Основи диференціальних ігор

О. Галганов, Н. Фордуй

Історія виникнення

Класична теорія ігор, творцями якої є Джон фон Нейман та Оскар Морґенштерн, оперує матричними іграми та зв'язаними з ними поняттями. Такі ігри зовсім не покривають всі можливості формалізувати процеси, в яких гравці приймають рішення. Наприклад, якщо розглядати задачу про оптимальну траєкторію руху корабля від керованої торпеди, яка його переслідує, то необхідно досліджувати не просто скінченний набір можливих дій кожного з гравців, а континуум можливих стратегій, кожна з яких відповідає деякій комбінації траєкторій руху корабля та торпеди.

Сам термін «диференціальна гра» було введено Руфусом Айзексом — одним з основоположників теорії диференціальних ігор.

Історія виникнення

Джон фон Нейман

Оскар Морґенштерн

Руфус Айзекс

Основні поняття

Рішення, що їх приймають гравці, полягають у виборі так званих **керувань**, від яких залежать **фазові координати**: їх значення у будь-який момент часу повністю визначає хід гри, характеризуючи положення гравців у деякому просторі — **фазовому просторі**.
Поточні значення фазових координат завжди відомі гравцям — тобто, це ігри з повною інформацією. Невідомим зазвичай є характер їх зміни: тобто, **керування** фазовими змінними гравцями.

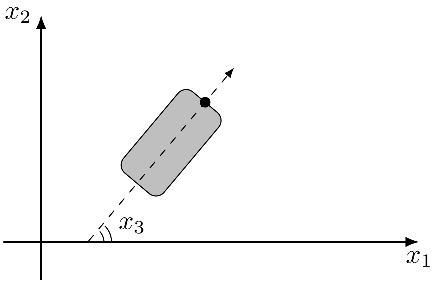
Приклад

Положення матеріальної точки на площині описується двома координатами x_1 та x_2. Нехай швидкість руху точки є сталою v, а гравець обирає напрямок швидкості φ та може змінювати його у будь-який момент часу — тобто, φ є керуванням. Тоді рух точки описується системою диференціальних рівнянь

$$\begin{cases} \dot{x}_1 = v\cos\varphi \\ \dot{x}_2 = v\sin\varphi \end{cases}$$

Приклад

Геометричне положення автомобіля на декартовій площині описується трьома фазовими координатами: x_1, x_2 — положення деякої точки автомобіля, x_3 — кут, який утворює вісь вздовж автомобіля з деяким фіксованим напрямком — наприклад, x_1.



Продовження приклада

Нехай A — максимальне можливе прискорення автомобіля, тоді прискорення може набувати значень $A\varphi_1$, де $\varphi_1 \in [0;1]$ і знаходиться під контролем гравця-водія. Можна ввести ще одну фазову координату x_4 — швидкість автомобіля. Як ще одну фазову координату x_5 можна ввести кривину (фізично — це кут повороту передніх коліс), керуванням якої є $W\varphi_2$, де $\varphi_2 \in [-1;1]$, а W — максимальна швидкість зміни x_5.
Система задає рух автомобіля у деякій диференціальній грі:

$$\begin{cases} \dot{x_1} = x_4 \cos x_3 \\ \dot{x_2} = x_4 \sin x_3 \\ \dot{x_3} = x_4 x_5 \\ \dot{x_4} = A\varphi_1, \ \varphi_1 \in [0;1] \\ \dot{x_5} = W\varphi_2, \ \varphi_2 \in [-1;1] \end{cases}$$

Опис руху

Вважаємо, що гра відбувається у *фазовому просторі* $\mathcal{E}$ — деякій області в $\mathbb{R}^n$ та на її межі. Рух точки $x = (x_1, x_2, ..., x_n)$ у фазовому просторі описується системою диференціальних рівнянь

$$\begin{cases} \dot{x_1}(t) = f_1(x_1(t), ..., x_n(t), u_1(x,t), ..., u_P(x,t), v_1(x,t), ..., w_E(x,t)) \\ \dot{x_2}(t) = f_2(x_1(t), ..., x_n(t), u_1(x,t), ..., u_P(x,t), v_1(x,t), ..., w_E(x,t)) \\ \dots \\ \dot{x_n}(t) = f_n(x_1(t), ..., x_n(t), u_1(x,t), ..., u_P(x,t), v_1(x,t), ..., w_E(x,t)) \\ x_1(0) = x_1^0, x_2(0) = x_2^0, ..., x_n(0) = x_n^0 \end{cases}$$

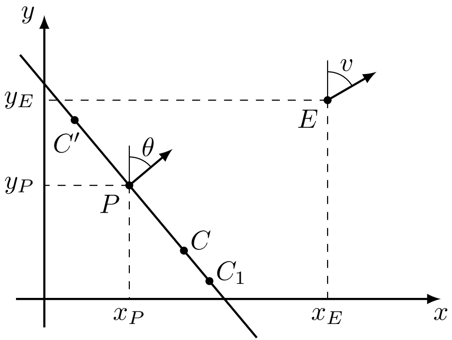
або, коротше,

$$\begin{cases} \dot{x}(t) = f(x(t), u(x,t), v(x,t)) \\ x(0) = x_0 \end{cases}$$

Ці рівняння називаються *рівняннями руху*. Функції f_j є заданими та вважаються достатньо гладкими.

