

Міністерство освіти і науки України
Національний технічний університет України
«Київський політехнічний інститут ім. І. Сікорського»
Інститут прикладного системного аналізу

Доповідь
з курсу «Теорія ігор»
на тему «Основи диференціальних ігор.
Ігри простого переслідування»

Виконали

студенти 3 курсу групи КА-81
Галганов Олексій
Фордуй Нікіта

Прийняла

доцент кафедри ММСА
Барановська Леся Валеріївна

Київ 2021

Зміст

1	Вступ	2
1.1	Передумови для розгляду диференціальних ігор	2
1.2	Опис ходу гри	3
1.3	Приклади задання фазових змінних	3
2	Формалізація ігор	5
2.1	Опис руху	5
2.2	Виграші та стратегії гравців	7

Розділ 1

Вступ

1.1 Передумови для розгляду диференціальних ігор

Класична теорія ігор, творцями якої є Джон фон Нейман та Оскар Морґенштерн, оперує матричними іграми та зв'язаними з ними поняттями. Для них сформульовано та доведено багато фундаментальних теорем — наприклад, щодо існування рівноваги Неша. Однак, ці результати зовсім не покривають всі можливості формалізувати процеси, в яких гравці приймають рішення. Наприклад, якщо розглядати задачу про оптимальну траєкторію 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.

Визначення поняття диференціальної гри буде базуватися на тому, що до їх розв'язання (чи хоча б дослідження) буде застосовуватися апарат математичного аналізу, як-от теорія диференціальних рівнянь. Сам термін «диференціальна гра» було введено Руфусом Айзексом — одним з основоположників теорії диференціальних ігор.

Прикладами диференціальних ігор, окрім наведеного вище, є й звичайні спортивні ігри на кшталт футболу, де, очевидно, недоречно розглядати лише дискретні моменти часу або скінченний набір стратегій, бо найменше відхилення від таких стратегій (яке, звичайно, може статися через людський фактор), породжує вже іншу стратегію. Зауважимо, що коли рішення протягом усієї гри приймає лише один гравець, то така задача фактично стосується *теорії керування* та *варіаційного числення*, тому такі випадки в теорії ігор не розглядаються.

1.2 Опис ходу гри

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*. Для визначення результату гри необхідно знати значення фазових координат у початковий момент. Також, в кожен момент часу гравець має звертати увагу на значення фазових змінних (що, по суті, і означає деяких рух гравців). В процесі гри фазові координати змінюються. Оскільки розглядається випадок скінченної кількості фазових змінних, то фазовий простір зручно ототожнювати з координатами точок в \mathbb{R}^n (або в деякій підмножині цього простору). Поточні значення фазових координат завжди відомі гравцям — тобто, це ігри з повною інформацією. Невідомим зазвичай є характер їх зміни: тобто, *керування* фазовими змінними гравцями. Хід гри характеризується рухом точки у фазовому просторі, причому гра завершується, якщо виконуються деякі умови: наприклад, потрапляння точки в деяку підмножину простору.

1.3 Приклади задання фазових змінних

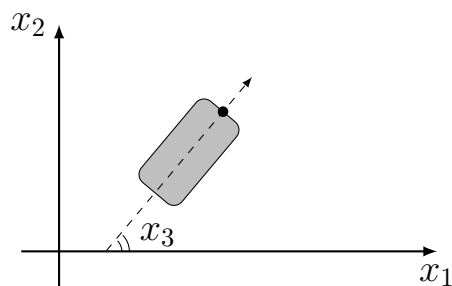
Розглянемо приклади задання фазових змінних у випадку керування лише одним об'єктом.

Приклад 1.1. 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, $\varphi \in \text{керуванням}$. Тоді рух точки описується системою диференціальних рівнянь

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

Інший гравець у будь-який момент часу може виміряти значення фазових координат x_1 та x_2 , але закону їх зміни (керування першого гравця) він не знає.

Приклад 1.2. 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 .



Якщо цей автомобіль є складовою диференціальної гри, то про нього треба знати більше. Нехай гравець керує ним за допомогою педалі акселератора, що задає прискорення, та керма, що задає напрямок руху. 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. Інший гравець у будь-який момент часу знає (чи може виміряти) значення фазових координат, але не значення керування φ_1 . Положення керма визначає кривину траєкторії руху автомобіля, значення якої, очевидно, є обмеженими. Таким чином, можна ввести кривину як ще одну фазову координату x_5 (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 .

