И. А. Крылов, Ф. Л. Черноусько, О методе последовательных приближений для решения задач оптимального управления, *Ж. вычисл. матем. и матем. физ.*, 1962, том 2, номер 6, 1132–1139

О МЕТОДЕ ПОСЛЕДОВАТЕЛЬНЫХ ПРИБЛИЖЕНИЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

И. А. КРЫЛОВ, Ф. Л. ЧЕРНОУСЬКО

(Москва)

Нахождение оптимального управления для процесса, описываемого системой обыкновенных дифференциальных уравнений n-го порядка, с помощью принципа максимума Л. С. Понтрягина [1] сводится к краевой задаче для системы уравнений $2n$-го порядка. Решение краевых задач для сложных нелинейных систем часто встречает значительные вычислительные трудности и требует большой затраты машинного времени. Поэтому представляют интерес методы решения задач оптимального управления, позволяющие обойти эти трудности.

Ниже предлагается метод последовательных приближений для определения оптимального управления в задачах со свободным правым концом, основанный на принципе максимума.

Для иллюстрации метода приводятся результаты численного решения одной вариационной задачи.

1. Пусть задана система обыкновенных дифференциальных уравнений

$$\frac{dx}{dt} = f(t, x, u). \tag{1}$$

Здесь $x = (x_1, \ldots, x_n)$ — n-мерная вектор-функция времени t (переменные x_i будем называть координатами), $u = (u_1, \ldots, u_m)$ — m-мерная вектор-функция времени (u_k — управляющие функции), $f = (f_1, \ldots, f_n)$ — n-мерная вектор-функция $(n + m + 1)$ переменных.

Координаты должны удовлетворять начальным условиям (t_0 — начальный момент времени, x_0 — заданный n-мерный вектор)

$$x(t_0) = x_0. \tag{2}$$

На управляющие функции могут быть наложены ограничения, которые обычно выражаются в виде

$$u \in U, \tag{3}$$

где U — замкнутое множество m-мерного пространства. Кусочно-непрерывные управления, удовлетворяющие ограничению (3), будем называть допустимыми.

Функционал, который надлежит минимизировать (или максимизировать), задаем в форме

$$J = \sum_{i=1}^{n} c_i x_i(T), \qquad \sum_{i=1}^{n} c_i^2 \neq 0, \qquad T > t_0. \tag{4}$$

К такому виду приводится (за счет увеличения числа переменных в системе (1)) широкий класс функционалов (см. [2]), в частности функционалы типа

$$J = \int_{t_0}^{T} F(t, x, u)\, dt.$$

Будут рассматриваться вариационные задачи двух типов: во-первых, задачи с фиксированным конечным моментом T (при этом на координаты $x(T)$ не накладывается никаких ограничений), во-вторых, задачи, в которых время T не фиксировано, а задана одна из координат (или их комбинация) в конце процесса управления. В последнем случае, без нарушения общности, будем считать

$$x_1(T) = a. \tag{5}$$

Задачи обоих типов являются вариационными задачами со свободным правым концом (в момент T задано лишь одно условие, служащее для определения окончания процесса).

Задача состоит в построении допустимого управления $u(t)$, при котором функционал J достигает наименьшего значения по сравнению с его значениями при других допустимых управлениях. Это управление будем называть оптимальным и обозначать через $u^*(t)$.

Введем функцию

$$H(t, x, p, u) = \sum_{i=1}^{n} p_i f_i(t, x, u). \tag{6}$$

Функции $p_i(t)$ («импульсы») должны удовлетворять системе уравнений, канонически сопряженной системе (1), с функцией Гамильтона H из формулы (6):

$$\frac{dp_i}{dt} = -\sum_{k=1}^{n} \frac{\partial f_k(t, x, u)}{\partial x_i} p_k. \tag{7}$$

Начальные условия для импульсов в случае, когда время фиксировано, имеют вид

$$p_i(T) = -c_i \qquad (i = 1, 2, \ldots, n). \tag{8}$$

Если в конце процесса задано условие (5), то первое из условий (8) заменится условием (см. [1], [2])

$$H(T) \equiv \sum_{i=1}^{n} p_i(T) f_i(T, x(T), u(T)) = 0, \tag{9}$$

а остальные условия (8) не изменятся.

