ЗАДАЧА ПАРАМЕТРИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Общая постановка задачи параметрического программирования

Пусть $x \in E^n$ – вектор переменных и $u \in E^l$ – вектор параметров являются элементами конечномерных евклидовых пространств соответственно с координатными представлениями $\| x \| = \left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \cdots \\ \xi_n \end{matrix} \right\|$ и $\| u \| = \left\| \begin{matrix} \nu_1 \\ \nu_2 \\ \cdots \\ \nu_l \end{matrix} \right\|$.

Рассмотрим следующую задачу, которую принято называть *задачей параметрического программирования*:

найти $\max_x F(x,u)$ (1)

при условиях: $f_s(x,u) \le 0,\ s = [1,m]$, (2)

и пусть x_u^* есть решение задачи (1)−(2) для некоторого фиксированного $u \in \Omega \subseteq E^l$.

Любую задачу, в формулировке которой используется x_u^*, будем называть задачей *в пространстве параметров* или параметрической задачей *второго уровня*.

Например, задачу

$$\max_u F(x_u^*, u) \text{ при условии } u \in \Omega. \tag{3}$$

В отличие от задач типа (3), задачу (1)−(2) будем называть задачей "*первого уровня*".

Следует отметить, что, хотя система задач (1)−(2) и (3) сводится к задаче математического программирования вида

$$\max F(x,u) \tag{4−5}$$
$$\text{при условиях}$$
$$f_s(x,u) \le 0,\ s = [1,m],\ (x,u) \in E^n \otimes E^l,$$

такое сведение может приводить к неприемлемому уровню усложнения задачи.

В качестве иллюстрации приведем следующий пример такой пары.

1. Задача *первого уровня* - задача линейного программирования с нелинейно входящими в ее условие параметрами, для которой $\|x\| = \begin{Vmatrix} \xi_1 \\ \xi_2 \end{Vmatrix}$ и $\|u\| = \begin{Vmatrix} \nu_1 \\ \nu_2 \end{Vmatrix}$ – координатные представления векторов $x \in E^2$ и $u \in E^2$.

$$\text{Найти } \max_x 2\xi_1 + 3\xi_2 \tag{6}$$

$$\begin{gathered} \text{при условиях: } \xi_1 \geq 0, \xi_2 \geq 0, \\ \xi_1 + \nu_1 \xi_2 \leq 6, \\ 2\xi_1 + \xi_2 \leq \nu_2, \end{gathered} \tag{7}$$

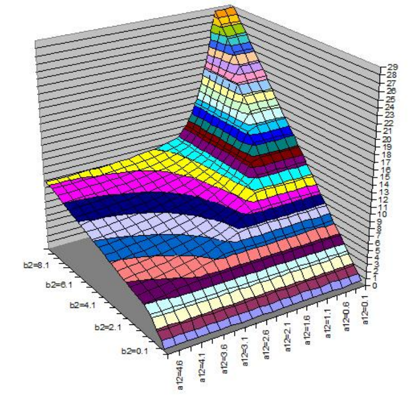
решение которой представляется зависимостями $\xi^*_{\nu_1,\nu_2}$ и $\xi^*_{\nu_1,\nu_2}$.

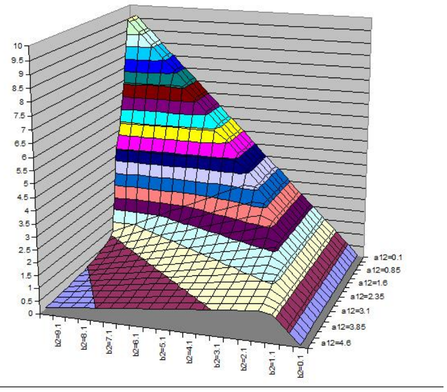
2. Задача *второго уровня*:

$$\max_u 2\xi^*_{1,\nu_1,\nu_2} + 3\xi^*_{2\,\nu_1,\nu_2} \tag{8}$$

при условиях:

$$0.1 \leq \nu_1 \leq 5, \quad 0.1 \leq \nu_2 \leq 10\,. \tag{9}$$





На этих рисунках приведены графические представления зависимостей

$$\max_u 2\xi^*_{1,\nu_1,\nu_2} + 3\xi^*_{2\,\nu_1,\nu_2} \qquad \text{и} \qquad \xi^*_{2\,\nu_1,\nu_2} \text{ от } \{\nu_1, \nu_2\}$$

– компонент вектора $u \in \Omega$, которые позволяют заключить, что данные зависимости непрерывные, нелинейные, невыпуклые и не дифференцируемые для всех $u \in \Omega$, а задача второго уровня имеет неединственное решение.

