kappa$ *са определени от (7) и (8).*

Следващият резултат показва как ограничената игра може да бъде решена с помощта на решението на съответната, подходящо дефинирана игра на краен хоризонт. По същество следствието твърди, че ако имаме решение на играта върху краен интервал и го ”разширим” чрез прилагане на линейните обратни връзки към неограничената игра, ще получим решението на ограничената игра върху безкраен хоризонт.

Следствие 2.1. *Нека Постоянното предположение (**) е изпълнено, нека* $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ *е допустима двойка управления за играта* $(9) \div (11)$ *и нека* $\kappa \geq k_0$ *да е достатъчно голямо, така че точката* $\bar{x}_\kappa \in \bar{\mathbf{x}}^\kappa$ *да принадлежи на* $\delta\bar{\mathbf{B}}^n$ *(*δ *е въведено в Твърдение 2.3), където* $\bar{\mathbf{x}}^\kappa$ *е траекторията, съответстваща на двойката* $(\bar{\mathbf{u}}^\kappa, \bar{\mathbf{v}}^\kappa)$*. Нека*

$$V_{\mathbf{U}}^\kappa(x_0, k_0) = J^\kappa(x_0, k_0, \bar{\mathbf{u}}^\kappa, \bar{\mathbf{v}}^\kappa),$$

тогава

$$V_{\mathbf{U}}^\kappa(x_0, k_0) = J(x_0, k_0, \hat{\mathbf{u}}, \hat{\mathbf{v}}) = V_{\mathbf{U}}(x_0, k_0),$$

където

$$(\hat{\mathbf{u}}, \hat{\mathbf{v}}) := \begin{cases} (\bar{\mathbf{u}}^\kappa, \bar{\mathbf{v}}^\kappa) & \text{за } k \in \{k_0, k_0 + 1, ..., \kappa\}, \\ (\mathbf{u}^c_{\kappa+1}, \mathbf{v}^c_{\kappa+1}) & \text{за } k \in \{\kappa + 1, \kappa + 2, ...\}, \end{cases}$$

със $u^c_k(x_k) = -K_u x_k$ *и* $v^c_k(x_k) = K_v x_k$ *за всяко* $k = \kappa + 1, \kappa + 2, ...$*, където матриците* K_u *и* K_v *са дефинирани от равенствата (15).*

3 Дискретна динамична игра

3.1 Постановка на играта

Нека означим с $\mathbb{N}$ множеството от всички неотрицателни цели числа и с $\mathbb{R}^n$, n-мерното евклидово пространство. Фиксираме вектор x_0 в отворено подмножество $G \subset \mathbb{R}^n$ и разглеждаме динамична игра с дискретно време между двама души на безкраен времеви хоризонт. Динамиката на играта се описва от следната система от диференчни уравнения:

$$x_{k+1} = f_k(x_k, u_k, v_k), \quad x_{k_0} = x_0, \quad u_k \in U_k, \quad v_k \in V_k, \quad k \in \mathbb{N}. \tag{17}$$

Въвеждаме следните означения и предположения:

- $x_0 \in \mathbb{R}^n$ е началният вектор на състоянието, елемент от отворено подмножество $G \subset \mathbb{R}^n$.
- $k_0 \in \mathbb{N}$ е началният момент, който за простота се приема за 0.
- $x_k \in \mathbb{R}^n$ обозначава състоянието на системата в момента k.
- $U_k \subseteq \mathbb{R}^{m_u}$ и $V_k \subseteq \mathbb{R}^{m_v}$, за всяко $k \in \mathbb{N}$ са непразни, затворени и изпъкнали множества, представляващи пространствата от управления за двамата играчи.
- Функцията $f_k : G \times \tilde{U}_k \times \tilde{V}_k \to \mathbb{R}^n$, за всяко $k \in \mathbb{N}$ е непрекъснато-диференцируема.
- $\tilde{U}_k$ и $\tilde{V}_k$ са отворени множества, които съдържат съответно U_k и V_k за всяко $k \in \mathbb{N}$, разширявайки домейна на функцията f_k, за да се подсигури нейната непрекъснато-диференцируемост.
- n, m_u и m_v са размерностите на съответните векторни пространства.

Играчите влияят върху еволюцията на системата чрез своя избор на функции $u : \mathbb{N} \to \mathbb{R}^{m_u}$ за първия играч и $v : \mathbb{N} \to \mathbb{R}^{m_v}$ за втория играч.

Наричаме двойката редици управления $(\mathbf{u}, \mathbf{v})$, където $\mathbf{u} := \{u_k\}_{k=0}^{\infty}$, $\mathbf{v} := \{v_k\}_{k=0}^{\infty}$, *допустима*, ако включванията $u_k \in U_k$ и $v_k \in V_k$ са изпълнени за всяко $k \in \mathbb{N}$. $\mathcal{U}$ и $\mathcal{V}$ обозначават множествата от всички допустими стратегии, съответно на първия и втория играч.

За дадена допустима двойка $(\mathbf{u}, \mathbf{v})$, уравнението (17) поражда *фазова траектория* $x_0, x_1, \dots$. Отбелязваме, че тази траектория може да бъде удължена до безкрайност или до минимално число k, за което е изпълнена следната връзка: $f_k(x_k, u_k, v_k) \notin G$ (ако съществува такова k). В първия случай наричаме тройката $(\mathbf{x}, \mathbf{u}, \mathbf{v})$ *допустим процес*, където $\mathbf{x} := \{x_k\}_{k=0}^{\infty}$.

За даден допустим процес $(\mathbf{x}, \mathbf{u}, \mathbf{v})$ елементите на фазовата траектория $\{x_k\}_{k=0}^{\infty}$ може да бъдат представени като:

$$x_{k+1} := f_k^{(u_k, v_k)} \circ f_{k-1}^{(u_{k-1}, v_{k-1})} \circ \dots \circ f_0^{(u_0, v_0)}(x_0), \quad k \in \mathbb{N}. \tag{18}$$

Символът "$\circ$" обозначава композицията на съответните изображения, а $f_k^{(u,v)}(x)$ обозначава $f_k(x, u, v)$.