Таким образом, из (8) и (9) получим условия для импульсов в задаче с нефиксированным временем T и с заданной координатой $x_1(T)$:

$$\begin{aligned} p_1(T) &= \frac{\sum_{i=2}^{n} c_i f_i(T, x(T), u(T))}{f_1(T, x(T), u(T))} \qquad (f_1 \neq 0), \\ p_i(T) &= -c_i \qquad (i = 2, 3, \ldots, n). \end{aligned} \tag{10}$$

Принцип максимума для сформулированных задач утверждает (см. [1], [2]), что искомое оптимальное управление $u^*(t)$ доставляет функции H максимум по $u \in U$ при любом $t \in [t_0, T]$, если x и p — решения системы (1), (7) при краевых условиях (2), (8) (или (2), (10), когда время T определяется условием (5)). Из принципа максимума можно, вообще говоря, выразить u как функцию t, x и p и подставить ее в уравнения (1), (6). Получится краевая задача для системы $2n$ уравнений с $2n$ неизвестными (1), (7) при граничных условиях (2), (8) (или (2), (5), (10)).

2. Ниже описывается предлагаемый способ решения поставленной задачи (для определенности рассматривается задача с фиксированным временем). Задаем в качестве первого приближения некоторое допустимое управление $u^{(1)}(t)$ $(t_0 \leqslant t \leqslant T)$ (t) (выбор его может быть основан на каких-либо физических соображениях). Подставляя в систему (1) вместо u управление $u^{(1)}(t)$, решаем эту систему при начальных условиях (2) в интервале $[t_0, T]$; полученное решение обозначим через $x^{(1)}(t)$. Подставляя $u^{(1)}$ и $x^{(1)}$ в систему (7), интегрируем ее «справа налево» (от T до t_0) при начальных условиях (8); полученное решение обозначим через $p^{(1)}(t)$. Если в выражение (6) подставить $x^{(1)}(t)$, $p^{(1)}(t)$, то H будет функцией от t и u. Второе приближение для управления $u^{(2)}(t)$ определяем из условия максимальности H по U:

$$H(t, x^{(1)}(t), p^{(1)}(t), u^{(2)}(t)) = \max_{u \in U} H(t, x^{(1)}(t), p^{(1)}(t), u).$$

Теперь подставим $u^{(2)}(t)$ в (1) и найдем $x^{(2)}(t)$, затем находим $p^{(2)}(t)$ и т. д. Если процесс последовательных приближений сходится, то продолжаем его до тех пор, пока последующие приближения не будут отличаться друг от друга в пределах заданной точности.

Если момент T не фиксирован, а задано условие (5), то процесс последовательных приближений проводится так же, как и выше; разница состоит лишь в начальных условиях для p. Для определения момента T используется условие (5), а условия (8) заменяются условиями (10); функции x и u в (5) и (10) подставляются из рассматриваемого приближения.

Описанный метод очень прост с вычислительной стороны, так как на каждом шаге требует лишь решения двух задач Коши: «слева направо» для системы (1) и «справа налево» для системы (7).

Отметим, что так как предлагаемый метод использует принцип максимума, то с его помощью можно решать как классические вариационные задачи, в которых оптимальное управление непрерывно, так и задачи с разрывным оптимальным управлением (с точками переключения).

3. Вопрос об условиях сходимости этого процесса последовательных приближений остается открытым. Даже в случае сходимости, вообще говоря, неизвестно, является ли найденное управление оптимальным, так как принцип максимума дает лишь необходимое условие оптимальности. Однако для вариационных задач, существование и единственность решения которых ясны из физических соображений, таким методом можно находить оптимальные управления.

Рассмотрим вариационную задачу с фиксированным временем для линейной системы, в которой f — линейные функции x и u с зависящими от времени коэффициентами. Так как в этом случае x и u не входят в систему (7), то импульсы однозначно определяются этой системой и условиями (8). Из линейности f по u следует, что управление определяется из принципа максимума независимо от x. Подставляя это управление в систему (1) и решая ее при условиях (2), находим оптимальную траекторию. Нетрудно видеть, что применение предлагаемого метода к этой задаче во втором приближении даст точное решение. (В первом же приближении определяются точные значения импульсов и оптимального управления, а во втором — координат.) Это обстоятельство позволяет надеяться на то, что метод сходится и для некоторых нелинейных задач.

