Мы начали семестр с показа численного решения некоторых задач теории дифференциальных игр: задача посадки самолёта в условиях ветрового возмущения, задачи преследования двумя объектами одного убегающего.

В каждой из этих задач были два игрока и цели их были противоположны: если один минимизировал некоторый показатель, то другой его максимизировал. Такие игры называют антагонистическими (или играми с противоположными интересами). Дифференциальные игры — поскольку динамика описывалась при помощи дифференциальных соотношений (чаще говорят — дифференциальных уравнений). Всю машину дифференциальных соотношений можно считать описанием одного объекта, который управляется двумя игроками.

Задачи управления также можно считать играми, но это игра с одним игроком.

В задачах управления мы различали два вида законов управления: программное управление и управление по принципу обратной связи. В первом случае управление есть функция $t \to u(t)$, во втором — функция $(t, x) \to u(t, x)$, где x — символ фазового состояния. Если помеху не учитывать, т.е. строго находиться в рамках постановки с одним игроком, то эти два варианта обеспечивают один и тот же наилучший результат.

Принципиальное отличие в дифференциальных играх: наилучшие результаты, полученные при помощи программного управления, как правило намного хуже тех, которые можно обеспечить в классе управлений по принципу обратной связи.

Переходим к дифференциальным играм.

Чтобы перед глазами был некоторый конкретный образ, рассмотрим модельную задачу.

Задача о преследовании двумя догоняющими одного убегающего

① Представим ситуацию воздушного боя. В горизонтальной плоскости два самолёта атакуют самолёт противника на встречных курсах. Пусть в начальный момент идеализированная картинка выглядит следующим образом:

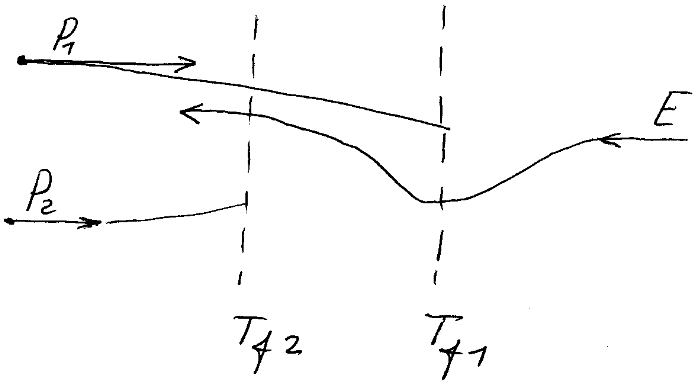
Через P_1, P_2 обозначаем преследователей (P — начальная буква слова pursuer), E — убегающий (evader). Начальные скорости велики настолько, что о промахе P_1 относительно E можно говорить только в момент расхождения движений P_1 и E по горизонтали, аналогично — о промахе P_2 относительно E. Пусть T_{f1} и T_{f2} — моменты такого расхождения. Они полностью определяются начальными данными.

В качестве промаха будем брать расстояние бокового отклонения: φ_1 и φ_2. Положим

$\varphi = \min\{\varphi_1, \varphi_2\}$.

Преследователи P_1 и P_2 действуют координированно. Объединённого преследователя обозначим буквой P. Игрок P минимизирует значение φ, игрок E максимизирует (антагонистическая игра!).

В процессе движения игроки P и E имеют возможность измерять положения и скорости друг друга. Ясно, что при помощи программного управления (объявленного заранее!) игрок P ничего хорошего не получит.



Игроку P надо управлять по принципу обратной связи. Аналогично, если да следуем за игроком E.

② Описание динамики. Пусть y_1 – разность между положениями E и P_1 по „вертикали" (боковое отклонение E от P_1). Аналогично, y_2 – между E и P_2.

Будем использовать в линейном приближении такую систему:

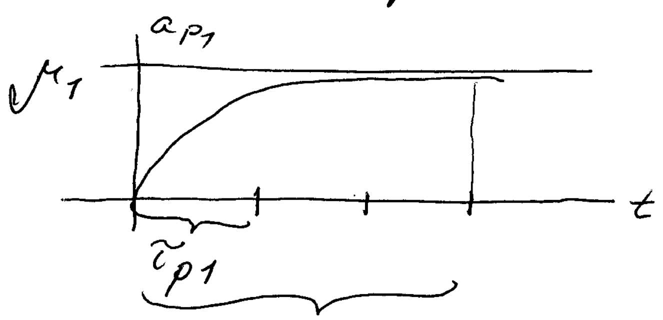
$$\ddot{y}_1 = -a_{P_1} + a_E \qquad \ddot{y}_2 = -a_{P_2} + a_E$$

$$\dot{a}_{P_1} = (u_1 - a_{P_1})/\tau_{P_1} \qquad \dot{a}_{P_2} = (u_2 - a_{P_2})/\tau_{P_2}$$

$$\dot{a}_E = (v - a_E)/\tau_E$$

$|u_1| \le \mu_1$, $|u_2| \le \mu_2$; $|v| \le \nu$.

