ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ

Глава 1. Введение

1.1. Динамика

Начнём с рассмотрения обыкновенного дифференциального уравнения (ОДУ):

$$\begin{cases} \dot{x}(t) = f(x(t)) \quad (t > 0) \\ x(0) = x^0 \end{cases} \tag{1}$$

В данном случае у нас есть начальная точка $x^0 \in \mathbb{R}^n$ и функция $f : \mathbb{R}^n \to \mathbb{R}^n$. Неизвестным является кривая $x : [0, \infty) \to \mathbb{R}^n$, которую мы интерпретируем как динамическую эволюцию состояния некоторой системы.

1.2. Контролируемая динамика

Мы немного обобщим и предположим, что теперь функция f также зависит от некоторых “контролирующих” параметров принадлежащих множеству $A \subset \mathbb{R}^m$, так чтобы $f : \mathbb{R}^n \times A \to \mathbb{R}^n$. Тогда, если мы выберем некоторый элемент $a \in A$ и рассмотрим динамику:

$$\begin{cases} \dot{x}(t) = f(x(t), a) \quad (t > 0) \\ x(0) = x^0, \end{cases} \tag{2}$$

мы получим эволюцию нашей системы в ситуации, когда параметр является константой равной a.

Далее мы можем изменять значение параметра вместе с изменением системы. Определим функцию $\alpha : [0, \infty) \to A$, называемую контролем. В таком случае мы получаем следующее ОДУ:

$$\begin{cases} \dot{x}(t) = f(x(t), \alpha(t)) \quad (t > 0) \\ x(0) = x^0, \end{cases} \tag{3}$$

при этом траекторию $x(\cdot)$ мы будем воспринимать как ответ системы.

1.3. Нотация

Мы будем писать:

$$f(x,a) = \begin{pmatrix} f^1(x,a) \\ \vdots \\ f^n(x,a) \end{pmatrix}, \tag{4}$$

чтобы изобразить компоненты f, аналогично:

$$x(t) = \begin{pmatrix} x^1(t) \\ \vdots \\ x^n(t) \end{pmatrix}. \tag{5}$$

Мы также обозначим:

$$\mathbb{A} = \{\alpha : [0,\infty) \to A \,|\, \alpha(\cdot)\,\text{измерима}\}, \tag{6}$$

чтобы обозначить множество всех возможных допустимых контролей, где:

$$\alpha(t) = \begin{pmatrix} \alpha^1(t) \\ \vdots \\ \alpha^m(t) \end{pmatrix}. \tag{7}$$

Обратите внимание, что решение ОДУ $x(\cdot)$ зависит от $\alpha(\cdot)$ и начальных условий. Поэтому, строго говоря, мы бы должны были писать $x(\cdot) = x(\cdot, \alpha(\cdot), x^0)$, чтобы отразить зависимость ответа системы $x(\cdot)$ от контроля и начального значения.

1.4. Награда

Наша общая задача будет определить “наилучший” контроль для нашей системы. Для этого нам потребуется ввести специальный функционал награды:

$$P[\alpha(\cdot)] := \int_0^T r(x(t), \alpha(t))dt + g(x(T)), \tag{8}$$

где $x(\cdot)$ решает ОДУ для контроля $\alpha(\cdot)$. Здесь $r : \mathbb{R}^n \times A \to \mathbb{R}$ и $g : \mathbb{R}^n \to \mathbb{R}$ заданы, и r называется скользящей наградой, а g – конечной наградой. Конечное время $T > 0$ также задано.

1.5. Постановка задачи

Нашей целью является нахождение контроля $\alpha^*(\cdot)$, который максимизирует награду. Другими словами мы хотим, чтобы:

$$P[\alpha^*(\cdot)] \geq P[\alpha(\cdot)] \tag{9}$$

для любого контроля $\alpha(\cdot) \in \mathbb{A}$. Такой контроль $\alpha^*(\cdot)$ называется оптимальным.

Такая задача ставит перед нами несколько математических вопросов:

1. Существует-ли оптимальный контроль?
2. Как мы можем математически характеризовать оптимальный контроль?
3. Как мы можем построить оптимальный контроль?

Глава 2. Примеры

2.1. Контроль производства и потребления

Допустим, мы владеем заводом, чью продукцию мы можем контролировать. Начнем с построения математической модели, положив:

$$x(t) = \text{количество продукции, произведенной в момент времени } t \geq 0. \tag{10}$$

Мы предположим, что мы потребляем какую-то долю продукции в каждый момент времени, и также повторно используем оставшуюся часть продукции. Обозначим:

$$\alpha(t) = \text{доля повторно использованной продукции в момент времени } t \geq 0. \tag{11}$$

Это будет нашим контролем, для которого есть естественные ограничения:

$$0 \leq \alpha(t) \leq 1 \text{ в каждый в момент времени } t \geq 0. \tag{12}$$

Имея такой контроль, соответствующая динамика описывается ОДУ:

$$\begin{cases} \dot{x}(t) = k\alpha(t)x(t) \\ x(0) = x^0. \end{cases} \tag{13}$$

константа $k > 0$ моделирует скорость роста нашего повторного использования. Рассмотрим следующий функционал награды:

$$P[\alpha(\cdot)] := \int_0^T (1 - \alpha(t))x(t)dt. \tag{14}$$

Смысл в том, что мы хотим максимизировать наше общее потребление продукции, при учёте того, что наше потребление в момент времени t равно

$(1-\alpha(t))x(t)$. Эта модель соответствует нашей общей постановке для $n = m = 1$, если мы примем:

$$\mathbb{A} = [0,1], f(x,a) = kax, r(x,a) = (1-a)x, g \equiv 0 \tag{15}$$

Как мы потом выясним, оптимальный контроль $\alpha^*(\cdot)$ задается **переключением**:

$$\alpha^*(\cdot) = \begin{cases} 1 & \text{, если } 0 \le t \le t^* \\ 0 & \text{, если } t^* < t \le T \end{cases} \tag{16}$$

2.2. Маятник

Рассмотрим маятник, для которого:

$$\theta(t) = \text{угол в момент времени } t. \tag{17}$$

Если внешние силы отсутствуют, то у нас есть следующие уравнения движения:

$$\begin{cases} \ddot{\theta}(t) + \lambda\dot{\theta}(t) + \omega^2\theta(t) = 0 \\ \theta(0) = \theta_1, \dot{\theta}(0) = \theta_2; \end{cases} \tag{18}$$

решением которых являются затухающие колебания, при условии $\lambda > 0$.