Сходимость итерационного процесса существенно зависит от выбора первого приближения. Если система (1) зависит от некоторого параметра λ и при $\lambda = \lambda_0$ найдено оптимальное управление, то его можно брать в качестве первого приближения для значений λ, близких к λ_0. Оказывается, что в ряде вариационных задач путем такого постепенного изменения параметра удается найти решение для большого интервала значений λ. Для улучшения сходимости можно прибегать к обычным способам улучшения сходимости итераций (см., например, [3]).

При практическом решении задач описанным методом можно хранить в памяти машины получающиеся при счете «слева направо» значения координат x в достаточно большом числе точек и пользоваться ими при интегрировании системы (7) для импульсов. Эти координаты также используются при вычислении нового приближения для управления u (из условия максимума функции Гамильтона (6)), которое, таким образом, необходимо запоминать при счете «справа налево». Вместо хранения координат, с точки зрения экономии машинной памяти, лучше интегрировать систему (1) вместе с системой (7) при счете «справа налево», задавая в качестве начальных условий $x(T)$ значения, полученные при решении системы (1) «слева направо». При этом в памяти машины хранятся только значения управления u. Наконец, возможен способ, когда в памяти машины не нужно хранить значения каких-либо функций. При этом интегрируется вся система в целом, причем на левом конце для k-го приближения задаются условия (2) для координат и полученные из предыдущего приближения значения импульсов $p^{(k-1)}(t_0)$, а на правом, соответственно, условия (8) (или (10)) для импульсов и $x^{(k)}(T)$ для координат. Первое приближение при этом нужно задать, как обычно, взяв какое-либо управление $u^{(1)}(t)$ и интегрируя только систему (1). В последующих приближениях управление находится из принципа максимума.

Отметим, что последний способ отличается от двух первых тем, что в нем всегда удовлетворяются уравнения и принцип максимума, но имеется невязка в граничных условиях на одном из концов, в то время как в двух первых удовлетворяются уравнения и граничные условия, но не удовлетворен принцип максимума. Невязка в выполнении принципа максимума и сводится к нулю в процессе итераций.

4. Приведем некоторые соображения, говорящие в пользу предложенного метода.

В работе Л. И. Розоноэра [2] выведена формула для приращения функционала (4), которое он получает при приращении управляющей функции δu в задаче с фиксированным временем T

$$\delta J \equiv \sum_{i=1}^{n} c_i \delta x_i (T) = - \int_{t_0}^{T} [H(t, x, p, u, + \delta u) - H(t, x, p, u)]\, dt - \varphi.$$

Для остаточного члена φ имеет место оценка [2]

$$|\varphi| \leqslant A \left(\int_{t_0}^{T} \sum_{k=1}^{m} |\delta u_k(t)|\, dt \right)^2 \equiv A \|\delta u\|^2,$$

где A — фиксированная константа, связанная с постоянными в условиях Липшица для функций f_i и $\partial H / \partial x_i$.

На каждом шаге предлагаемого процесса последовательных приближений новое управление $u + \delta u$ выбирается из условия максимальности H по u. Тогда из приведенной формулы для δJ видно, что на каждом шаге процесса главная часть приращения минимизируемого функционала: 1) неположительна: 2) максимальна по абсолютной величине среди главных частей других возможных приращений δJ, полученных за счет допустимых приращений управления δu. Другими словами, предлагаемый метод обеспечивает, с точностью до малых высшего порядка относительно приращения управления, изменение функционала в нужную сторону и притом наиболее быстрое по сравнению с другими методами, учитывающими лишь первую вариацию. Малость δu понимается здесь в смысле интегральной нормы, фигурирующей в оценке остаточного члена ($|\delta u|$ может быть конечной величиной, но на малом интервале времени).