Здесь a_{P_1}, a_{P_2} – боковые ускорения преследователей P_1 и P_2, a_E – боковое ускорение убегающего E. Далее: u_1, u_2 – управления преследователей. Их можно рассматривать как электрические сигналы, которые через дифференциальные уравнения первого порядка (т.е. через однократное интегрирование) преобразуются в действительные ускорения. Константы τ_{p1} и τ_{p2} определяют время этого преобразования. Эти константы называются постоянными времени переходного процесса:



Аналогично учитывается инерционность отработки управляющего воздействия v убегающим E.

Очень часто мы (рассуждая за преследователя) можем не знать детали описания динамики убегающего. А именно, в нашем случае можем не знать постоянную τ_E. Страхуясь, можем считать её близкой к нулю. Типичный приём: выкидываем соотношение $\dot{a}_E = (v - a_E)/\tau_E$, а соотношения для уклонения возьмём в виде

$$\ddot{y}_1 = -a_{p1} + v \qquad \ddot{y}_2 = -a_{p2} + v$$

$$\dot{a}_{p1} = (u_1 - a_{p1})/\tau_{p1} \qquad \dot{a}_{p2} = (u_2 - a_{p2})/\tau_{p2}$$

$$|u_1| \le \mu_1, \ |u_2| \le \mu_2; \quad |v| \le \nu.$$

Сейчас управление v отрабатывается безынерционно.

Постановка задачи за первого игрока (игрока P)

Первым в дальнейшем называем игрока, который стремится минимизировать гарантированный результат. В нашем случае это игрок P. Другого игрока, который стремится максимизировать, будем называть вторым.

Предположим, что первый игрок выбрал некоторую функцию (стратегию) $U(t, x)$. Здесь в нашем случае фазовый вектор

$$x = (y_1, \dot{y}_1, a_{p1}, y_2, \dot{y}_2, a_{p2}, a_E)^T.$$

Применяя U, первый игрок может столкнуться с некоторой реализацией $v(\cdot)$ управления второго игрока. Получаем результат $\varphi(t_0, x_0; U, v(\cdot))$. При закреплённой функции U наихудший результат для первого игрока будет

$$\varphi^{(1)}(t_0, x_0; U) = \sup_{v(\cdot)} \varphi(t_0, x_0; U, v(\cdot)) \qquad (1)$$

Тогда наилучший гарантированный результат для первого игрока при начальной позиции (t_0, x_0) есть

$$\varphi^{(1)}(t_0, x_0) = \inf_{U} \varphi^{(1)}(t_0, x_0; U) \qquad (2)$$

Мы подсчитали наилучший гарантированный результат для начальной позиции (t_0, x_0). Это можно сделать для всех начальных позиций. Получим в итоге функцию $\Gamma^{(1)}: (t, x) \to \Gamma^{(1)}(t, x)$. Назовём её функцией наилучшего гарантированного результата первого игрока.

А где оптимальное управление обратной связи?

Смотрим формулу (2). Если здесь min рассчитается, то получим V^*. Но подчеркнём, что функция $(t, x) \to V^*(t, x)$ может зависеть от (t_0, x_0). Формально это видно из (2).

Оказалось, что это и на самом деле так.

Постановка задачи за второго игрока (игрока E).

Рассуждая за второго игрока, запишем формулы, аналогичные (1), (2).

Зафиксируем функцию (стратегию) $V(t, x)$.

Пусть

$$\varphi^{(2)}(t_0, x_0; V) = \inf_{u(\cdot)} \varphi(t_0, x_0; V, u(\cdot)), \qquad (3)$$

$$\varphi^{(2)}(t_0, x_0) = \sup_{V} \varphi^{(2)}(t_0, x_0; V). \qquad (4)$$

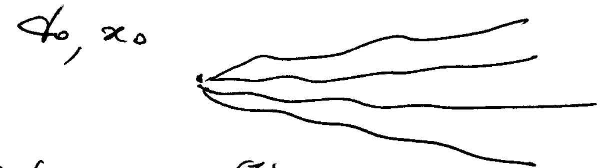
Формула (4) даёт наилучший гарантированный результат второго игрока для позиции (t_0, x_0). Делая (t_0, x_0) свободными, получим функцию $(t, x) \to \Gamma^{(2)}(t, x)$. Это функция наилучшего гарантированного результата второго игрока

Оказывается, что для широкого класса дифференциальных игр справедливо равенство $P^{(1)} = P^{(2)}$.