Предвид динамиката (17), разглеждаме следната динамична игра върху безкраен хоризонт с дискретно време:

$$\min_{\mathbf{u}\in\mathcal{U}} J(x_0, \mathbf{u}, \mathbf{v}), \quad \max_{\mathbf{v}\in\mathcal{V}} J(x_0, \mathbf{u}, \mathbf{v}), \tag{19}$$

чиято целева функция J, се определя като:

$$J(x_0, \mathbf{u}, \mathbf{v}) = \sum_{k=0}^{\infty} g_k(x_k, u_k, v_k). \tag{20}$$

От (19) е видно, че първият играч се стреми да "минимизира" тази целева функция, докато целта на втория играч е да я "максимизира". Тук функциите $g_k : G \times \tilde{U}_k \times \tilde{V}_k \to \mathbb{R}$, $k \in \mathbb{N}$, се приемат за непрекъснато-диференцируеми.

За да изясним смисъла на тази задача, за всяко $K \in \mathbb{N}$ полагаме

$$J_K(x_0, \mathbf{u}, \mathbf{v}) = \sum_{k=0}^{K} g_k(x_k, u_k, v_k).$$

Определение 3.1. *Допустимият процес* $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ *се нарича* слабо изпреварващо оптимален*, ако за всяко $\varepsilon > 0$, за всяко положително цяло число K и за всеки допустим процес $(\mathbf{x}, \mathbf{u}, \mathbf{v})$ съществува положително цяло число $\kappa > K$, такова че*

$$J_\kappa(x_0, \bar{\mathbf{u}}, \mathbf{v}) - \varepsilon \le J_\kappa(x_0, \bar{\mathbf{u}}, \bar{\mathbf{v}}) \le J_\kappa(x_0, \mathbf{u}, \bar{\mathbf{v}}) + \varepsilon.$$

Двойката $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$, която образува слабо изпреварващ оптимален процес, се нарича *слабо изпреварващо равновесие на Неш.*

Нека $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ е слабо изпреварващ оптимален процес. За всеки $k \in \mathbb{N}$ и за всеки вектор ξ, означаваме с $x^{k,\xi} = (x_k, x_{k+1}, \ldots)$ траекторията, породена от (17) с начално състояние $x_k = \xi$ в момента k, нейните елементи могат да се представят по следния начин:

$$x_{s+1}^{k,\xi} := f_s^{(\bar{u}_s,\bar{v}_s)} \circ f_{s-1}^{(\bar{u}_{s-1},\bar{v}_{s-1})} \circ \ldots \circ f_k^{(\bar{u}_k,\bar{v}_k)}(\xi), \quad s = k, k+1, \ldots .$$

Бележка 3.1. *Горната дефиниция, предоставяща смисъла на задачата е модификация на тази в [3] в контекста на динамична игра. В този ред, подобно на статията, въвеждаме следните предположения и дефиниции.*

Предположение 3.1. *За всяко $k \in \mathbb{N}$ съществуват реално число $\alpha_k > 0$ и редица $\{\beta_s^k\}_{s=k}^{\infty}$ с $\sum_{s=k}^{\infty} \beta_s^k < \infty$, такива че $\alpha_k \mathbf{B}(\bar{x}_k) \subset G$, за всеки вектор $\xi \in \alpha_k \bar{\mathbf{B}}(\bar{x}_k)$, траекторията $x^{k,\xi}$ е безкрайна и*

$$\sup_{\xi\in\alpha_k\bar{\mathbf{B}}(\bar{x}_k)} \left\| \frac{\partial}{\partial \xi} g_s(x_s^{k,\xi}, \bar{u}_s, \bar{v}_s) \right\| \le \beta_s^k,$$

където $\alpha_k \bar{\mathbf{B}}(\bar{x}_k)$ е затвореното кълбо в $\mathbb{R}^n$ с център в $\bar{x}_k$ и радиус α_k.

Предположение 3.1 налага, че реда

$$\sum_{s=k}^{\infty} \frac{\partial}{\partial \xi} g_s(x_s^{k,\xi}, \bar{u}_s, \bar{v}_s), \quad k = 1, 2, ...$$

е абсолютно сходящ, равномерно по отношение на $\xi \in \alpha_k \bar{\mathbf{B}}(\bar{x}_k)$.

От тъждеството

$$g_s(x_s^{k,\xi}, \bar{u}_s, \bar{v}_s) = g_s(f_{s-1}^{(\bar{u}_{s-1}, \bar{v}_{s-1})} \circ f_{s-2}^{(\bar{u}_{s-2}, \bar{v}_{s-2})} \circ ... \circ f_k^{(\bar{u}_k, \bar{v}_k)}(\xi), \bar{u}_s, \bar{v}_s)$$

и от правилото за диференциране на съставна функция имаме:

$$\frac{\partial}{\partial \xi} g_s(x_s^{k,\xi}, \bar{u}_s, \bar{v}_s) = \frac{\partial}{\partial x} g_s(x_s^{k,\xi}, \bar{u}_s, \bar{v}_s) \prod_{i=s-1}^{k} \frac{\partial}{\partial x} f_i(x_i^{k,\xi}, \bar{u}_i, \bar{v}_i), \tag{21}$$

където се използват следните означения:

$$\prod_{i=s-1}^{k} A_i := \begin{cases} A_{s-1} A_{s-2} ... A_k, & \text{ако } s > k \\ I, & \text{ако } s \leq k. \end{cases}$$

Тук $A_i := \frac{\partial}{\partial x} f_i(x_i^{k,\xi}, \bar{u}_i, \bar{v}_i)$ и I обозначава единичната матрицата с размерност $n \times n$. Освен това тук използваме символа $\prod$ вместо обичайния символ $\prod$, за да покажем, че "нарастването" на текущия индекс i е с обратен знак (тъй като $s > k$).