Приклад (гра «водій-вбивця»)

Гра відбувається на площині. Переслідувач P рухається зі сталою швидкістю w_P, радіус кривини його траєкторії обмежений заданою величиною R. Керування P — це вибір значення кривини в кожний момент часу. Рух утікача E простий: швидкість w_E фіксована, керуванням є вибір напрямку швидкості $v(t)$. E є більш маневреним, ніж P. Фазовими координатами в цій грі є пари x_P, y_P і x_E, y_E для опису положення P та E відповідно, та θ — напрямок руху P.



Приклад (гра «водій-вбивця»)

Керування E — це вибір кута швидкості v. Керування P записати дещо складніше. Проведемо через точку (x_P, y_P) пряму $C'PC$, $|C'P| = |PC| = R$, перпендикулярну до вектору швидкості P. P обирає миттєвий центр кривини своєї траєкторії у довільній точці C_1 цієї прямої, що лежить за межами відрізку $C'C$ (оскільки радіус кривини обмежений). Керування $u(t)$ будемо вважати рівним за модулем $R/|PC_1|$, додатним для точок C_1, що знаходяться правіше від P, та від'ємним для тих, що знаходяться лівіше. Остаточно, маємо такі рівняння руху:

$$\begin{cases} \dot{x_P}(t) = w_P \sin\theta(t) \\ \dot{y_P}(t) = w_P \cos\theta(t) \\ \dot{x_E}(t) = w_E \sin v(t) \\ \dot{y_E}(t) = w_E \cos v(t) \\ \dot{\theta}(t) = \frac{w_P}{R} u(t), \ u(t) \in [-1; 1] \end{cases}$$

Виграші

Мета диференціальної гри визначається виграшем, який залежить від траєкторій гравців. Позначимо ці траєкторії як функції від часу як $x(t)$ та $y(t)$. Зауважимо, що диференціальні ігри є *антагоністичними* (або ж, *іграми з нульовою сумою*).
Якщо гра триває деякий заздалегідь визначений час T, то виграш гравця E визначається як $H(x(t), y(T))$, де $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ — деяка функція (нагадаємо, що розмірність $\mathcal{E}$ — n).

Приклади виграшів

1. $H(x(T), y(T)) = \|x(T) - y(T)\|$
2. $H(x(T), y(T)) = \min\limits_{0 \le t \le T} \|x(t) - y(t)\|$
3. $H(x(T), y(T)) = t_* = \min\{t \ge 0 : (x(t), y(t)) \in \mathcal{T}\}$

Приклад

Розглянемо переслідування на площині з простим рухом, що описується системою

$$\begin{cases} \dot{x_1} = u_1, \dot{x_2} = u_2,\ u_1^2 + u_2^2 \le \alpha^2 \\ \dot{y_1} = v_1, \dot{y_2} = v_2,\ v_1^2 + v_2^2 \le \beta^2 \\ x_1(0) = x_1^0, x_2(0) = x_2^0, y_1(0) = y_1^0, y_2(0) = y_2^0 \end{cases}$$

Якщо $\alpha > \beta$, то гравець P може гарантувати

$$\forall l \ge 0 : \min \{t \ge 0 : \|P(t) - E(t)\| \le l\} < +\infty$$

Якщо $\alpha \le \beta$, то в разі $\|P(0) - E(0)\| > l$ для всіх $l \ge 0$ гравець E, рухаючись від P по прямій з максимальною швидкістю, зможе уникнути захоплення гравцем P.

Поняття стратегії

Означення

Стратегіями у диференціальній грі є вибір керувань u та v як функцій від часу t та фазових координат x у системі рівнянь руху

$$\begin{cases} \dot{x}(t) = f(x(t), u(x,t), v(x,t)) \\ x(0) = x_0 \end{cases}$$

Керування вважаються кусково-гладкими як компроміс між забезпеченням існування розв'язку, (його може не існувати у класі неперервних функцій) та його єдиності (вона може порушуватися, якщо не вимагати неперервності розв'язку).
Позначатимемо через P та E множини кусково-неперервних стратегій (керувань) гравців P та E.

Ситуація

Надалі для спрощення розглядатимемо не один вектор x, а два вектори x та y, що відповідатимуть руху кожного з гравців. Тоді систему можна записати як

$$\begin{cases} \dot{x}(t) = f(x(t), u(x, y, t)) \\ \dot{y}(t) = g(x(t), v(x, y, t)) \\ x(0) = x_0, y(0) = y_0 \end{cases}$$

Означення

Набір $S = \{x_0, y_0, u(\cdot), v(\cdot)\}$, де x_0, y_0 — початкові умови, а $u \in \mathrm{P}$, $v \in \mathrm{E}$ — керування, називається *ситуацією* в диференціальній грі.

Умова існування та єдиності траекторій

Якщо розглядати траєкторії, що залежать лише від часу t та накладати на f та g умови обмеженості та ліпшицевості по x та y, тобто

$$\|f(x_1, u) - f(x_2, u)\| \leq \alpha \cdot \|x_1 - x_2\|,$$
$$\|g(y_1, v) - g(y_2, v)\| \leq \beta \cdot \|y_1 - y_2\|,$$

то за теоремою про існування та єдиність роз'язку задачі Коші, для кожної ситуації S буде існувати єдина пара траєкторій $x(t), y(t)$, для якої

$$\begin{cases} \dot{x}(t) = f(x(t), u(t)) \\ \dot{y}(t) = g(y(t), v(t)) \\ x(0) = x_0, y(0) = y_0 \end{cases}$$

Формальне визначення виграшів та їх види

Означення

Користуючись означенням ситуації, можна ввести *виграш* в ситуації $S = \{x_0, y_0, u(\cdot), v(\cdot)\}$ як функцію $K(x_0, y_0, u(\cdot), v(\cdot))$.