Свойства решений задач параметрического программирования

Пусть x_u^* есть решение задачи (1)−(2) для фиксированного u, а задача второго уровня сформулирована в виде

$$\max_u F(x_u^*,u) \quad u \in \Omega, \tag{10}$$

где $F(x,u)$ – некоторая функция, зависящая как от $x \in E^n$, так и от $u \in E^l$.

Как постановка, так и процедура решения задачи (10) *могут* в значительной степени осложняться следующими специфическими свойствами зависимости x_u^*.

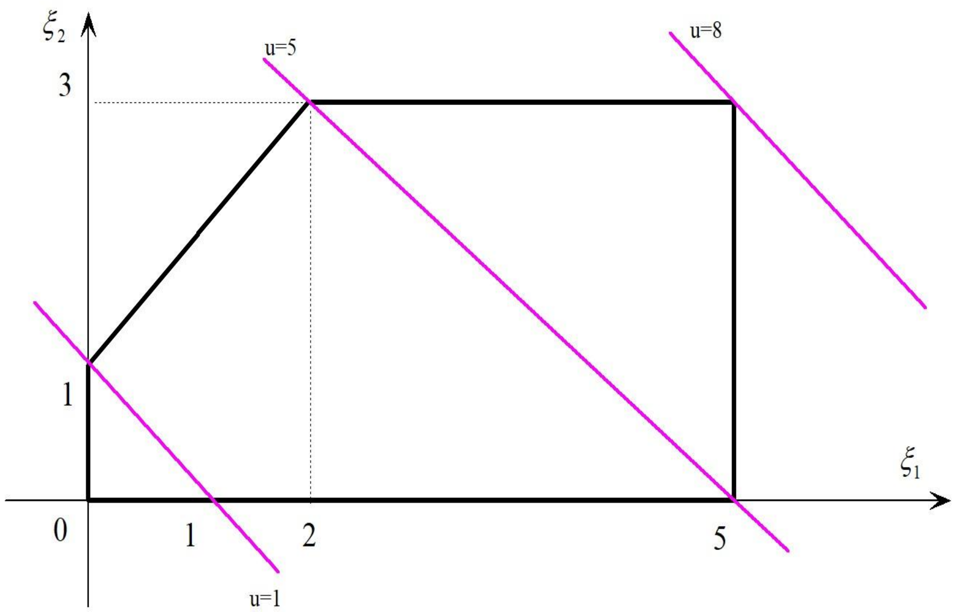
1. Практической *невозможностью* (за исключением, быть может, некоторых тривиальных случаев) аналитического решения задачи первого уровня (1)−(2), а, значит, также и постановки, исследования и решения в явном виде задачи второго уровня (10).
2. Несовпадением в общем случае области определения зависимости x_u^* и множества Ω, поскольку система условий задачи первого уровня (1)−(2) может оказаться *противоречивой* для некоторых $u \in \Omega \subseteq E^l$.
3. *Нефункциональностью* (неоднозначностью) зависимости x_u^* для тех $u \in \Omega \subseteq E^l$, при которых задача первого уровня (1)−(2) имеет решение, но не единственное.
4. *Негладкостью* зависимости x_u^* в силу того, что условия задачи первого уровня (1)−(2) могут содержать ограничения типа *неравенство*. Более того, даже существование непрерывных производных у функций $f(x,u)$ и $y_s(x,u)$ достаточно высокого порядка не гарантирует необходимой гладкости (а иногда даже и непрерывности) зависимости x_u^* и, следовательно, входящих в формулировку задачи второго уровня условий.

Эти свойства решений задач параметрического программирования могут существенно усложнить процедуру их поиска.

Причины, порождающие подобные свойства можно проиллюстрировать следующим примером.