Следвайки отново [3], дефинираме *спрегнатата редица* $\psi := \{\psi_k\}_{k=1}^{\infty}$:

$$\psi_k = \sum_{s=k}^{\infty} \frac{\partial}{\partial \xi} g_s(x_s^{k,\xi}, \bar{u}_s, \bar{v}_s)_{|\xi = \bar{x}_k}, \quad k = 1, 2, ... \, .$$

Бележка 3.2. *Предположение 3.1 всъщност налага, че* $\|\psi_k\| < \infty, \; k = 1, 2, ...$ *Освен това получаваме от (21), че*

$$\psi_k = \sum_{s=k}^{\infty} \frac{\partial}{\partial x} g_s(\bar{x}_s, \bar{u}_s, \bar{v}_s) \prod_{t=s-1}^{k} \frac{\partial}{\partial x} f_t(\bar{x}_t, \bar{u}_t, \bar{v}_t) \tag{22}$$

$$\text{(като се има предвид, че } x_s^{k,\bar{x}_k} = \bar{x}_s)$$

Равенство (22) налага, че така дефинирана, спрегнатата редица удовлетворява *спрегнато уравнение*

$$\psi_k = \psi_{k+1} \frac{\partial}{\partial x} f_k(\bar{x}_k, \bar{u}_k, \bar{v}_k) + \frac{\partial}{\partial x} g_k(\bar{x}_k, \bar{u}_k, \bar{v}_k), \quad k = 1, 2, ... \, . \tag{23}$$

3.2 Предварителни резултати

Теорема 3.1, изложена в [3] като Теорема 2.2, предоставя локално условие за максимум за Хамилтониана, което играе важна роля в доказателството на основния резултат в тази глава. Това условие е изпълнено за определено решение $\psi = \{\psi_k\}_{k=1}^{\infty}$, на спрегнатото уравнение (23). За да се идентифицира „правилното“ решение на това уравнение, са необходими допълнителни условия, обикновено под формата на условие за трансверзалност на ψ_k, при $k \to \infty$. За формулировката на теоремата ще ни е необходимо следното определение:

Определение 3.2 (Допирателен конус на Булиган $T_S(\bar{y})$). *Нека Y е банахово пространство и S е непразно затворено подмножество на Y. Нека $\bar{y}$ е произволна точка от S. Множеството $T_S(\bar{y})$, дефинирано като съвкупността от всички точки $w \in Y$, за които съществуват редица от положителни реални числа $\{t^\mu\}_{\mu=1}^{\infty} \downarrow 0$ и редица от точки $\{w^\mu\}_{\mu=1}^{\infty} \subset Y$, сходяща към w, удовлетворяващи включването $\bar{y} + t^\mu w^\mu \in S$ за всяко $\mu = 1, 2, ...$, се нарича* допирателен конус на Булиган *към затвореното множество S в точката $\bar{y} \in S$ (виж [4], Глава 4.1).*

Теоремата гласи:

Теорема 3.1 (Локално условие за максимум, Теорема 2.2 в [3]). *Нека Предположение 3.4 е изпълнено, нека двойката $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ е слабо изпреварващ оптимален процес (в контекста на задачата в [3]) и нека спрегнатата редица ψ да бъде определена от (32). Тогава за всяко $k \in \mathbb{N}$ е изпълнено следното локално* условие за максимум*:*

$$\Big(\frac{\partial}{\partial v} g_k(\bar{x}_k, \bar{v}_k) + \psi_{k+1} \frac{\partial}{\partial v} f_k(\bar{x}_k, \bar{v}_k)\Big) w \le 0 \quad \textit{за всеки } w \in T_{V_k}(\bar{v}_k),$$

където $T_{V_k}(\bar{v}_k)$ е допирателният конус на Булиган, въведен от Определение 3.2.

Следващият резултат е следствие на Теорема 3.1. За формулировката му въвеждаме матриците Z_k, $k \in \mathbb{N}$, дефинирани като:

$$Z_k := \frac{\partial}{\partial x} f_{k-1}(\bar{x}_{k-1}, \bar{u}_{k-1}, \bar{v}_{k-1}) \frac{\partial}{\partial x} f_{k-2}(\bar{x}_{k-2}, \bar{u}_{k-2}, \bar{v}_{k-2}) ... \frac{\partial}{\partial x} f_0(\bar{x}_0, \bar{u}_0, \bar{v}_0). \tag{24}$$

Следствие 3.1. *Нека Предположение 3.1 е изпълнено, нека тройката $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ е слабо изпреварващ оптимален процес и нека спрегнатата редица $\psi = \{\psi_k\}_{k=1}^{\infty}$ е дефинирана от (22). Тогава, за всяко $k \in \mathbb{N}$ е изпълнено следното локално* условие за максимум*:*

$$\Big(\frac{\partial}{\partial v} g_k(\bar{x}_k, \bar{u}_k, \bar{v}_k) + \psi_{k+1} \frac{\partial}{\partial v} f_k(\bar{x}_k, \bar{u}_k, \bar{v}_k)\Big) w \le 0 \quad \text{за всеки } w \in T_{V_k}(\bar{v}_k), \tag{25}$$

както и условието за трансверзалност

$$\lim_{k \to +\infty} \psi_k Z_k = 0. \tag{26}$$

За разглежданата игра дефинираме *функцията на Хамилтон* $\mathcal{H}_k : G \times \mathbb{R}^{m_u} \times \mathbb{R}^{m_v} \times \mathbb{R}^n \to \mathbb{R}$, включваща спрегнатата променлива ψ. Функцията $\mathcal{H}_k$ е дефинирана за всяко $k \in \mathbb{N}$ като:

$$\mathcal{H}_k(x, u, v, \psi) := g_k(x, u, v) + \psi f_k(x, u, v).$$

Използвайки тази функция, отношенията (17), (23) и (25) могат да бъдат изразени по следния начин:

$$\bar{x}_{k+1} = \frac{\partial}{\partial \psi_{k+1}} \mathcal{H}_k(\bar{x}_k, \bar{u}_k, \bar{v}_k, \psi_{k+1}), \tag{27}$$

$$\psi_k = \frac{\partial}{\partial x_k} \mathcal{H}_k(\bar{x}_k, \bar{u}_k, \bar{v}_k, \psi_{k+1}), \tag{28}$$

$$\frac{\partial}{\partial v} \mathcal{H}_k(\bar{x}_k, \bar{u}_k, \bar{v}_k, \psi_{k+1}) w \le 0 \quad \text{за всеки } w \in T_{V_k}(\bar{v}_k). \tag{29}$$

3.3 Необходимо условие за оптималност

Отличителна черта на задачите на оптималното управление с дискретно време в сравнение с техните аналози с непрекъснато време е локалният характер на принципа на максимум на Понтрягин върху ограничени времеви хоризонти. По-конкретно, при отсъствието на допълнителни условия от типа за вдлъбнатост (когато се адресира задача за максимизиране), условието за максимум, свързано с Хамилтониана, служи само като необходимо условие за локален максимум. Ние демонстрираме, че този принцип е приложим и за задачи на безкраен хоризонт.