Наведемо строгі означення чотирьох видів виграшів.

Означення

1. *Термінальний виграш.* Задано деяке число $t > 0$ та неперервна по x та y функція $H(x, y)$. Виграш в ситуації $\{x_0, y_0, u(\cdot), v(\cdot)\}$ визначається як:

$$K(x_0, y_0, u(\cdot), v(\cdot)) = H(x(T), y(T))$$

Означення

2. *Мінімальний результат*. Задано деяке число $t > 0$ та неперервна по x та y функція $H(x, y)$. Виграш в ситуації $\{x_0, y_0, u(\cdot), v(\cdot)\}$ визначається як

$$K(x_0, y_0, u(\cdot), v(\cdot)) = \min_{0 \le t \le T} H(x(t), y(t))$$

3. *Інтегральний виграш*. Нехай $\mathcal{T}$ — деяка підмножина $\mathbb{R}^n \times \mathbb{R}^n$, $H(x, y)$ — неперервна функція. Нехай в ситуації $\{x_0, y_0, u(\cdot), v(\cdot)\}$ t_* — перший момент потрапляння траєкторії $(x(t), y(t))$ на $\mathcal{T}$. Тоді

$$K(x_0, y_0, u(\cdot), v(\cdot)) = \int_0^{t_*} H(x(t), y(t))dt$$

де при $t_* = +\infty$ покладається $K = +\infty$.

Види виграшів

Означення

4 *Якісний виграш*. Нехай $\mathcal{T}$ та $\mathcal{L}$ — деякі підмножини $\mathbb{R}^n \times \mathbb{R}^n$, а t_* — перший момент потрапляння траєкторії $(x(t), y(t))$ на $\mathcal{T}$ в ситуації $\{x_0, y_0, u(\cdot), v(\cdot)\}$. Тоді

$$K(x_0, y_0, u(\cdot), v(\cdot)) = \begin{cases} 1, & \text{якщо } (x(t_*), y(t_*)) \in \mathcal{L} \\ 0, & \text{якщо } t_* = +\infty \\ -1, & \text{якщо } (x(t_*), y(t_*)) \notin \mathcal{L} \end{cases}$$

Нормальна форма диференціальної гри

Нарешті, можна дати означення нормальної форми диференціальної гри.

Означення

Нормальною формою диференціальної гри $\Gamma(x_0, y_0)$, заданої на просторі стратегій $\mathrm{P} \times \mathrm{E}$, називається система

$$\Gamma(x_0, y_0) = \langle x_0, y_0, \mathrm{P}, \mathrm{E}, K(x_0, y_0, u(\cdot), v(\cdot)) \rangle$$

де $K(x_0, y_0, u(\cdot), v(\cdot))$ — функція виграшу, визначена будь-який з чотирьох способів вище.

Кожній парі $(x_0, y_0) \in \mathbb{R}^n \times \mathbb{R}^n$ відповідає своя гра в нормальній формі, тобто, фактично, визначається двопараметрична сім'я ігор, що залежать від (x_0, y_0).

Простий рух на площині

Розглянемо найпростіші моделі задач переслідування — диференціальні ігри на площині з двома учасниками: переслідувачем P та утікачем E, траєкторії яких відповідно позначатимемо $x(t)$ та $y(t)$. Під *простим рухом* мається на увазі, що закони їх руху описуються системою

$$\begin{cases} \dot{x} = u, & \|u\| \leq \alpha \\ \dot{y} = v, & \|v\| \leq \beta \end{cases}$$

Тут $\|z\| = \sqrt{z_1^2 + z_2^2}$. Такі закони руху означають, що гравці рухаються з обмеженою швидкістю, але напрямок руху можуть змінювати довільно. Проінтегрувавши рівняння, можна явно записати траєкторії руху як

$$x(t) = x(0) + \int\limits_0^t u(s)ds, \; y(t) = y(0) + \int\limits_0^t v(s)ds$$

Приклад

Умова

Нехай $u(t) = \begin{pmatrix} -\sin t \\ 2\cos 2t \end{pmatrix}$, $v(t) = \begin{pmatrix} -\sqrt{2}\sin t \\ \sqrt{2}\cos t \end{pmatrix}$, $x(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $y(0) = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}$, а гра триває до моменту $T = 2\pi$. Знайти значення функції виграшу мінімального результату з $H(x, y) = \|x - y\|$.