Рассмотрим следующую задачу параметрического программирования: $\forall u \in \mathbf{R}$ максимизировать по $x = \begin{Vmatrix} \xi_1 \\ \xi_2 \end{Vmatrix} \in E^2$ функцию $F = \xi_2$

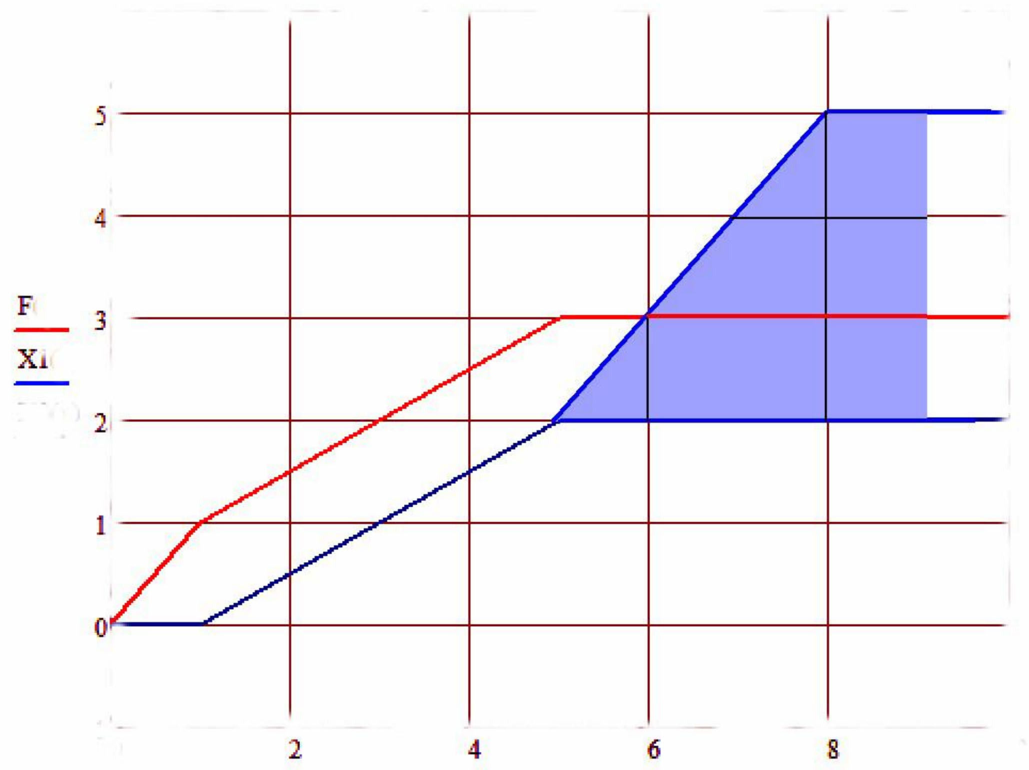
при условиях: $\begin{matrix} 0 \le \xi_1 \le 5, & 0 \le \xi_2 \le 3, \\ \xi_1 - \xi_2 \le -1, & \xi_1 + \xi_2 \le u\,. \end{matrix}$



Геометрическая интерпретация задачи параметрического программирования.

Решение этой задачи представляется такими зависимостями

u	F_u^*	ξ_{1u}^*	ξ_{2u}^*
$(-\infty, 0)$	*не сущ.*	*не сущ*	*не сущ*
$[0, 1)$	u	0	u
$[1, 5)$	$\frac{u+1}{2}$	$\frac{u-1}{2}$	$\frac{u+1}{2}$
$[5, 8]$	3	$\forall[2, u-3]$	3
$(8, +\infty)$	3	$\forall[2, 5]$	3



Использование метода штрафных функций для построения сглаженных аппроксимаций решений задач первого уровня

Идея этого подхода заключается в замене зависимости x_u^* достаточно гладкой и определенной для всех $u \in \Omega \subseteq E^l$ и любых положительных τ функцией $x(\tau, u)$, такой что $\lim\limits_{\tau \to +0} x(\tau, u) = x_u^*$,

В качестве аппроксимирующей зависимости $x(\tau, u)$ предлагается использовать решение задачи первого уровня, получаемое методом *гладких штрафных функций*.