В работе Л. И. Шатровского [4] предлагается другой метод решения задач оптимального управления, основанный на классической формуле для первой вариации функционала. При выводе этой формулы $|\delta u|$ предполагается малым всюду на интервале $[t_0, T]$. В вариационных задачах с ограничениями на управление часто приходится рассматривать разрывные управляющие функции и, следовательно, конечные значения $|\delta u|$, и поэтому классические формулы вариационного исчисления здесь непригодны.

Предлагаемый метод одинаковым образом применим как к классическим вариационным задачам, так и к задачам с ограничениями на управление.

Л. И. Шатровский [4] предлагает простой прием для приближенного учета граничных условий на правом конце (при $t = T$), а также ограничений на координаты x_i. Суть этого приема состоит в том, что вместо исходной задачи рассматривается задача со свободным правым концом, а к минимизируемому функционалу J добавляются члены, быстро возрастающие при нарушении ограничений или граничных условий. Этот прием, конечно, может использоваться и при применении метода, предлагаемого в данной работе. Так, если в фиксированный момент T заданы граничные условия $x(T) = X$, то минимизируемый функционал J'' вспомогательной задачи со свободным правым концом можно взять в виде

$$J' = J + \sum_{i=1}^{n} A_i [x_i(T) - X_i]^2,$$

где A_i — достаточно большие положительные числа (весовые коэффициенты).

5. В качестве иллюстрации применения метода рассмотрим модельную задачу о плоском движении тяжелого тела в сопротивляющейся среде. Процесс управления осуществляется при помощи двух управляющих функций, одна из которых непрерывна, а другая — релейного типа (эти функции вводятся ниже). Уравнения движения имеют вид

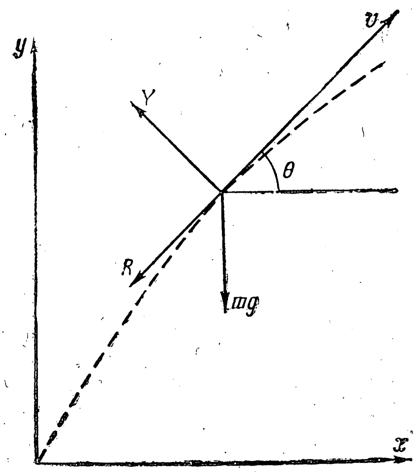
$$\frac{dx}{dt} = v \cos\theta, \qquad \frac{dy}{dt} = v \sin\theta, \tag{11}$$

$$m\frac{dv}{dt} = -R - mg \sin\theta, \qquad mv\frac{d\theta}{dt} = Y - mg\cos\theta.$$

Здесь t — время, x, y, v, θ — соответственно, дальность, высота, модуль и угол наклона скорости к оси x, m — масса тела, g — ускорение силы тяжести, R и Y — сила сопротивления и подъемная сила (фиг. 1). Как обычно, принимаем

$$R = \frac{1}{2}\rho v^2 S C_x, \qquad Y = \frac{1}{2}\rho v^2 S C_y,$$

где ρ, S, C_x, C_y — плотность среды (постоянная), характерная площадь тела и аэродинамические коэффициенты. Будем считать, что тело симметрично относительно продольной и поперечной осей. Тогда, очевидно, C_x является четной, а C_y — нечетной периодической функцией угла атаки α с периодом π (угол атаки — угол между вектором скорости и продольной осью тела). Для простоты положим



Фиг. 1

$$C_x = A - B\cos 2\alpha, \qquad A > B > 0,$$
$$C_y = C\sin 2\alpha, \qquad C > 0,$$

что означает аппроксимацию поляры (зависимости между C_x и C_y) эллипсом. При малых значениях угла атаки α эти формулы переходят в общепринятые (C_1, C_2, C_3— постоянные):

$$C_x = C_1 + C_2\alpha^2, \qquad C_y = C_3\alpha.$$

Вместо трех постоянных A, B, C введем максимальное качество $K = \max[C_y(\alpha)/C_x(\alpha)]$ и угол атаки α_0, при котором оно достигается. Включая третью постоянную множителем в S, окончательно получим

$$C_x = 1 - \cos 2\alpha_0 \cos 2\alpha,$$
$$C_y = K\sin 2\alpha_0 \sin 2\alpha. \tag{12}$$

Пусть в процессе движения можно изменять характерную площадь S, придавая ей одно из двух значений (в зависимости от величины η):

$$S = S_0(1 + \eta b), \quad S_0 > 0, \quad b > 0; \quad \eta = 0; 1.$$

Угол атаки α и параметр η считаем управляющими функциями, причем первая из них может принимать произвольные значения, а вторая — одно из двух допустимых значений.