За получаването на основния резултат в настоящата глава, ни е необходимо да допуснем следното основно предположение. Да разгледаме тройката $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$, която представлява слабо изпреварващ оптимален процес и спрегнатата редица $\psi = \{\psi_k\}_{k=1}^{\infty}$, определена съгласно (22).

Предположение 3.2. *Следните условия са изпълнени:*

(i) Функцията $g_k(\bar{x}_k, \cdot, \bar{v}_k) : U_k \to \mathbb{R}$ е изпъкнала.
Когато $\psi_{k+1}^j > 0$ (j-компонентата на ψ_{k+1}), j-компонентата $f_k^j(\bar{x}_k, \cdot, \bar{v}_k) : U_k \to \mathbb{R}$ на векторната функция $f_k(\bar{x}_k, \cdot, \bar{v}_k) : U_k \to \mathbb{R}^n$ е изпъкнала, а когато $\psi_{k+1}^j \leq 0$, тя е вдлъбната.

(ii) Функцията $g_k(\bar{x}_k, \bar{u}_k, \cdot) : V_k \to \mathbb{R}$ е вдлъбната.
Когато $\psi_{k+1}^j \geq 0$ (j-компонентата на ψ_{k+1}), j-компонентата $f_k^j(\bar{x}_k, \bar{u}_k, \cdot) : V_k \to \mathbb{R}$ на векторната функция $f_k(\bar{x}_k, \bar{u}_k, \cdot) : V_k \to \mathbb{R}^n$ е вдлъбната, а когато $\psi_{k+1}^j < 0$, тя е изпъкнала.

Бележка 3.3. *Предположение 3.2 налага, че за всяко $k \in \mathbb{N}$:*

- *функцията $\mathcal{H}_k(\bar{x}_k, \cdot, \bar{v}_k, \psi_{k+1}) : U_k \to \mathbb{R}$ е изпъкнала;*
- *функцията $\mathcal{H}_k(\bar{x}_k, \bar{u}_k, \cdot, \psi_{k+1}) : V_k \to \mathbb{R}$ е вдлъбната.*

При тази постановка на задачата и направените предположения, представяме основният резултат в тази глава. Той формулира необходимо условие за оптималност от вид принцип на максумума на Понтрягин за дискретната игра върху безкраен хоризонт.

Теорема 3.2 (Необходимо условие за оптималност)**.** *Нека предположения 3.1 и 3.2 са изпълнени, нека двойката управления $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ определя слабо изпреварващо равновесие на Неш за играта (17) $\div$ (19) и нека спрегнатата редица $\psi = \{\psi_k\}_{k=1}^{\infty}$ е определена от (22). Тогава спрегнатата редица $\psi = \{\psi_k\}_{k=1}^{\infty}$ е решение на спрегнатата система (28) и за всяко $k \in \mathbb{N}$, са изпълнени следните условия:*

(i) условие min – max

$$\min_{u \in U_k} \mathcal{H}_k(\bar{x}_k, u, \bar{v}_k, \psi_{k+1}) = \mathcal{H}_k(\bar{x}_k, \bar{u}_k, \bar{v}_k, \psi_{k+1}) = \max_{v \in V_k} \mathcal{H}_k(\bar{x}_k, \bar{u}_k, v, \psi_{k+1});$$

(ii) условие за трансверзалност

$$\lim_{k \to +\infty} \psi_k Z_k = 0,$$

където Z_k се определя от (24).

3.4 Дискретна задача на оптималното управление

В широкия спектър динамични игри, задачите на оптималното управление без смущения се появяват като отделна и специализирана подгрупа.

В тази перспектива задачата на оптималното управление се разглежда като игра с единствен играч, като целта на този играч остава същата – да определи "най-добрите" входни функции. Това преформулиране ни дава възможност да разгледаме задачата на оптималното управление като ситуация, в която влиянието на множество играчи е отсъстващо.

Разделът установява ново достатъчно условие за оптималност, което се базира на определени предположения, и демонстрира значението му за разрешаване на задачата върху безкраен времеви хоризонт. Разглеждат се поредица от условия и изводи, които са фундаментални за теоретичното разбиране на оптималността в контекста на задачата.

3.4.1 Постановка на задачата

В описания сценарий, системата с дискретно време (17) е формулирана без външни влияния, т.е. стойностите на управленията на втория играч $v_k, k \in \mathbb{N}$ (представляващи смущенията), са зададени като нули, т.е. фиксираме $\mathbf{v} = \mathbf{0}$. Ще считаме също, че множеството G е цялото пространство $\mathbb{R}^n$.

Ясно е, че в този случай, динамиката на системата се определя, както следва:

$$x_{k+1} = f_k(x_k, u_k), \quad x_{k_0} = x_0, \quad u_k \in U_k, \quad k \in \mathbb{N}. \tag{30}$$

Развитието на системата ще зависи от избора на функция само на единствения играч, която представлява редица от управления, обозначена със

$$\mathbf{u} := (u_0, u_1, \ldots, u_k, \ldots).$$

Предположение 3.3. *Всяка допустима редица от управления* $\mathbf{u}$ *(т.е. елементите ѝ да удовлетворяват включването* $u_k \in U_k,\ k \in \mathbb{N}$*) е елемент на банахово пространство* $(\mathbb{X}, \|\cdot\|)$.

Над всички допустими процеси $(\mathbf{x}, \mathbf{u})$, разглеждаме следната задача на оптималното управление:

$$J(x_0, \mathbf{u}) \to \min, \tag{31}$$

където е ясно, че критерият на задачата се определя като

$$J(x_0, \mathbf{u}) := \sum_{k=0}^{\infty} g_k(x_k, u_k).$$

Наричаме допустимата редица управления $\mathbf{u}^*$ *локално оптимална*, ако съществува околност $N \subset \mathbb{R}^m$ на $\mathbf{u}^*$, таква че за всяка допустима редица управления $\mathbf{u} \in N$ имаме изпълнено следното неравенството

$$J(x_0, \mathbf{u}^*) \leq J(x_0, \mathbf{u}).$$

Ако това неравенство е изпълнено за всички допустими редици управления $\mathbf{u}$, тогава редицата управления $\mathbf{u}^*$ се нарича *глобално оптимална*.