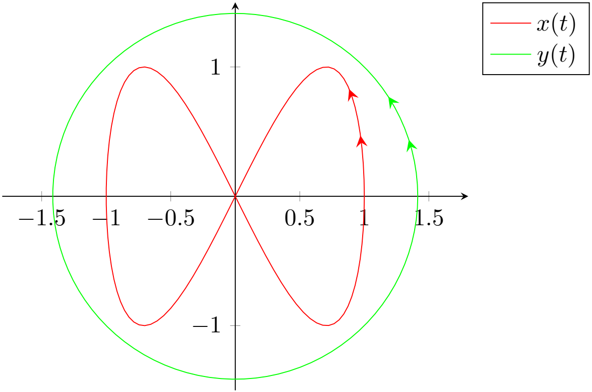
Знайдемо рівняння траєкторій:

$$x(t) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_0^t \begin{pmatrix} -\sin s \\ 2\cos 2s \end{pmatrix} ds = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} \cos s \\ \sin 2s \end{pmatrix} \Bigg|_0^t = \begin{pmatrix} \cos t \\ \sin 2t \end{pmatrix}$$

$$y(t) = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix} + \int_0^t \begin{pmatrix} -\sqrt{2}\sin s \\ \sqrt{2}\cos s \end{pmatrix} ds = \begin{pmatrix} \sqrt{2}\cos t \\ \sqrt{2}\sin t \end{pmatrix}$$

Приклад

На декартовій площині ці траєкторії матимуть вигляд:



Значення $K = \min\limits_{0 \le t \le 2\pi} \|x(t) - y(t)\|$ можна знайти чисельно: $K \approx 0.282394$ при $t \approx 0.850448$.

Простий рух в $\mathbb{R}^n$

Тепер розглянемо гру переслідування вже не на площині $\mathbb{R}^2$, а в $\mathbb{R}^n$:

$$\begin{cases} \dot{x} = u, & \|u\| \le \alpha \\ \dot{y} = v, & \|v\| \le \beta \\ x(0) = x_0, \ y(0) = y_0 \end{cases}$$

Тут усі величини є n-вимірними векторами, і, як раніше, $x(t)$ — траєкторія руху переслідувача P, $y(t)$ — утікача E.

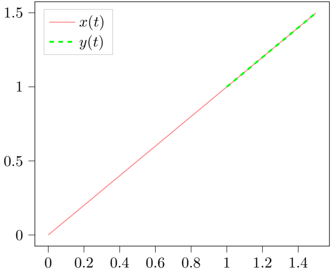
Нехай $\alpha > \beta$, тобто, переслідувач може рухатися швидше за утікача. Тоді можна довести, що яку б стратегію $v(t)$ не обрав утікач E, переслідувач P наздожене його не пізніше, ніж за $\frac{\|x_0 - y_0\|}{\alpha - \beta}$, використовуючи стратегію $u(t) = -\frac{\alpha}{\|x(t)-y(t)\|}(x(t) - y(t))$, причому переслідування буде найдовшим, якщо E обере «раціональну» стратегію $v(t) = -\frac{\beta}{\|x(t)-y(t)\|}(x(t) - y(t))$.

Приклад

Умова

Нехай $\alpha = 3, \beta = 1$, гра починається з $x_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ та $y_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, обидва гравці обирають керування, вказані вище.

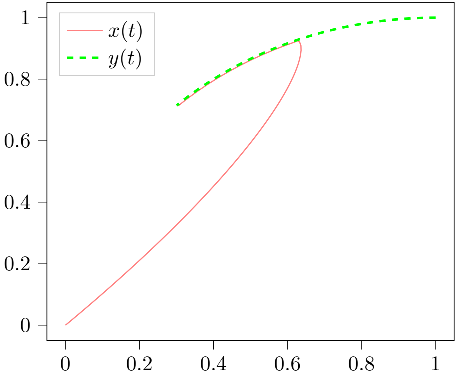
Розв'язавши чисельно (методом Рунге-Кутта) відповідну систему диференціальних рівнянь, отримаємо такі траєкторії руху:



P наздогнав E у момент часу $T \approx 0.7071$. За отриманою формулою точним значенням T є $\sqrt{2}/2$.

Приклад

Якщо ж утікач обиратиме «нераціональне» керування, наприклад, $\dot{y} = -\beta \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$, то отримаємо такі траєкторії руху:



В такому випадку P наздожене E в момент $T \approx 0.385$.

Лінійна диференціальна гра

У випадку простого руху в $\mathbb{R}^n$ замість системи диференціальних рівнянь, що описують рух переслідувача та утікача розглядалося одне диференціальне рівняння, що описувало динаміку різниці їх положень.

Означення

Лінійною диференціальною грою називається гра з фазовим простором $\mathbb{R}^n$, що описується рівнянням

$$\begin{cases} \dot{z}(t) = Az(t) - u(t) + v(t) \\ z(0) = z_0 \\ u \in U, \; v \in V \end{cases}$$

де A — деяка стала матриця порядку $n \times n$, U та V — опуклі компактні підмножини $\mathbb{R}^n$. Також задано матрицю π, що є матрицею проекції на ортогональне доповнення до термінальної множини.

Лінійна диференціальна гра

Ця гра відбувається таким чином: в кожний момент часу t утікач E знає параметри гри (A, U, V, z_0, π) та обирає своє керування $v(t) \in V$, повідомляючи про свій вибір переслідувача P, який, в свою чергу, обирає керування $u(t) \in U$. Якщо існує такий момент часу $T > 0$, коли переслідувач P за будь-яких дій утікача E забезпечує виконання умови $\pi z(\tau) = 0$ для деякого $\tau \in [0; T]$, то кажуть, що в переслідувач наздоганяє утікача. Наведемо умови, за яких це відбувається. Для цього треба ввести декілька нових означень.