Отже, маємо систему, що задає рух автомобіля у деякій диференціальній грі:

$$\begin{cases} \dot{x}_1 = x_4 \cos x_3 \\ \dot{x}_2 = x_4 \sin x_3 \\ \dot{x}_3 = x_4 x_5 \\ \dot{x}_4 = A\varphi_1, \varphi_1 \in [0; 1] \\ \dot{x}_5 = W\varphi_2, \varphi_2 \in [-1; 1] \end{cases}$$

Два перших рівняння — координатний запис швидкості руху автомобіля, третє означає, що швидкість зміни напрямку руху дорівнює добутку швидкості руху на кривину траєкторії, а два останні задають керування гравцем швидкості руху та швидкості зміни кривини траєкторії руху. Звісно, вони не є точними з фізичної точки зору (тому що, наприклад, не враховують тертя), але є досить простими для аналізу.

Розділ 2

Формалізація ігор

2.1 Опис руху

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

$$\begin{cases} \dot{x}_1(t) = f_1(x_1(t), \dots, x_n(t), u_1(x, t), \dots, u_P(x, t), v_1(x, t), \dots, w_E(x, t)) \\ \dot{x}_2(t) = f_2(x_1(t), \dots, x_n(t), u_1(x, t), \dots, u_P(x, t), v_1(x, t), \dots, w_E(x, t)) \\ \dots \\ \dot{x}_n(t) = f_n(x_1(t), \dots, x_n(t), u_1(x, t), \dots, u_P(x, t), v_1(x, t), \dots, w_E(x, t)) \\ x_1(0) = x_1^0, x_2(0) = x_2^0, \dots, x_n(0) = x_n^0 \end{cases} \quad (1)$$

або, коротше,

$$\begin{cases} \dot{x}(t) = f(x(t), u(x, t), v(x, t)) \\ x(0) = x_0 \end{cases} \quad (2)$$

Ці рівняння називаються *рівняннями руху*. Функції f_j є заданими та вважаються достатньо гладкими. Функції від часу u та v називаються *керуванням* та змінюються, відповідно, першим та другим гравцем, яких позначатимемо P та E . Ці позначення не є випадковими: вони походять від слів pursuer (переслідувач) та evader (утікач), оскільки саме такі задачі переслідування-втечі дали початок розвитку теорії диференціальних ігор.

Фазові координати x_1, \dots, x_n описують стан гри в тому сенсі, що якщо зупинити гру в будь-який момент часу, зафіксувати значення фазових координат та почати нову гру з цієї зафіксованої точки фазового простору, то її хід буде таким самим, як у початкової гри після моменту зупинки. Зокрема, значення x_0 фазових координат на початку гри, власне, є

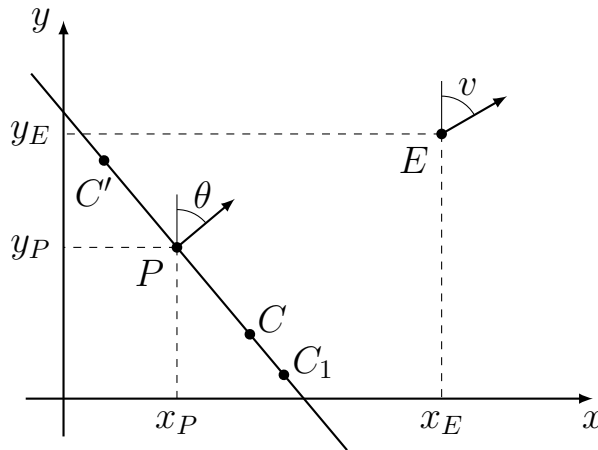
всім необхідним набором початкових даних, тому, навіть при однакових f_j та керуваннях, за різних початкових умов отримуватимемо, взагалі кажучи, різні «партії» гри.

Приклад 2.1. Якщо позначити через (x_P, y_P) координати гравця P , через (x_E, y_E) — гравця E , через w_P та w_E їх сталі швидкості руху, а керування напрямком швидкості через $u(t)$ та $v(t)$ відповідно, то отримаємо такі рівняння руху:

$$\begin{cases} \dot{x}_P(t) = w_P \cos u(t) \\ \dot{y}_P(t) = w_P \sin u(t) \\ \dot{x}_E(t) = w_E \cos v(t) \\ \dot{y}_E(t) = w_E \sin v(t) \\ (x_P(0), y_P(0)) = (x_P^0, y_P^0) \\ (x_E(0), y_E(0)) = (x_E^0, y_E^0) \end{cases}$$

Такий рух називається «переслідуванням на площині з простим рухом гравців».

Приклад 2.2 (гра «водій-вбивця»). В цьому випадку гра також 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 в деякому сенсі 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 .