Локално (глобално) оптималната редица управления $\mathbf{u}^*$ води до локално (глобално) оптимална фазова траектория $\mathbf{x}^*$, породена от динамиката на системата (30). Получената двойка $(\mathbf{x}^*, \mathbf{u}^*)$ наричаме *локално (глобално) оптимален процес*.

Нека $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ е произволен осъществим процес (т.е. $J(x_0, \bar{\mathbf{u}}) < \infty$). За всяко положително цяло число k и всеки вектор ξ, представяме елементите на породената от (30) траектория $x^{k,\xi} = (x_k, x_{k+1}, \ldots)$ с начално състояние $x_k = \xi$ в момента k, като

$$x_{s+1}^{k,\xi} := f_s^{(\bar{u}_s)} \circ f_{s-1}^{(\bar{u}_{s-1})} \circ \cdots \circ f_k^{(\bar{u}_k)}(\xi), \quad s = k, k+1, \ldots .$$

Подобно на [3], въвеждаме следното предположение, представляващо свободния от смущения вариант на Предположение 3.1:

Предположение 3.4. *За всяко положително цяло число k, съществуват реално число $\alpha_k > 0$ и редица $\{\beta_s^k\}_{s=k}^{\infty}$ с $\sum_{s=k}^{\infty} \beta_s^k < \infty$, такива че следното неравенството е изпълнено:*

$$\sup_{\xi \in \alpha_k \bar{\mathbf{B}}(\bar{x}_k)} \left\| \frac{\partial}{\partial \xi} g_s(x_s^{k,\xi}, \bar{u}_s) \right\| \leq \beta_s^k, \quad s = k, k+1, \ldots ,$$

където $\alpha_k \bar{\mathbf{B}}(\bar{x}_k)$ е затвореното кълбо в $\mathbb{R}^n$ с център $\bar{x}_k$ и радиус α_k.

В съответствие с [3], вариантът без смущения на *спрегнатата редица* $\psi := \{\psi_k\}_{k=1}^{\infty}$ се дефинира както:

$$\psi_k = \sum_{s=k}^{\infty} \frac{\partial}{\partial \xi} g_s(x_s^{k,\xi}, \bar{u}_s)_{|\xi = \bar{x}_k} = \sum_{s=k}^{\infty} \frac{\partial}{\partial x} g_s(\bar{x}_s, \bar{u}_s) \prod_{i=s-1}^{k} \frac{\partial}{\partial x} f_i(\bar{x}_i, \bar{u}_i). \tag{32}$$

С прилагането на второто равенство в (32) става очевидно, че спрегнатата редица, дефинирана по този начин, удовлетворява *спрегнатото уравнение*.

$$\psi_k = \frac{\partial}{\partial x} g_k(\bar{x}_k, \bar{u}_k) + \psi_{k+1} \frac{\partial}{\partial x} f_k(\bar{x}_k, \bar{u}_k), \quad k = 1, 2, \ldots .$$

Използвайки спрегнатата редица ψ, за всяко $k \in \mathbb{N}$ въвеждаме съответната за задачата *Хамилтонова функция* $\mathcal{H}_k : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$, като:

$$\mathcal{H}_k(x, u, \psi) := g_k(x, u) + \psi f_k(x, u).$$

3.4.2 Връзка на Хамилтоновата с целевата функция

За представянето на резултатите в този раздел, ще въведем следното определение:

Определение 3.3 (Производна по посока на допирателен конус на Булиган $T_S(\bar{y})$)**.** *Нека Y е банахово пространство, S е непразно затворено подмножество на Y, $\bar{y}$ е произволна точка от S и w е произволен елемент от допирателния конус на Булиган $T_S(\bar{y})$, въведен от Опредление 3.2. Казваме, че функцията $h : S \to \mathbb{R}$ е диференцируема по посока w, ако следната граница съществува:*

$$\lim_{t^\mu \downarrow 0} \frac{h(\bar{y} + t^\mu w^\mu) - h(\bar{y})}{t^\mu},$$

където $\{t^\mu\}_{\mu=0}^{\infty} \downarrow 0$, редицата $\{w^\mu\}_{\mu=0}^{\infty} \to w$ при $\mu \to +\infty$, и $\bar{y} + t^\mu w^\mu \in S$ за всеки $\mu = 1, 2, \ldots$. Означаваме тази граница с $dh(\bar{y}; w)$ и я наричаме производна на h по посока $w \in T_S(\bar{y})$ в точката $\bar{y} \in S$ *(виж [5]).*

Бележка 3.4. *Дефиницията 3.3 е еквивалентна на дефиницията на условна производна за случая на еднозначно изображение (виж, [5]).*

Следното твърдение осигурява връзка между съответните за задачата функция на Хамилтон $\mathcal{H}$ и целевата функция J, подчертавайки как вариациите в управлението влияят върху целевата функция чрез Хамилтониана. Твърдението по същество казва, че производната по посока на $T_{\mathbf{U}}(\bar{\mathbf{u}})$ на критерия J, се получава като сумираме скаларните произведения на съответните компоненти на частната производна на Хамилтониана относно управлението и съответните елементи на вектора на смущения $\mathbf{p}$, където $\mathbf{p} \in T_{\mathbf{U}}(\bar{\mathbf{u}})$.