Операції над множинами

Означення 1

Сумою множин (за Мінковським) A і B називається множина $C = A + B = \{a + b : a \in A, b \in B\}$.

Означення 2

Різницею множин (за Мінковським) A і B називається найбільша така множина $C = A \dot{-} B$, що $B + C \subset A$

Означення 3

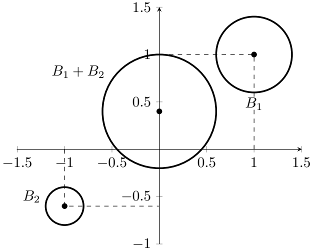
Добутком множини A на число $\lambda \in \mathbb{R}$ називається множина $\lambda \cdot A = \{\lambda \cdot a : a \in A\}$.

Приклад

Нехай $B_r(a) = \{x \in \mathbb{R}^n : \|x - a\| \le r\}$ — куля радіуса r з центром в точці a. Для $r \in \mathbb{R}$ та $a \in \mathbb{R}^n$ має місце $r \cdot B_1(0) + \{a\} = B_{|r|}(a)$. Сумою двох куль $B_{r_1}(a_1)$ та $B_{r_2}(a_2)$ є множина

$$M = B_{r_1}(a_1) + B_{r_2}(a_2) = \{x_1 + x_2 : \|x_1 - a_1\| \le r_1, \|x_2 - a_2\| \le r_2\}$$

Для $x = x_1 + x_2 \in M$: $\|(x_1 + x_2) - (a_1 + a_2)\| \le \|x_1 - a_1\| + \|x_2 - a_2\| \le r_1 + r_2$, тобто $M = B_{r_1 + r_2}(a_1 + a_2)$.



Інтеграл багатозначного відображення

Означення

Нехай $W(t)$ — неперервна функція з дійсним аргументом, значеннями якої є компактні підмножини $\mathbb{R}^n$ (*багатозначне відображення*). *Інтегралом* за проміжком $[a;b]$ від неї називається множина $\int\limits_a^b W(t)dt$, яку можна розуміти в сенсі рімановї суми $\lim\limits_{n\to\infty} \sum\limits_{i=0}^{n} \Delta t_i \cdot W(t_i^*)$, де $\{\Delta t_i\}_{i=1}^{n}$ — довжини відрізків, на які розбивається $[a;b]$, $t_i^* \in \Delta t_i$ — деякі точки з цих відрізків, а сума розуміється в сенсі суми множин за Мінковським.

Інтеграл багатозначного відображення, приклад

Умова

Нехай $W(t) = B_t(0)$, $[a;b] = [0;T]$. Знайдемо $\int\limits_0^T W(t)dt$.

$\sum\limits_{i=0}^{n} \Delta t_i \cdot W(t_i^*) = \sum\limits_{i=0}^{n} \Delta t_i \cdot B_{t_i^*}(0) = \sum\limits_{i=0}^{n} B_{t_i^* \Delta t_i}(0) = B_{\sum\limits_{i=0}^{n} t_i^* \Delta t_i}(0)$ — це куля з центром в 0 та радіусом $\sum\limits_{i=0}^{n} t_i^* \Delta t_i$. Вираз для радіуса є інтегральною сумою для $\int\limits_0^t t dt = \frac{T^2}{2}$, тому $\int\limits_0^T B_t(0)dt = B_{\frac{T^2}{2}}(0)$.

Умови того, що P наздожене E

Нехай у лінійній диференціальній грі виконуються дві умови:

1. Для всіх $t > 0$: $W(t) = \pi e^{At}U \buildrel{*}\over{-} \pi e^{At}V \neq \varnothing$.
2. Існує такий момент часу T_0, що $\pi e^{AT_0} z_0 \in \int\limits_0^T W(T_0 - s)ds$.

Можна довести, що в разі виконання цих умов переслідувач наздожене утікача.

Наведемо алгоритм застосування методу Понтрягіна:

1. Знайти множину $W(t) = \pi e^{At} U \dot{-} \pi e^{At} V$.
2. Знайти множину $\Omega(t) = \int\limits_0^t W(s)ds$.
3. Знайти T_0, для якого $\pi e^{AT_0} z_0 \in \Omega(T_0)$.
4. Знайти функцію $w(t) \in W(t)$ таку, що $\pi e^{AT_0} z_0 = \int\limits_0^{T_0} w(s)ds$.
5. Знайти керування $u(t)$ як розв'язок $\pi e^{A(T_0-s)} u(s) - \pi e^{A(T_0-s)} v(s) = w(T_0 - s)$ при заданому керуванні $v(t) \in V$.
6. Знайти розв'язок задачі Коші $\dot{z} = Az - u(t) + v(t),\ z(0) = z_0$ на відрізку $[0; T_0]$.

Метод Понтрягіна у грі з простим рухом

Умова

Розглянемо гру з простим рухом:

$$\begin{cases} \dot{x} = u, & \|u\| \leq \alpha \\ \dot{y} = v, & \|v\| \leq \beta \end{cases}$$

Переслідування закінчується, коли $x - y = 0$.

Зробимо заміну $z = x - y$: $\dot{z} = u - v = -(-u) + (-v)$, переслідування закінчується, коли $z = x - y = 0$. В такому випадку оператор π є тотожнім, матриця A — нульовою, тому й композиція $\pi e^{At} = I$ — теж тотожній оператор. $U = B_\alpha(0)$, $V = B_\beta(0)$ — кулі з центрами в 0 та з радіусами α та β відповідно. Таким чином, $W(t) = \pi e^{At} U \dot{-} \pi e^{At} V = U \dot{-} V = B_{\alpha-\beta}(0)$, при $\alpha \geq \beta$ ця множина є непорожньою.