Как известно, решением задачи первого уровня (1)−(2) при использовании метода штрафных функций является

$$x(\tau,u) = \arg\max_x A(\tau,x,u)\,, \tag{11}$$

где вспомогательная функция метода гладких штрафных функций выбирается следующего вида:

$$A(\tau,x,u) = F(x,u) - \sum_{i=1}^{m} P[\tau,\, f_i(x,u)]\ , \tag{12}$$

uде

$$\lim_{\tau\to+0} P(\tau,s) = \begin{bmatrix} +\infty,\ s>0, \\ 0,\ s<0 \end{bmatrix}$$

достаточно гладкая штрафная функция $P(\tau,s)$ удовлетворяет также и неравенствам

$$\frac{\partial P}{\partial s} > 0 \qquad \text{и} \qquad \frac{\partial^2 P}{\partial s^2} > 0 \quad \forall s. \tag{13}$$

Тогда $x(\tau,u)$ – точку максимума вспомогательной функции A можно найти из условия *стационарности* этой вспомогательной функции

$$\operatorname{grad}_x A(\tau,x,u) = o\,. \tag{14}$$

Теоретической основой предлагаемого подхода является использование локальных тейлоровских аппроксимаций зависимостей и теоремы о неявных функциях.

Если штрафная функция $P(\tau, s)$ строго выпукла по s для любых $\tau > 0$ и, по крайней мере, дважды непрерывно дифференцируема по всем своим аргументам, то оказываются справедливыми следующие свойства.

1. Решение уравнений (14), являющихся условиями стационарности функции (12), $\bar{x}(\tau, u)$, существует и локально единственно для любых $u \in \Omega \subseteq E^l$ и $\tau > 0$, и потому зависимость $\bar{x}(\tau, u)$ является функциональной.

2. Для зависимости $\bar{x}(\tau, u)$ верно равенство $\lim\limits_{\tau \to +0} \bar{x}(\tau, u) = x_u^*$.

3. Если, наконец, функции $F(x, u)$ и $f_i(x, u)$ имеют непрерывные частные производные до второго порядка включительно, то из теоремы о *неявных функциях*, примененной к условиям стационарности (14), следует непрерывная дифференцируемость функции $\bar{x}(\tau, u)$ по всем ее аргументам.

Принципиальную применимость рассматриваемого подхода продемонстрируем на следующем примере.

Максимизировать 3ξ при условиях $0 \le \xi \le u \quad \forall u \in \mathbf{R}$.

Построим вспомогательную функцию, использовав штрафную функцию:

$$A(\tau, \xi, u) = 3\xi - \tau e^{-\frac{\xi}{\tau}} - \tau e^{\frac{\xi - u}{\tau}} \tag{15}$$

Условие стационарности функции (15) по ξ для этой вспомогательной функции будет

$$3 + e^{-\frac{\xi}{\tau}} - e^{\frac{\xi - u}{\tau}} = 0 \qquad \Rightarrow \qquad \xi(\tau, \nu) = \tau \cdot \ln \frac{2}{\sqrt{9 + 4e^{-\frac{u}{\tau}}} - 3} \tag{16}$$

График зависимости $\xi(\tau,u)$ (формулы (16) для разных значений τ) показан на следующем рисунке. Здесь $B(u)=\lim\limits_{\tau\to+0}A(\tau,u)\quad u\geq 0$.