Твърдение 3.1 (Връзка $\mathcal{H} \leftrightarrow J$)**.** *Нека предположения 3.3 и 3.4 са изпълнени, нека двойката $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ е осъществим процес и нека съответната спрегната редица ψ е определена от (32). Тогава за всяко положително цяло число τ и индексно множество от цели числа $\mathcal{K} := \{k_1, k_2, ..., k_\tau\} \subset \mathbb{N}$ с $0 \leq k_1 < k_2 < \cdots < k_\tau$ е изпълнено следното равенство:*

$$dJ(x_0, \bar{\mathbf{u}}; \mathbf{p}) = \sum_{i=1}^{\tau} \Big(\frac{\partial}{\partial u}\mathcal{H}_{k_i}(\bar{x}_{k_i}, \bar{u}_{k_i}, \psi_{k_i+1})\Big) p_{k_i},$$

за всеки вектор $\mathbf{p} := (p_0, p_1, p_2, ..., p_k, ...) \in T_{\mathbf{U}}(\bar{\mathbf{u}})$, където

$$\mathbf{U} := \Big\{(u_0, u_1, u_2, ..., u_k, ...) : u_s \in U_s, \text{ ако } s \in \mathcal{K}; \quad u_s = \bar{u}_s, \text{ ако } s \notin \mathcal{K}\Big\},$$

където за всяко $s \in \mathbb{N}$

$$U_s := \begin{cases} \{\bar{u}_s\}, & \text{ако } s \notin \mathcal{K}, \\ U_{k_i}, & \text{ако } s = k_i \quad \text{за някои } i \in \{1, 2, ..., \tau\}. \end{cases}$$

3.4.3 Достатъчно условие за оптималност

Теоремата, която следва, представя основния резултат в този раздел. Тя е в съответствие с принципите на теорията за оптимално управление, където Хамилтонианът е значим за характеризацията на оптималните траектории на състояние и управление. Условията, описани в теоремата, като изпъкналост и ограниченост на критерия и условие за минимум, свързано с функциятата на Хамилтон са интегрални. При тяхното удовлетворяване те колективно потвърждават, че процесът $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ минимизира критерия J, осигурявайки решение на поставената задача.

Теорема 3.3. *Нека предположения 3.3 и 3.4 са изпълнени, нека двойката $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ е осъществим процес и нека съответната спрегната редица $\psi = \{\psi_k\}_{k=1}^{\infty}$ се определя от (32). Нека критерият J е изпъкнал и съществуват константи $c, r > 0$, така че $J(x_0, \mathbf{u}) \leq c$ за всички $\mathbf{u} \in r\bar{\mathbf{B}}(\bar{\mathbf{u}})$ и нека следното условие за минимум е изпълнено*

$$\mathcal{H}_k(\bar{x}_k, \bar{u}_k, \psi_{k+1}) \leq \mathcal{H}_k(\bar{x}_k, u, \psi_{k+1}), \quad u \in U_k \text{ за всяко } k \in \mathbb{N}. \tag{33}$$

Тогава $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ е глобално оптимален процес

Бележка 3.5. *Бихме искали да отбележим, че предположението за изпъкналост във формулировката на Теорема 3.3 не означава, че всички функции g_k, $k \in \mathbb{N}$, са изпъкнали. Наистина, нека разгледаме следния прост илюстративен пример:*

$$J(x_0, \mathbf{u}) = \Big(\frac{1}{2}x_0^2 - \frac{1}{2}u_0^2\Big) + \Big(\frac{1}{2}x_1^2 + \frac{1}{2}u_1^2\Big) \to \min,$$

предмет на

$$x_{k+1} = x_k + 2u_k, \quad x_0 \in \mathbb{R}, \quad u_k \in [-1, 1], \quad k \in \mathbb{N}.$$

Тук,

$$g_0(x_0, u_0) = \frac{1}{2}x_0^2 - \frac{1}{2}u_0^2 \quad \text{и} \quad g_1(x_1, u_1) = \frac{1}{2}x_1^2 + \frac{1}{2}u_1^2.$$

Ясно е, че g_0 не е изпъкнала функция. Въпреки това критерият J е изпъкнал по отношение на (u_0, u_1), защото

$$\begin{aligned} J(x_0, \mathbf{u}) &= \left(\frac{1}{2}x_0^2 - \frac{1}{2}u_0^2\right) + \left(\frac{1}{2}\left(x_0 + 2u_0\right)^2 + \frac{1}{2}u_1^2\right) \\ &= x_0^2 + 2x_0u_0 + \frac{3}{2}u_0^2 + \frac{1}{2}u_1^2. \end{aligned}$$

Авторска справка

Основни приноси на дисертационния труд през погледа на автора:

- Установени са достатъчни условия за съществуване на равновесие на Неш за линейно-квадратичната диференциална игра между двама душ и върху безкраен хоризонт, в случая, когато управлението на минимизиращия играч е ограничено.

- Демонстрирано е, че тези резултати са приложими в контекста на диференциалните игри на Щакелберг. Практическата приложимост на тези резултати е илюстрирана с пример на парична политика.

- За линейно-квадратичната игра с дискретно време и безкраен хоризонт, при която управлението на играча, стремящ се към минимизация, е отново ограничено са установени достатъчни условия за оптималност. Основополагащо за извеждането на тези условия е използването на полученият принцип за оптималност на Белман за игра.

- Получен е ”мин-макс” резултат за линейно-квадратичната игра без ограничения. Показано е, че задачата върху безкраен времеви хоризонт е еквивалентна на определена задача в рамките на краен времеви интервал.

- Получено е необходимо условие от тип принцип на максимума на Понтрягин за дискретна игра с нулева сума вурху безкраен времви хоризонт.

- Получено е ново достатъчно условие за оптималност за задачата на оптималното управление с дискретно време.