Теперь введем крутящий момент $\alpha(\cdot)$, для которого есть физическое ограничение:

$$|\alpha| \le 1 \tag{19}$$

Тогда наша динамика становится:

$$\begin{cases} \ddot{\theta}(t) + \lambda\dot{\theta}(t) + \omega^2\theta(t) = \alpha(t) \\ \theta(0) = \theta_1, \dot{\theta}(0) = \theta_2; \end{cases} \tag{20}$$

Определим $x_1(t) = \theta(t), x_2(t) = \dot{\theta}(t)$ и $x(t) = (x_1(t), x_2(t))$. Тогда мы можем записать эволюцию нашей системы как:

$$\dot{x}(t) = \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} \dot{\theta} \\ \ddot{\theta} \end{pmatrix} = \begin{pmatrix} x_2 \\ -\lambda x_2 - \omega^2 x_1 + \alpha(t) \end{pmatrix} = f(x, \alpha). \tag{21}$$

Мы также введём:

$$P[\alpha(\cdot)] = -\int_0^\tau 1dt = -\tau, \tag{22}$$

для:

$$\tau = \tau(\alpha(\cdot)) = \text{первый момент, когда } x(\tau) = 0 \tag{23}$$

Мы хотим максимизировать $P[\cdot]$, тем самым мы хотим минимизировать время, за которое маятник придет в спокойствие.

Глава 3. Линейный оптимальный контроль

3.1. Существование оптимального контроля

Рассмотрим ОДУ:

$$\begin{cases} \dot{x}(t) = Mx(t) + N\alpha(t) \\ x(0) = x^0, \end{cases} \tag{24}$$

для данных матриц $M \in \mathbb{M}^{n\times n}$ и $N \in \mathbb{M}^{n\times m}$. Обозначим за $\mathbb{A}$ куб $[-1,1]^m \subset \mathbb{R}^m$. Определим:

$$P[\alpha(\cdot)] = -\int_0^\tau 1dt = -\tau, \tag{25}$$

где $\tau = \tau(\alpha(\cdot))$ обозначает первый момент, когда ОДУ доходит до 0. Если траектория никогда не доходит до 0, то $\tau = \infty$.

3.2. Задача оптимального контроля

Нам дана начальная точка $x^0 \in \mathbb{R}^n$, и мы хотим найти оптимальный контроль $\alpha^*(\cdot)$ такой, что:

$$P[\alpha^*(\cdot)] = \max_{\alpha(\cdot)\in\mathbb{A}} P[\alpha(\cdot)]. \tag{26}$$

Тогда:

$$\tau^* = -\rho[\alpha^*(\cdot)]\text{это минимальное время для касания нуля.} \tag{27}$$

Теорема 1. ***Существование оптимального контроля.*** *Пусть $x^0 \in \mathbb{R}^n$. Тогда существует оптимальное переключение $\alpha^*(\cdot)$.*

3.3. Принцип максимума для оптимального контроля

Действительно интересная практическая задача теперь понять, как вычислить оптимальный контроль $\alpha^*(\cdot)$.

Определение 3.1. *Определим $K(t, x^0)$ как множество достижимости в момент времени t. То есть $K(t, x^0) = \{x^1 \mid$ существует $\alpha(\cdot) \in \mathbb{A}$, которое переключает из x^0 в x^1 в момент времени $t\}$.*

Так как $x(\cdot)$ решает ОДУ, то $x^1 \in K(t, x^0)$ тогда и только тогда:

$$x^1 = X(t)x^0 + X(t)\int_0^t X^{-1}N\alpha(s)ds = x(t) \tag{28}$$

для некоторого контроля $\alpha(\cdot) \in \mathbb{A}$.

Теорема 2. ***Геометрия множества** K. Множество $K(t, x^0)$ выпуклое и замкнутое.*

Теорема 3. ***Принцип максимума Понтрягина**. Существует ненулевой вектор h такой, что*

$$h^T X^{-1}(t)N\alpha^*(t) = \max_{a\in\mathbb{A}}\{h^T X^{-1}Na\} \tag{29}$$

для каждого момента времени $0 \le t \le \tau^*$.

Теорема 4. ***Принцип максимума Понтрягина в сопряженной форме**. Пусть $\alpha^*(\cdot)$ это оптимальный контроль, и $x^*(\cdot)$ это соответсвующий ему ответ. Тогда существует функция $p^*(\cdot) : [0, \tau^*] \to \mathbb{R}^n$ такая, что*

$$\dot{x}^*(t) = \bigtriangledown_p H(x^*(t), p^*(t), \alpha^*(t)) \tag{30}$$

$$\dot{p}^*(t) = - \bigtriangledown_x H(x^*(t), p^*(t), \alpha^*(t)) \tag{31}$$

$$H(x^*(t), p^*(t), \alpha^*(t)) = \max_{a \in \mathbb{A}} H(x^*(t), p^*(t), a) \tag{32}$$

.