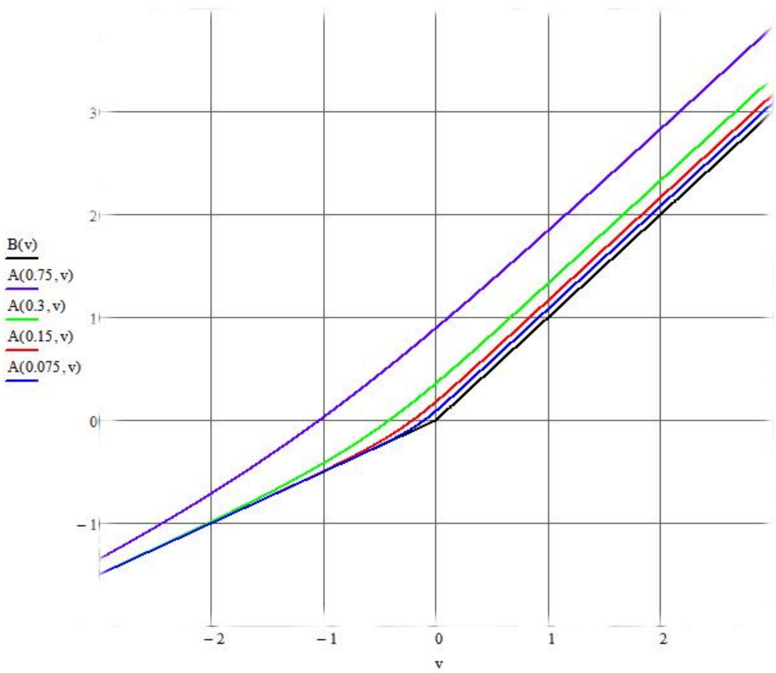


Схема решения задач параметрического программирования

Как известно, основное свойство метода штрафных функций состоит в том, что

$$\lim_{\tau \to +0} A(\tau, x(\tau, u), u) = F(x_u^*, u),$$

где x_u^* − решение задачи (1)−(2). Поэтому в качестве целевого функционала, итерационная оптимизация которого позволяет получить оценку решения задачи второго уровня, можно принять $A(\tau, x(\tau, u), u)$.

Как было отмечено ранее, основным затруднением при реализации процедуры вида

$$u_{k+1} = u_k + \sigma_k w_k,\ k = 0, 1, 2, \ldots \tag{17}$$

является оценка w_k – направления улучшающей вариации и σ_k – величины шага по данному направлению, а главным преимуществом − возможность использования численных методов, не требующих явного аналитического представления для

$$E(\tau, u) = A(\tau, x(\tau, u), u).$$

Далее, если решение задачи второго уровня сведено к оптимизации функционала $E(\tau,u)$, то при фиксированном $\tau > 0$ для определения w_k можно применить, например, метод Ньютона.

Данный метод, как известно, использует квадратичную тейлоровскую аппроксимацию, требующую значений как оптимизируемого функционала, так и его градиента и гессиана в точке $u_k \in \Omega \subset E^l$. Поскольку в данном случае функция $\bar{x}(r,u)$ определена неявно уравнениями

$$\frac{\partial A}{\partial \xi_i} = 0 \qquad \text{или} \qquad \frac{\partial f}{\partial \xi_i} + \sum_{s=1}^{m} \frac{\partial P}{\partial y_s} \frac{\partial y_s}{\partial \xi_i} = 0 \quad \forall i = [1,n], \tag{18}$$

то для компонент градиента $E(r,u)$ мы имеем

$$E'_{v_j} = \frac{\partial A}{\partial v_j} + \sum_{i=1}^{n} \frac{\partial A}{\partial \xi_i} \frac{\partial \bar{\xi}_i}{\partial v_j}, \ \forall j = [1,l],$$

а в силу (18) окончательно получаем

$$E'_{v_j} = \frac{\partial A}{\partial v_j}(\tau, \bar{x}(u), u) \quad \forall j = [1,l]. \tag{19}$$

Формула для компонент матрицы Гессе, будет несколько более сложной:

$$E''_{\nu_j \nu_t} = \frac{\partial^2 A}{\partial \nu_j \partial \nu_t} + \sum_{i=1}^{n} \frac{\partial^2 A}{\partial \xi_i \partial \nu_j} \frac{\partial \overline{\xi}_i}{\partial \nu_t} \qquad \forall j, t = [1, l]. \tag{20}$$

Необходимые для использования (20) значения $\frac{\partial \overline{\xi}_i}{\partial \nu_t}$ $\quad \forall i = [1, n], \forall j = [1, l]$ можно найти, решив систему линейных уравнений:

$$\sum_{i=1}^{n} \frac{\partial^2 A}{\partial \xi_i \partial \xi_j} \frac{\partial \overline{\xi}_i}{\partial \nu_t} = -\frac{\partial^2 A}{\partial \xi_j \partial \nu_t} \qquad \forall j, t = [1, l], \tag{21}$$

которая получается при дифференцировании (18) по всем компонентам $u \in E^l$.

Наконец, оценки w_k и σ_k находим, например, по формулам

$$\| w_k \| = -\| E'' \|^{-1} \| E' \| \qquad \text{и} \qquad \sigma_k = \arg\min_{\sigma > 0} E(u_k + \sigma w_k).$$