Итак, главные моменты:

1). Отдельно рассматриваем задачи за первого и второго игроков.

2). Взяв начальную позицию (t_0, x_0) и зафиксировав стратегию $U(t, x)$ первого игрока, воображаем пучок всех возможных движений из позиции (t_0, x_0), каждое из них соответствует некоторой функции $v(\cdot)$.



Значение $\varphi^{(1)}(t_0, x_0; U)$ даёт гарантию, соответствующую этому пучку, т.е. стратегии U.

Аналогично за второго игрока.

3). Говорим о наилучших гарантированных результатах и об оптимальных стратегиях.

Основная недоработка наших выкладок состоит в том, что мы не определили что такое движение из начальной позиции (t_0, x_0) при выбранных функциях $U(t,x)$, $v(\cdot)$ в выкладках за первого игрока и что такое движение в силу $V(t,x)$, $u(\cdot)$, когда рассуждаем за второго игрока.

Поэтому наши выкладки, связанные с определением наилучших гарантированных результатов, следует признать нестрогими.

Строгая постановка задач за первого и второго игроков в дифференциальной игре

Введём понятия допустимых стратегий игроков и дадим строгие определения наилучших гарантированных результатов.

① Пусть динамика описывается соотношением

$$\dot{x} = f(t, x, u, v), \quad x \in R^n, \quad u \in P, \quad v \in Q \qquad (5)$$

Здесь x — фазовый вектор; u, v — управляющие воздействия первого и второго игроков; P, Q — ограничения на мгновенные значения (компакты в R^p и R^q). Функцию f предполагаем достаточно хорошей: так, чтобы при любых кусочнонепрерывных функциях $u(t)$, $v(t)$ со значениями в P и Q можно было бы говорить об единственном решении, выходящем в момент (произвольный) t_* из произвольной точки x_*, и чтобы такое решение можно было продолжить на весь промежуток, для чего подробно оговорили вектор функции $u(t)$ и $v(t)$. Подобные условия (при отсутствии второго игрока) были сформулированы в Л.6.

Будем считать, что игра заканчивается в момент T. Первый игрок заинтересован в минимизации значения заданной функции платы $\varphi(x(T))$, а второй игрок старается, чтобы значение $\varphi(x(T))$ было как можно больше. Функцию φ полагаем непрерывной.

Итак, имеем игру с фиксированным моментом окончания и терминальной функцией платы.

В примере с двумя регистраторами у нас были два момента, отсчитываемых уровнях. Таким образом, сейчас рассматриваем несколько более простую ситуацию.

② Для постановки задачи за первого игрока мы должны ввести множество $\mathcal{U}$ его допустимых стратегий. В качестве $\mathcal{U}$ возьмём совокупность всех функций $(t, x) \to V(t,x)$ со значениями в P. Подчеркнём, что мы не предполагаем даже непрерывность функции V.

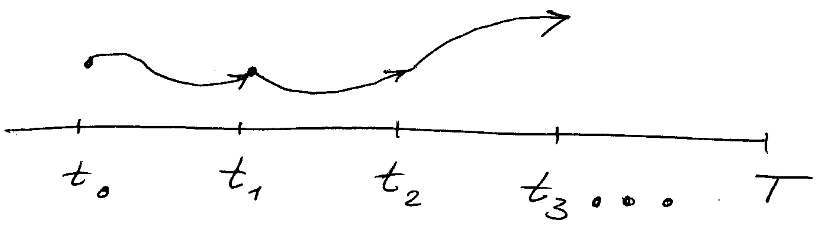
Пусть $\Delta^{(1)}$ — некоторое разбиение промежутка $[t_0, T]$, которое выбирает первый игрок.

Зафиксируем начальную позицию (t_0, x_0).

Предположим, что выбраны некоторые $V \in \mathcal{U}$ и $\Delta^{(1)}$. Дискретная схема управления первого игрока действует следующим образом.