Апробация на дисертацията

Резултатите в дисертацията се съдържат в следните статии:

Публикувани статии:

M. I. Krastanov, R. Rozenov, and B. K. Stefanov, *On a constrained infinite-time horizon linear quadratic game*, Dynamic Games and Applications, Volume 13, Issue 3, Birkhäuser, 2023

M. I. Krastanov, B. K. Stefanov, *On Decision Making under Uncertainty*, Lecture Notes in Computer Science, 209-220, ISSN 0302-9743, Springer, 2023

Статии приети за публикуване:

M. I. Krastanov and B. K. Stefanov, *A Sufficient Condition for a Discrete-Time Optimal Control Problem*, Lecture Notes in Computer Science, Springer

Статии предадени за публикуване:

M. I. Krastanov, R. Rozenov, and B. K. Stefanov, *On a Linear-Quadratic Game with Constrained Control over a Discrete Infinite-Time Horizon*, SIAM Journal on Control and Optimization

Изнесени доклади на научни форуми:

A Linear-Quadratic Differential Game,
Spring Scientific Session, Faculty of Mathematics and Informatics, Sofia University, 27 March 2021 (based on a joint work with M. Krastanov and R. Rozenov)

A Constrained Linear-Quadratic Differential Game,
Seminar on Optimization, Faculty of Mathematics and Informatics, Sofia University, 12 April 2021 (based on a joint work with M. Krastanov and R. Rozenov)

A Linear-Quadratic Game,
13th International Conference on Large-Scale Scientific Computations, LSSC 2021, June 7-11, 2021, Sozopol, Bulgaria, (based on a joint work with M. Krastanov and R. Rozenov)

On a Constrained Infinite-Time Horizon Linear-Quadratic Game,
Reporting Session, Conference on New Scalable Algorithms and Applications, 1 December, 2021, Sofia, Bulgaria (based on a joint work with M. Krastanov and R. Rozenov)

On a Infinite Time-Horizon Linear-Quadratic Game with Constrained Control,
Seminar on Optimization, Faculty of Mathematics and Informatics, Sofia University, 7 March

2022 (based on a joint work with M. Krastanov and R. Rozenov)

On a infinite time-horizon linear-quadratic game with constrained control, Spring Scientific Session, Faculty of Mathematics and Informatics, Sofia University, 26 March 2022 (based on a joint work with M. Krastanov and R. Rozenov)

On the Decision Making Under Uncertainty, 10th International Conference on Numerical Methods and Applications, Borovetz, August 22-26 2022,(based on a joint work with M. Krastanov)

A necessary condition for the existence of a Nash equilibrium for a discrete controllable system, Seminar on Optimization, Faculty of Mathematics and Informatics, Sofia University, 31 October 2022 (based on a joint work with M. Krastanov)

A Sufficient Condition for Optimality for a Discrete-Time Optimal Control Problem, Spring Scientific Session, Faculty of Mathematics and Informatics, Sofia University, 25 March 2023 (based on a joint work with M. Krastanov)

A Sufficient Condition for the Existence of a Nash Equilibrium, 14th International Conference on Large-Scale Scientific Computations LSSC 2023, June 5-9, 2023, Sozopol, Bulgaria, (based on a joint work with M. Krastanov)

A Sufficient Optimality Condition for a Constrained Infinite-Time Horizon LQ Game , 16th International Workshop on Well-Posedness of Optimization Problems and Related Topics , July 3-7, 2023, Borovets, Bulgaria, (based on a joint work with M. Krastanov and R. Rozenov)

On an Infinite-Time Horizon Linear-Quadratic Game with Constrained Control, Seminar Operations Research, Probability and Statistics, IMI-BAS, 17 October 2023 (based on a joint work with M. Krastanov and R. Rozenov)

A Sufficient Condition for a Discrete-Time Optimal Control Problem, Reporting Session, Department Operations Research, Probability and Statistics, IMI-BAS, 19 December 2023 (based on a joint work with M. Krastanov)

Декларация за автентичност

- Като автор на тази дисертация, декларирам, че представените в нея резултати са плод на моята собствена работа или в сътрудничество с моите съавтори.
- Когато съм се консултирал с публикуваната работа на други, това винаги е ясно указано.
- Където съм цитирал от работата на други, източникът винаги е посочен.

Боян Стефанов

Литература

[1] A. Al-Tamimi, F. L. Lewis, and M. Abu-Khala, ”Model-free Q-learning designs for linear discrete-time zero-sum games with application to H-infinity control,” Automatica, vol. 43, 2007.

[2] P. J. Antsaklis and A. N. Michel, ”Linear Systems,” Birkhäuser, 2006.

[3] S. M. Aseev, M. I. Krastanov, and V. M. Veliov, ”Optimality Conditions for Discrete-Time Optimal Control on Infinite Horizon,” Pure Appl. Funct. Anal., Nov 2017, pp. 395-409.

[4] J. P. Aubin and H. Frankowska, ”Set-Valued Analysis,” Birkhäuser, May 2009.

[5] J. P. Aubin, ”Lipschitz Behavior of Solutions to Convex Minimization Problems,” Math. Oper. Res., vol. 9, no. 1, Feb 1984, pp. 87-111.

[6] T. Başar and P. Bernhard, ”H^{∞}- Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach,” Birkhäuser, Sep 1991.

[7] T. Başar and G. J. Olsder, ”Dynamic Noncooperative Game Theory,” SIAM, Oct 1999.

[8] R. Bellman, ”Dynamic Programming,” Princeton University Press, 1957.

[9] A. Ben-Tal and A. Nemirovski, ”Robust Optimization,” Princeton University Press, Princeton, NJ, 2009.

[10] D. P. Bertsekas, ”Dynamic Programming and Optimal Control 4th Edition,” Athena Scientific, Belmont, Massachusetts, May 2017.

[11] D. Bertsimas and M. Sim, ”The Price of Robustness,” Operations Research, vol. 52, no. 1, 2004, pp. 35–53.

[12] V. G. Boltyanskij, ”Optimal Control of Discrete Systems,” Halsted Press, 1978.

[13] J. F. Bonnans and A. Shapiro, ”Perturbation Analysis of Optimization Problems,” Springer Ser. Oper. Res., Jul 2000.

[14] A. Borzi and F. C. Campana, ”On the SQH Method for Solving Differential Nash Games,” J. Dyn. Control Syst., May 2021, pp. 739-735.

[15] B. M. Chen, T. I. Baskin, and B. Hu, ”Robust Flight Control: A Design Challenge,” Journal of Guidance, Control, and Dynamics, vol. 30, no. 1, 2007.

[16] C.L. Chen and J. Cruz, ”Stackelburg Solution for Two-Person Games with Biased Information Patterns,” IEEE Transactions on Automatic Control, 1972.

[17] J. H. Cochrane, ”The New Keynesian Liquidity Trap,” J. Monet. Econ., vol. 92, no. 1, pp. 47-63, 2017.

[18] J. Dennis, K. Leitemo, and U. Soderstrom, ”Methods for Robust Control,” J. Econ. Dyn. Control, Mar 2009, pp. 1604-1616.