Начальные условия берем в виде

$$t = 0, \qquad x = y = 0, \qquad v = v_0, \qquad \theta = \theta_0 > 0.$$

Вариационная задача заключается в определении такого закона управления $\alpha(t)$ и $\eta(t)$, при котором величина дальности полета $x(T)$ была бы максимальной в момент T, когда тело снова пересечет ось абсцисс (т. е. при $y(T) = 0$). Таким образом, это — вариационная задача со свободным правым концом и нефиксированным временем.

Перейдем к безразмерным переменным, выбрав в качестве единицы скорости начальную скорость v_0, единицы времени — величину v_0/g и единицы массы — m. Тогда система уравнений (11) примет следующий вид (безразмерные переменные обозначены теми же буквами, что и соответствующие им размерные):

$$\frac{dx}{dt} = v\cos\theta, \qquad \frac{dy}{dt} = v\sin\theta,$$
$$\frac{dv}{dt} = -\sigma v^2 C_x(1 + \eta b) - \sin\theta,$$
$$\frac{d\theta}{dt} = \sigma v C_y(1 + \eta b) - \frac{\cos\theta}{v}. \tag{13}$$

Начальные условия при этом преобразуются в

$$t = 0, \qquad x = 0, \qquad y = 0, \qquad v = 1, \qquad \theta = \theta_0. \tag{14}$$

Здесь σ — безразмерный параметр, характеризующий отношение аэродинамической силы к весу:

$$\sigma = \frac{\rho S v_0^2}{2mg}.$$

Запишем функцию Гамильтона для поставленной вариационной задачи:

$$H = p_x v\cos\theta + p_y v\sin\theta - p_v \sigma v^2 C_x(1+\eta b) - \\ - p_v \sin\theta + p_\theta \sigma v C_y (1+\eta b) - p_\theta \frac{\cos\theta}{v}. \tag{15}$$

Максимизируется функционал $J = x(T)$ при условии $y(T) = 0$. Поэтому начальные условия для импульсов получим в форме (10):

$$p_x(T) = 1, \quad p_v(T) = 0, \quad p_\theta(T) = 0, \quad p_y(T) = -\operatorname{ctg}\theta(T).$$

Так как H не зависит явно от x и y, то

$$p_x \equiv 1, \quad p_y \equiv -\operatorname{ctg}\theta(T).$$

Для импульсов p_v и p_θ остаются два уравнения:

$$\frac{dp_v}{dt} = -\cos\theta + \operatorname{ctg}\theta(T)\sin\theta + 2p_v\sigma v C_x(1+\eta b) - \\ - p_\theta \sigma C_y(1+\eta b) - p_\theta \frac{\cos\theta}{v^2}, \tag{16}$$

$$\frac{dp_\theta}{dt} = v\sin\theta + v\cos\theta\operatorname{ctg}\theta(T) + p_v\cos\theta - p_\theta\frac{\sin\theta}{v}$$

с начальными условиями

$$t = T, \quad p_v = 0, \quad p_\theta = 0. \tag{17}$$

Система уравнений (13), (16) при условиях (14), (17) решалась описанным выше методом, причем на каждом шаге процесса управляющие функции α, η находились из условия максимальности выражения

$$(1+\eta b)(p_\theta C_y - v p_v C_x), \tag{18}$$

полученного из H (15) отбрасыванием слагаемых и множителей, не зависящих от управления. Однако при $t = T$, в силу граничных условий, $p_v = p_\theta = 0$ и непосредственное определение α и η из (18) невозможно. Заметим, что из уравнений (16), с учетом естественного условия $0 > \theta(T) > (-\pi)/2$, следует для момента $t = T$