Метод Понтрягіна у грі з простим рухом

$\Omega(t) = \int\limits_0^t W(s)ds = B_{t(\alpha-\beta)}(0)$, тому

$$\pi e^{AT_0} z_0 \in \Omega(T_0) \Leftrightarrow z_0 \in B_{T_0(\alpha-\beta)}(0) \Leftrightarrow$$
$$\Leftrightarrow \|z_0\| = T_0(\alpha-\beta) \Leftrightarrow T_0 = \frac{\|z_0\|}{\alpha-\beta}$$

Зауважимо, що такий само час було отримано раніше іншими міркуваннями.
Також можна знайти керування переслідувача $u(t)$ для заданого керування утікача $v(t)$ та отримати динаміку різниці координат гравців:

$$x(t) - y(t) = (\beta-\alpha)\cdot\frac{x_0-y_0}{\|x_0-y_0\|}t + (x_0-y_0)$$

Контрольний метод Понтрягіна

Умова

Нехай рух гравців в $\mathbb{R}^n$, $n \geq 2$, описується системою

$$\begin{cases} \ddot{x} + \alpha \dot{x} = a, & \|a\| \leq \rho \\ \ddot{y} + \beta \dot{y} = b, & \|b\| \leq \sigma \end{cases}$$

де $\alpha, \beta, \rho, \sigma$ — додатні числа. Переслідувач наздоганяє утікача, якщо $x = y$. Кожне рівняння системи описує рух точки одиничної маси під дією сили-керування з урахуванням тертя, що лінійно залежить від швидкості.

Перейдемо до системи диференціальних рівнянь першого порядку за допомогою замін $z^1 = x - y$, $z^2 = \dot{x}$, $z^3 = \dot{y}$:

$$\begin{cases} \dot{z}^1 = z^2 - z^3 \\ \dot{z}^2 = -\alpha z^2 + a \\ \dot{z}^3 = -\beta z^3 + b \end{cases}$$

Контрольний приклад Понтрягіна

Керування u та v задаються формулами $u = (0, -a, 0)^T$, $v = (0, 0, b)^T$, тому $U = \left\{(0, -a, 0)^T : \|a\| \le \rho\right\}$, $V = \left\{(0, 0, b)^T : \|b\| \le \sigma\right\}$. Оператор π задано як $\pi : (z^1, z^2, z^3)^T \mapsto (z^1, 0, 0)^T$, а матриця A дорівнює $\begin{pmatrix} 0 & 1 & -1 \\ 0 & -\alpha & 0 \\ 0 & 0 & -\beta \end{pmatrix}$.

Знайдемо e^{At} (за допомогою перетворення Лапласа):

$$e^{At} = \mathcal{L}\left\{(pI - A)^{-1}\right\} \Leftrightarrow (pI - A)^{-1} = \mathcal{L}^{-1}\left\{e^{At}\right\}$$

$$pI - A = \begin{pmatrix} p & -1 & 1 \\ 0 & p+\alpha & 0 \\ 0 & 0 & p+\beta \end{pmatrix}$$

Контрольний приклад Понтрягіна

$$(pI - A)^{-1} = \frac{1}{p(p+\alpha)(p+\beta)} \begin{pmatrix} (p+\alpha)(p+\beta) & (p+\beta) & -(p+\alpha) \\ 0 & p(p+\beta) & 0 \\ 0 & 0 & p(p+\alpha) \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{1}{p} & \frac{1}{p(p+\alpha)} & -\frac{1}{p(p+\beta)} \\ 0 & \frac{1}{p+\alpha} & 0 \\ 0 & 0 & \frac{1}{p+\beta} \end{pmatrix} \Rightarrow e^{At} = \begin{pmatrix} 1 & \frac{1-e^{-\alpha t}}{\alpha} & -\frac{1-e^{-\beta t}}{\beta} \\ 0 & e^{-\alpha t} & 0 \\ 0 & 0 & e^{-\beta t} \end{pmatrix}$$

Тепер можна записати $\pi e^{At}(z^1, z^2, z^3) = z^1 + \frac{1-e^{-\alpha t}}{\alpha} z^2 - \frac{1-e^{-\beta t}}{\beta} z^3$, звідки

$$\pi e^{At} U = \left\{ \frac{1-e^{-\alpha t}}{\alpha} \cdot (-a) : \|a\| \leq \rho \right\}$$

$$\pi e^{At} V = \left\{ -\frac{1-e^{-\beta t}}{\beta} \cdot b : \|b\| \leq \sigma \right\}$$

Контрольний приклад Понтрягіна

Отже, $\pi e^{At}U$ — куля з радіусом $\frac{1-e^{-\alpha t}}{\alpha}\rho$ і центром в нулі, а $\pi e^{At}V$ — куля з радіусом $\frac{1-e^{-\beta t}}{\beta}\sigma$ і центром в нулі, тому $W(t) = \pi e^{At}U \dot{-} \pi e^{At}V$ — куля з радіусом $\frac{1-e^{-\alpha t}}{\alpha}\rho - \frac{1-e^{-\beta t}}{\beta}\sigma$ і центром теж в нулі. Радіус $W(t)$ буде додатнім при $\rho > \sigma$ та $\frac{\rho}{\alpha} > \frac{\sigma}{\beta}$.