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 відрізка $C'C$ (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:

$$\begin{cases} \dot{x}_P(t) = w_P \sin \theta(t) \\ \dot{y}_P(t) = w_P \cos \theta(t) \\ \dot{x}_E(t) = w_E \sin v(t) \\ \dot{y}_E(t) = w_E \cos v(t) \\ \dot{\theta}(t) = \frac{w_P}{R} u(t), \quad u(t) \in [-1; 1] \end{cases}$$

2.2 Виграші та стратегії гравців

Мета диференціальної гри визначається виграшем, який деяким чином залежить від траєкторій, що пройшли гравці до завершення гри. 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 $H : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ — деяка функція (нагадаємо, що розмірність $\mathcal{E} = n$). Наприклад, якщо

$$H(x(T), y(T)) = \|x(T) - y(T)\|$$

то гра описуватиме процес переслідування, в якому метою гравця E є відхилення від гравця P на момент кінця гри на максимально можливу відстань. В цьому випадку антагоністичність означає, що метою P є, навпаки, максимальне зближення з E на момент $t = T$. Також, можна в якості H використовувати

$$H(x(T), y(T)) = \min_{0 \leq t \leq T} \|x(t) - y(t)\|$$

Це означатиме, що гравцю E потрібно не просто віддалитися від P в останній момент гри, а й триматися якнайдалі від P протягом усього часу гри. Це буде так званий *мінімальний виграш*.

Гра також може завершуватися, коли обидва гравці потраплять до деякої підмножини $\mathcal{T} \subset \mathcal{E}$. Тоді в якості виграшу гравця E можна покласти

$$t_* = \min \{t \geq 0 : (x(t), y(t)) \in \mathcal{T}\}$$

t_* — перший момент потрапляння гравців до \mathcal{T} (або «захоплення гравцем P »). Такий виграш називається *термінальним виграшем*. Якщо для всіх $t \geq 0$ вони ніколи не потрапляють до \mathcal{T} , то виграш гравця E дорівнює $+\infty$. Наприклад, якщо в якості \mathcal{T} взяти гіперсферу радіуса $l \geq 0$,

то метою гравця P буде якнайшвидше зближення з E на відстань l . Можна також поставити задачу пошуку таких множин початкових умов, за яких t_* гарантовано буде скінченним або нескінченним. В такому випадку можна ввести *якісний вигравш*, що набуває значень ± 1 в залежності від того, чи вдалося E уникнути захоплення гравцем P .

Приклад 2.3. Розглянемо переслідування на площині з простим рухом, що описується системою

$$\begin{cases} \dot{x}_1 = u_1, \dot{x}_2 = u_2, u_1^2 + u_2^2 \leq \alpha^2 \\ \dot{y}_1 = v_1, \dot{y}_2 = v_2, v_1^2 + v_2^2 \leq \beta^2 \\ x_1(0) = x_1^0, x_2(0) = x_2^0, y_1(0) = y_1^0, y_2(0) = y_2^0 \end{cases}$$

Тут $P(t) = (x_1, x_2)$ — координати гравця P , $E(t) = (y_1, y_2)$ — координати гравця E . u_1, u_2 та v_1, v_2 — їх керування відповідно, причому з умови швидкості руху (зміни координат гравців) обмежені максимальними значеннями α та β . Обидва гравці, обираючи керування, змінюють напрямок руху.

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

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

тобто, наближення до E на будь-яку відстань за скінченну кількість часу. Для цього достатньо рухатися з максимальною швидкістю α в тому ж напрямку, що і E .

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

Означення 1. *Стратегіями* у диференціальній грі є вибір керувань u та v як функцій від часу t та фазових координат x у системі рівнянь руху (1).

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 .

Надалі для спрощення розглядатимемо не один вектор x , а два вектори x та y , що відповідатимуть руху кожного з гравців. Тоді систему (2) можна записати як

$$\begin{cases} \dot{x}(t) = f(x(t), u(x, y, t)) \\ \dot{y}(t) = g(x(t), v(x, y, t)) \\ x(0) = x_0, y(0) = y_0 \end{cases} \quad (1)$$

Означення 2. Набір $S = \{x_0, y_0, u(\cdot), v(\cdot)\}$, де x_0, y_0 — початкові умови, а $u \in P, v \in E$ — керування, називається *ситуацією* в диференціальній грі.

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

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

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

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

Означення 3. Користуючись означенням ситуації, можна ввести *виграш* в ситуації $S = \{x_0, y_0, u(\cdot), v(\cdot)\}$ як функцію $K(x_0, y_0, u(\cdot), v(\cdot))$.

Приклади виграшів було наведено вище. Траєкторії $x(t)$ та $y(t)$ в них визначаються саме з ситуації S . Наведемо строгі означення.

Означення 4. *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

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

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

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

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 $K = +\infty$.

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

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

Нарешті, можна дати означення нормальної форми диференціальної гри.

Означення 5. Нормальною формою диференціальної гри $\Gamma(x_0, y_0)$, заданої на просторі стратегій $P \times E$, називається система

$$\Gamma(x_0, y_0) = \langle x_0, y_0, P, E, K(x_0, y_0, u(\cdot), v(\cdot)) \rangle \quad (2)$$

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