$$\frac{dp_v}{dt} = 0, \quad \frac{dp_\theta}{dt} = \frac{v}{\sin\theta} < 0.$$

Поэтому для достаточно малых отрицательных значений разности $t - T$ условие максимальности (18) приводится, с точностью до малых высшего порядка, к условиям $\eta = 1$ и $C_y(\alpha) = C_{y\,\max}$, что, в силу (12), соответствует значению $\alpha(T) = \pi/4$. Эти условия вполне понятны физически: для достижения максимальной дальности в последний момент выгоднее увеличить максимально подъемную силу, не заботясь о сопротивлении.

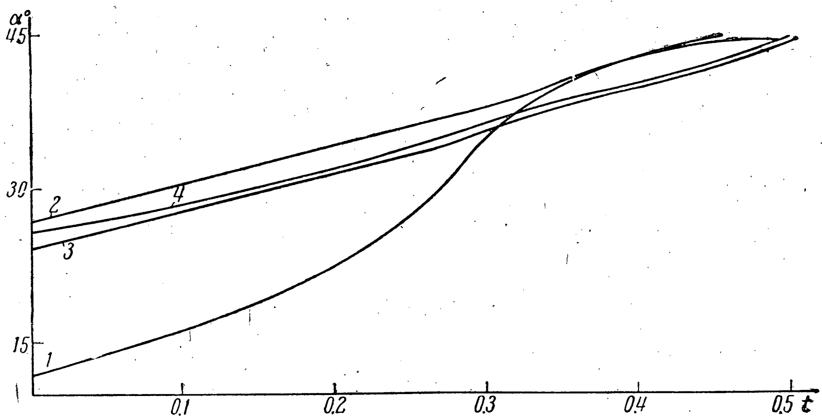
Ниже приводятся результаты численного решения поставленной задачи на электронной вычислительной машине для значений параметров

$$b = 0.2, \quad \alpha_0 = 10°, \quad \theta_0 = 10°.$$

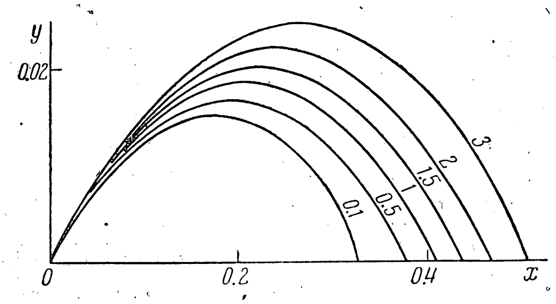
Параметры K и σ варьировались в широких пределах, причем в качестве первого приближения для какого-либо варианта задавалось оптимальное управление, найденное для предыдущего варианта с близкими значениями K и σ.

На фиг. 2 изображен типичный пример сходимости процесса последовательных приближений. На ней представлены зависимости управления $\alpha(t)$ для одного из вариантов, причем номер кривой указывает номер приближения. Пятое и последующие приближения практически не отличались от четвертого. Значения максимизируемого функционала (дальности) для последовательных итераций были следующими:

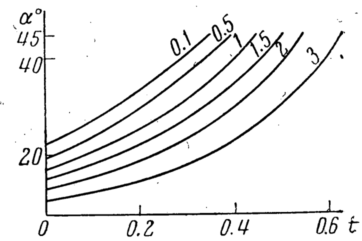
$$x_1 = 0.426, \quad x_2 = 0.441, \quad x_3 = 0.450, \quad x_4 = 0.451.$$