$\Omega(t) = \int\limits_0^t W(s)ds$ буде кулею з радіусом:

$$\int\limits_0^t \left(\frac{1-e^{-\alpha s}}{\alpha}\rho - \frac{1-e^{-\beta s}}{\beta}\sigma\right)ds = \rho\int\limits_0^t \frac{1-e^{-\alpha s}}{\alpha}ds - \sigma\int\limits_0^t \frac{1-e^{-\beta s}}{\beta}ds =$$

$$= \frac{\rho}{\alpha^2}\left(\alpha t + e^{-\alpha t} - 1\right) - \frac{\sigma}{\beta^2}\left(\beta t + e^{-\beta t} - 1\right) = r(t)$$

Контрольний приклад Понтрягіна

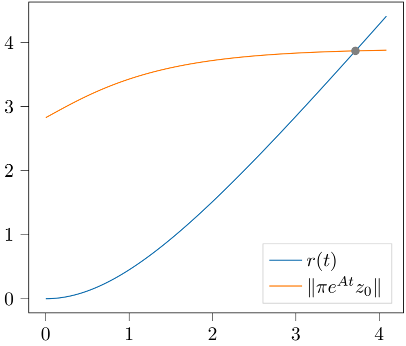
Далі необхідно знайти таке (найменше) значення T_0, для якого точка $\pi e^{AT_0}(z_0^1, z_0^2, z_0^3)$ належить $\Omega(T_0)$. З геометричних міркувань це буде найменший корінь рівняння

$$\frac{\rho}{\alpha^2}\left(\alpha t + e^{-\alpha t} - 1\right) - \frac{\sigma}{\beta^2}\left(\beta t + e^{-\beta t} - 1\right) = \left\|\pi e^{At} z_0\right\|$$

Контрольний приклад Понтрягіна

Наступний крок будемо проводити на конкретному прикладі. Нехай гра відбувається на площині з $\alpha = 1$, $\beta = 2$, $\rho = 2$, $\sigma = 1$ та початковими умовами $x(0) = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \dot{x}(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, y(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \dot{y}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Чисельно можна знайти значення $T_0 \approx 3.715$.



Метод розв'язуючих функцій

Метод розв'язуючих функцій належить А.О. Чикрію. Розглядається не просто лінійна диференціальна гра, а більш загальна *квазілінійна* гра виду:

$$\begin{cases} \dot{z} = Az + \varphi(u, v) \\ z(0) = z_0 \\ z \in \mathbb{R}^n, u \in U, v \in V \end{cases}$$

де $\varphi(u, v) : U \times V \to \mathbb{R}^n$ — неперервна за обома змінними функція.

Зауваження

У випадку $\varphi(u, v) = -u + v$ отримуємо лінійну диференціальну гру.

Метод розв'язуючих функцій

Термінальна множина має вид $M^* = M^0 + M$, де M^0 — деякий лінійний підпростір $\mathbb{R}^n$, а M — компактна підмножина ортогонального доповнення M^0, π — проектор на $(M^0)^\perp$.
Нехай $\varphi(U, v) = \{\varphi(u, v) : u \in U\}$ для фіксованої $v \in V$, $W(t, v) = \pi e^{At}\varphi(U, v)$, $W(t) = \bigcap\limits_{v \in V} W(t, v), t \geq 0$.

Зауваження

У випадку $\varphi(u, v) = -u + v$ $W(t) = \pi e^{At}U \mathbin{\dot{-}} \pi e^{At}V$.

Позначення

Вводяться позначення:

$$\gamma(t) \in W(t) \text{ — вимірна функція}$$

$$\xi(t, z, \gamma(\cdot)) = \pi e^{At} + \int\limits_0^t \gamma(\tau) d\tau$$

$$\alpha(t, \tau, z, v, \gamma(\cdot)) =$$

$$= \sup \{\alpha \geq 0 : [W(t-\tau, v) - \gamma(t-\tau)] \cap \alpha [M - \xi(t, z, \gamma(\cdot))] \neq \varnothing\}$$

$$T(z, \gamma(\cdot)) = \inf \left\{ t \geq 0 : \int\limits_0^t \inf_{v \in V} \alpha(t, \tau, z, v, \gamma(\cdot)) d\tau \geq 1 \right\}$$

Означення

$\alpha(t, \tau, z, v, \gamma(\cdot))$ називається *розв'язуючою функцією*.