Имея управляющее воздействие $V(t_0, x_0)$, первый игрок держит его постоянным на первом промежутке $[t_0, t_1)$ разбиения $\Delta^{(1)}$. На этом же промежутке реализуется некоторое кусочно-непрерывное управление $v(t)$ со значениями в Q. Подставляя $V(t_0, x_0)$ вместо u в правую часть системы (5) и функцию $v(t)$ вместо v, получим решение $t \to x(t)$ системы (5) на $[t_0, t_1)$.

В момент t_1 имеем положение $x(t_1)$. Вычисляем $U(t_1, x(t_1))$, сталкиваем его с некоторой кусочно-непрерывной функцией $v(t)$ на $[t_1, t_2)$ и получим продолжение нашего решения на $[t_1, t_2)$ и т.д., пока не дойдём до момента T.



В итоге имеем движение $t \to x(t; t_0, x_0, U, \Delta^{(1)}, v(\cdot))$ и значение $\varphi(x(T; t_0, x_0, U, \Delta^{(1)}, v(\cdot)))$ функции φ в момент T.

Теперь можно сосчитать

$$\varphi^{(1)}(t_0, x_0; U, \Delta^{(1)}) = \sup_{v(\cdot)} \varphi(x(T; t_0, x_0, U, \Delta^{(1)}, v(\cdot))) \qquad (6)$$

Величину $\varphi^{(1)}(t_0, x_0; U, \Delta^{(1)})$ можно назвать гарантией первого игрока, когда он применяет стратегию U в дискретной схеме управления с разбиением $\Delta^{(1)}$.

Далее можно варьировать диаметр δ разбиения промежутка $[t_0, T]$ и рассматривать различные разбиения $\Delta^{(1)}$ с диаметром

$\operatorname{diam} \leq \delta$. Определим

$$\varphi_\delta^{(1)}(t_0, x_0, U) = \sup_{\Delta^{(1)}} \varphi^{(1)}(t_0, x_0, U, \Delta^{(1)}) \qquad (7)$$

Здесь перебираем все разбиения $\Delta^{(1)}$ с диаметром $\operatorname{diam} \leq \delta$. Получаем гарантию стратегии U при начальной позиции (t_0, x_0) при использовании любого $\Delta^{(1)}$ с $\operatorname{diam} \leq \delta$.

Пусть

$$\varphi^{(1)}(t_0, x_0, U) = \overline{\lim_{\delta \to 0}}\, \varphi_\delta^{(1)}(t_0, x_0, U) \qquad (8)$$

и, наконец

$$\varphi^{(1)}(t_0, x_0) = \inf_{U} \varphi^{(1)}(t_0, x_0, U) \qquad (9)$$

Значение $\varphi^{(1)}(t_0, x_0)$ есть наилучшая гарантия первого игрока для позиции (t_0, x_0). Если в (2) $\inf$ достигается на некоторой функции U^*, то это есть наилучшая стратегия первого игрока для этой начальной позиции.

Делая (t_0, x_0) свободными, получаем функцию $(t_0, x_0) \to \varphi^{(1)}(t_0, x_0)$, которую обозначим $\Gamma^{(1)}$.

Наша цель — нахождение функции $\Gamma^{(1)}$ и оптимальных стратегий U^* (или близких к оптимальным, если inf в (9) не достигается). Но мы должны помнить, что U^* зависит от (t_0, x_0).

③ Симметричная постановка за второго игрока

Начинаем с $V(t,x)$, $\Delta^{(2)}$, определяем

$$\varphi^{(2)}(t_0, x_0; V, \Delta^{(2)}) = \inf_{u(\cdot)} \varphi(x(T; t_0, x_0, V, \Delta^{(2)}, u(\cdot))) \qquad (10)$$

$$\varphi^{(2)}_\delta(t_0, x_0; V) = \inf_{\Delta^{(2)}} \varphi^{(2)}(t_0, x_0, V, \Delta^{(2)}) \qquad (11)$$

$$\varphi^{(2)}(t_0, x_0, V) = \varliminf_{\delta \to 0} \varphi^{(2)}_\delta(t_0, x_0, V) \qquad (12)$$

Наконец,

$$\varphi^{(2)}(t_0, x_0) = \sup_{V} \varphi^{(2)}(t_0, x_0, V) \qquad (13)$$

Делая (t_0, x_0) свободными, получаем функцию $\Gamma^{(2)}$.

В задаче за второго игрока нас интересует функция $\Gamma^{(2)}$ и оптимальные стратегии V^*, но они могут зависеть от (t_0, x_0).