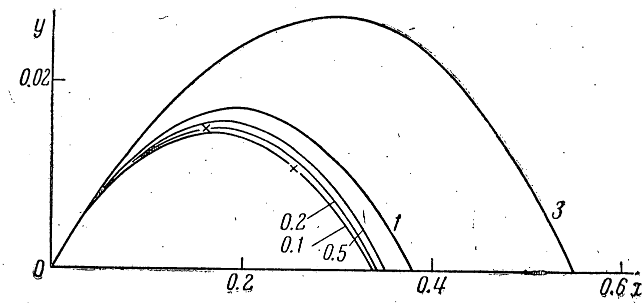
Фиг. 2



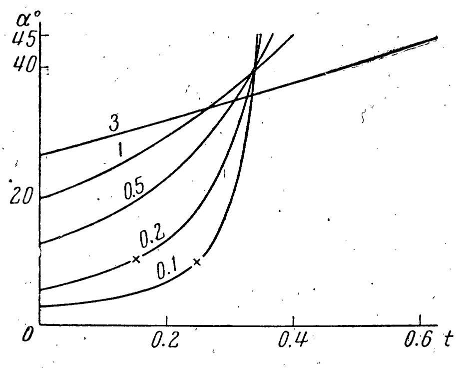
Фиг. 3



Фиг. 4



Фиг. 5



Фиг. 6

На фиг. 3 представлены оптимальные траектории при различных значениях σ (цифры при кривых) и фиксированном значении максимального качества $K = 1$. Оптимальные управления α (t) для этой серии вариантов даны на фиг. 4. С увеличением σ (отношения аэродинамических сил к весу) оказывается выгодным уменьшать углы атаки, проигрывая в подъемной силе, но зато добиваясь уменьшения лобового сопротивления. Дальность полета возрастает с увеличением σ. Для всех приведенных здесь вариантов оказалось $\eta (t) \equiv 1$, что соответствует наибольшему из двух возможных значений характерной площади.

Серия оптимальных траекторий при фиксированном значении $\sigma = 0.5$ и различных значениях максимального качества K, указанных на графике, изображена на фиг. 5. Дальность полета, естественно, возрастает с увеличением качества. На фиг. 6 представлены соответствующие оптимальные управления. При малых значениях качества K движение происходит на малых углах атаки. Кроме того, здесь на начальной части траектории (до точек, указанных звездочкой для $K = 0.1$ и $K = 0.2$) управляющая функция $\eta = 0$ (характерная площадь минимальна). Оба этих фактора уменьшают силу сопротивления (они одновременно уменьшают и подъемную силу, но она несущественна при малых K). Для больших K, наоборот, сила сопротивления становится менее существенной по сравнению с подъемной силой. Поэтому здесь движение по оптимальной траектории происходит при больших углах атаки и с $\eta \equiv 1$.

Авторы выражают глубокую благодарность Н. Н. Моисееву за ценные советы и обсуждения, В. Г. Сраговичу и В. И. Войтинскому, прочитавшим рукопись и сделавшим ряд полезных замечаний.

Поступила в редакцию
14. 05. 1962

Цитированная литература

1. Л. С. Понтрягин, В. Г. Болтянский, Р. В. Гамкрелидзе, Е. Ф. Мищенко. Математическая теория оптимальных процессов. М., Физматгиз, 1961.
2. Л. И. Розоноэр. Принцип максимума Л. С. Понтрягина в теории оптимальных систем. I, II. Автоматика и телемеханика, 1959, **20**, № 10, 1320—1334; № 11, 1441—1458.
3. И. С. Березин, Н. П. Жидков. Методы вычислений. Т. II. М., Физматгиз, 1959.
4. Л. И. Шатровский. Об одном численном методе решения задач оптимального управления. Ж. вычисл. матем. и матем. физ., 1962, **2**, № 3, 488—491.

СПОСОБ ПОСЛЕДОВАТЕЛЬНОГО ПОИСКА ДЛЯ ПРИБЛИЖЕННОГО РЕШЕНИЯ НЕКОТОРЫХ ЗАДАЧ ОПТИМАЛЬНОГО ПРОЕКТИРОВАНИЯ

И. Б. МОЦКУС

(*Каунас*)

§ 1. Введение

Основные технико-экономические показатели любого производственного объекта определяются заданием некоторого числа независимых переменных, которые назовем параметром состояния объекта U.

Совокупность таких значений параметра U, осуществление которых возможно имеющимися в распоряжении средствами, назовем множеством возможных состояний объекта D. Предполагается, что множеству D принадлежит конечное число дискретных значений U.

Каждому значению параметров $U \in D$ (каждому возможному состоянию объекта) соответствует некоторая величина F, которую назовем целевой функцией.