Метод розв'язуючих функцій

Можна довести наступне: якщо $W(t) \neq \varnothing$ для всіх $t \geq 0$, M — опукла множина, $T(z_0, \gamma_0(\cdot)) < +\infty$ для деякого початкового положення z_0 та деякої $\gamma_0(\cdot)$, то за час $T(z_0, \gamma_0(\cdot))$ гравці потрапляють у термінальну множину.
Також можна довести, що якщо гра є лінійною, $W(t) = \pi e^{At} U \dot{-} \pi e^{At} V \neq \varnothing$, існує неперервна $r(t) : [0; +\infty) \to [0; +\infty)$ та число $l \geq 0$ такі, що $\pi e^{At} U = r(t) S$, $M = lS$, де S — одинична куля в $(M^0)^\perp$ з центром в нулі, то при $\xi(t, z, \gamma(\cdot)) \notin lS$, розв'язуюча функція α може бути знайдена як найбільший додатний корінь квадратного рівняння

Контрольний приклад Понтрягіна

Як і у минулому розв'язку цього прикладу перейдемо від системи

$$\begin{cases} \ddot{x} + \alpha\dot{x} = \rho u, & \|a\| \leq 1 \\ \ddot{y} + \beta\dot{y} = \sigma v, & \|b\| \leq 1 \end{cases}$$

до системи з $z_1 = x - y$, $z_2 = \dot{x}$, $z_3 = \dot{y}$:

$$\begin{cases} \dot{z}_1 = z_2 - z_3 \\ \dot{z}_2 = -\alpha z_2 + \rho u \\ \dot{z}_3 = -\beta z_3 + \sigma v \end{cases}$$

Термінальна множина $M^* = \{z : z_1 = 0\} = M^0 + \{0\}$, ортогональне доповнення $(M^0)^\perp = \{z : z_2 = z_3 = 0\}$, тому $\pi = \begin{pmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, де I та 0 — тотожній та нульовий оператор відповідно.

Контрольний приклад Понтрягіна

Оскільки у вихідній системі рівнянь $x, y \in \mathbb{R}^n$, то $z \in \mathbb{R}^{3n}$, то $\pi : \mathbb{R}^{3n} \to (M^0)^\perp$. Матриця системи $A = \begin{pmatrix} 0 & E & -E \\ 0 & -\alpha E & 0 \\ 0 & 0 & -\beta E \end{pmatrix}$, $U = \left\{ \begin{pmatrix} 0 \\ \rho u \\ 0 \end{pmatrix} : \|u\| \leq 1 \right\}$, $V = \left\{ \begin{pmatrix} 0 \\ 0 \\ \sigma v \end{pmatrix} : \|v\| \leq 1 \right\}$. Аналогічно минулому прикладу,

$$\pi e^{At} U = \frac{1 - e^{-\alpha t}}{\alpha} \rho S, \ \pi e^{At} V = \frac{1 - e^{-\beta t}}{\beta} \sigma S$$

$$W(t) = \left(\frac{1 - e^{-\alpha t}}{\alpha} \rho - \frac{1 - e^{-\beta t}}{\beta} \sigma \right) S = \omega(t) S$$

Радіус цієї кулі невід'ємний при $\rho \geq \sigma$ та $\frac{\rho}{\alpha} \geq \frac{\sigma}{\beta}$.

Контрольний приклад Понтрягіна

Поклавши $\gamma(t) = 0$, отримаємо $\xi(t,z,0) = z_1 + \frac{1-e^{-\alpha t}}{\alpha} z_2 - \frac{1-e^{-\beta t}}{\beta} z_3$. Ця задача задовольняє всі умови для пошуку розв'язуючої функції через:

$$\left\| \frac{1-e^{-\beta t}}{\beta} \sigma v - \alpha \cdot \xi(t,z,0) \right\| = \frac{1-e^{-\alpha t}}{\alpha} \rho$$

Можна показати, що

$$\min_{\|v\| \le 1} \alpha(t,\tau,z,v,0) = \frac{\omega(t-\tau)}{\|\xi(t,z,0)\|}$$

і мінімум досягається при $v = -\frac{\xi(t,z,0)}{\|\xi(t,z,0)\|}$

Контрольний приклад Понтрягіна

Час, коли переслідувач наздожене утікача, визначається як

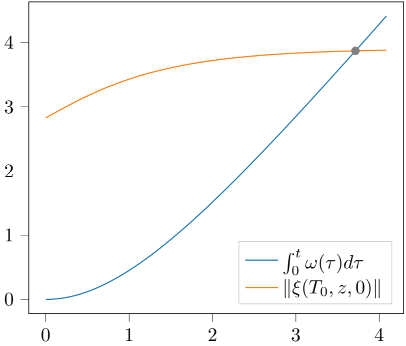
$$T(z,0) = \min\left\{ t \geq 0 : \int\limits_0^t \frac{\omega(t-\tau)}{\|\xi(t,z,0)\|} d\tau = 1 \right\}$$

або ж як найменший додатний корінь рівняння

$$\|\xi(t,z,0)\| = \int\limits_0^t \left(\frac{1-e^{-\alpha\tau}}{\alpha}\rho - \frac{1-e^{-\beta\tau}}{\beta}\sigma \right) d\tau$$

Контрольний приклад Понтрягіна

Розглянемо тепер конкретний приклад. Нехай ця гра відбувається на площині з $\alpha = 1$, $\beta = 2$, $\rho = 2$, $\sigma = 1$ та початковими умовами $x(0) = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \dot{x}(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, y(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \dot{y}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Чисельно можна знайти значення $T_0 \approx 3.715$.



Підсумок

У цій доповіді ми дослідили історію появи теорії диференціальних ігор та деякі задачі, що пояснюють необхідність виникнення цієї теорії. Також ознайомилися з основами теорії багатозначних відображень та двома методами для розв'язання лінійних диференціальних ігор переслідування: методом Понтрягіна та методом розв'язуючих функцій Чикрія. На прикладі контрольного прикладу Понтрягіна (лінійна гра) впевнилися, що обидва методи дають однаковий результат — час завершення переслідування.