[19] J. L. Doob, ”Stochastic Processes,” Wiley, New York, 1953.

[20] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, ”Feedback Control Theory,” Dover Publications, 2009.

[21] E. Dockner, S. Jorgensen, N. Van Long, and G. Sorger, ”Differential Games in Economics and Management Science,” Cambridge University Press, Jun 2012.

[22] W. Durham and D. A. Lawrence, ”Survey of Robust Control for Unmanned Aircraft,” Journal of Guidance, Control, and Dynamics, vol. 30, no. 2, 2007.

[23] J. Engwerda, “LQ Dynamic Optimization and Differential Games”. In: John Wiley and Sons Ltd. 2005

[24] Sir R. A. Fisher, ”Statistical Methods for Research Workers,” Oliver and Boyd, Edinburgh, 1925.

[25] D. Fudenberg and J. Tirole, ”Game Theory,” The MIT Press, 1991.

[26] J. P. Geering, ”Optimal Control with Engineering Applications,” Springer-Verlag, Berlin Heidelberg, Mar 2000.

[27] M. Giannoni, ”Robust Optimal Monetary Policy in a Forward-Looking Model with Parameter and Shock Uncertainty,” J. Appl. Econometrics, vol. 22, pp. 179-213, 2007.

[28] V. Y. Glizer and O. Kelis, ”Singular Linear-Quadratic Zero-Sum Differential Games and H^{∞}- Control Problems,” Birkhäuser, Apr 2007.

[29] R. Goebel and M. Subbotin, ”Continuous Time Linear Quadratic Regulator With Control Constraints via Convex Duality,” IEEE Trans. Autom. Control, vol. 52, no. 5, pp. 886-892, 2007.

[30] L. Hansen and T. Sargent, ”Robustness,” Princeton University Press, 2008.

[31] G. E. Ivanov, ”Saddle Point for Differential Games With Strongly Convex-Concave Integrand,” Math. Notes, vol. 62, Nov 1997, pp. 607-622.

[32] A. Kolmogorov, ”Foundations of the Theory of Probability,” Chelsea Publishing Company, New York, 2nd edition, 1956.

[33] I. A. Krylov and F. L. Chernousko, ”On the method of successive approximations for solution of optimal control problems,” USSR Comput. Math. Math. Phys., vol. 2, no. 6, 1963, pp. 1371-1382.

[34] P. S. Laplace, ”Théorie Analytique des Probabilités,” 1812.

[35] A.J. Laub, ”A Schur Method for Solving Algebraic Riccati Equations”, IEEE Transactions on Automatic Control, volume 24, number 6, pp 913–921, 1979.

[36] F. L. Lewis and B. L. Stevens, ”Aircraft Control and Simulation,” John Wiley and Sons, Inc, 2016.

[37] A. A. Lyubushin, ”Modifications of the method of successive approximations for solving optimal control problems,” USSR Comput. Math. Math. Phys., vol. 22, no. 1, 1982, pp. 29-34.

[38] A. de Moivre, ”The Doctrine of Chances,” 1718.

[39] J. F. Nash, ”Non-Cooperative Games,” Annals of Mathematics, vol. 54, no. 2, 1950, pp. 286–295.

[40] J. von Neumann and O. Morgenstern, ”Theory of Games and Economic Behavior”, Princeton University Press, 1944.

[41] J. Neyman, ”On the Two Different Aspects of the Representative Method,” Journal of the Royal Statistical Society, vol. 97, 1934, pp. 558-625.

[42] A. Onatski and N. Williams, ”Modeling Model Uncertainty,” J. Eur. Econ. Assoc., vol. 1, no. 5, Jan 2003, pp. 1087-1122.

[43] M. J. Osborne and A. Rubinstein, ”A Course in Game Theory,” The MIT Press, 1994.

[44] I. R. Petersen, V. A. Ugrinovskii, and A. V. Savkin, ”Robust Control Design: An Optimal Control Approach,” Springer, 2003.

[45] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, ”The Mathematical Theory of Optimal Processes,” Interscience Publishers, 1962.

[46] R. Rozenov, ”Optimal Fiscal Adjustment under Uncertainty, International Monetary Fund,” Int. Monetary Fund, Mar 2016.

[47] S. J. Rubio, ”On Coincidence of Feedback Nash Equilibria and Stackelberg Equilibria in Economic Applications of Differential Games”, J. Optim. Theory Appl., vol. 128, no. 1, Jan 2006, pp. 203-220.

[48] W. Rudin, ”Principles of Mathematical Analysis”, McGraw-Hill Science/Engineering/Math, 1976.

[49] C. E. Shannon, ”A Mathematical Theory of Communication,” Bell System Technical Journal, vol. 27, no. 3, 1948, pp. 379-423.

[50] H. E. Taylor and S. Karlin, ”A First Course in Stochastic Processes,” Academic Press, New York, 1975.

[51] A. T. Tran, ”Nonlinear Optimal Flight Control Design under Various Constraints,” PhD thesis, 2018.

[52] I. Werning, ”Managing a Liquidity Trap: Monetary and Fiscal Policy,” MIT Working Paper, Cambridge, MA, vol. 128, Mar 2012.

[53] B. M. Wilamowski and J. D. Irwin, ”Industrial Electronics Handbook,” CRC Press, 2011.

[54] R. J. Williams, ”Sufficient Conditions for Nash Equilibria in N-Person Games over Reflexive Banach Spaces,” J. Optim. Theory Appl., vol. 30, no. 3, Mar 1980, pp. 383-394.

[55] J. C. Willems and H. Blom, ”Mathematical Control Theory: Deterministic Finite Dimensional Systems,” Springer, Berlin, 1976.

[56] M. Woodford, ”Optimal Interest-Rate Smoothing,” Rev. Econ. Stud., vol. 70, Oct 2003, pp. 861-886.

[57] L. A. Zadeh, ”Fuzzy Sets,” Information and Control, vol. 8, 1965, pp. 338-353.

[58] K. Zhou and J. C. Doyle, ”Essentials of Robust Control,” Prentice Hall, 1997.

[59] B. Øksendal, ”Stochastic Differential Equations: An Introduction with Applications,” 6th edition, Springer, 2003.