

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 $\|x\| = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \dots \\ \xi_n \end{pmatrix}$ и $\|u\| = \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_l \end{pmatrix}$.

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

$$\text{найти } \max_x F(x, u) \quad (1)$$

$$\text{при условиях: } f_s(x, u) \leq 0, s = [1, m], \quad (2)$$

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

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

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

$$\begin{aligned} &\max F(x, u) \\ &\text{при условиях} \\ &f_s(x, u) \leq 0, s = [1, m], (x, u) \in E^n \otimes E^l, \end{aligned} \quad (4-5)$$

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 $\|x\| = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$ и $\|u\| = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ – координатные представления векторов $x \in E^2$ и $u \in E^2$.

$$\text{Найти } \max_x 2\xi_1 + 3\xi_2 \quad (6)$$

при условиях: $\xi_1 \geq 0, \xi_2 \geq 0,$

$$\xi_1 + v_1 \xi_2 \leq 6,$$

$$2\xi_1 + \xi_2 \leq v_2, \quad (7)$$

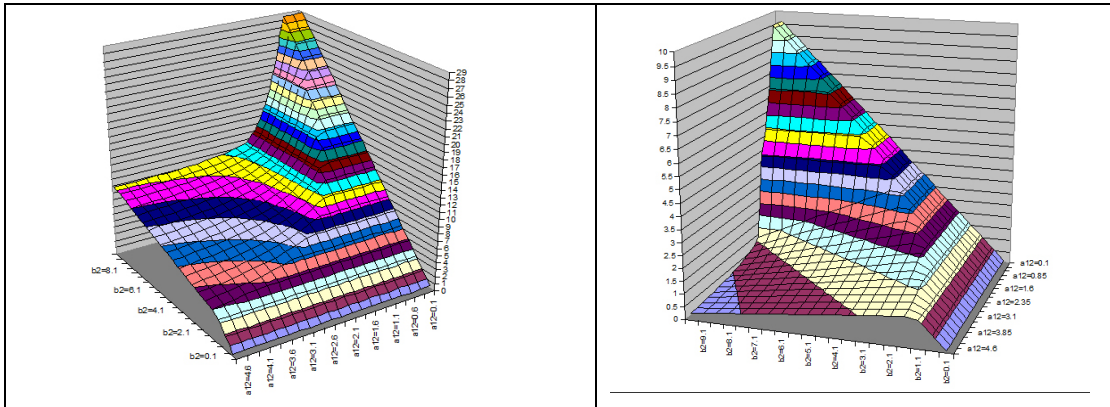
решение которой представляется зависимостями ξ_{v_1, v_2}^* и ξ_{v_1, v_2}^* .

2. Задача *второго уровня*:

$$\max_u 2\xi_{1, v_1, v_2}^* + 3\xi_{2, v_1, v_2}^* \quad (8)$$

при условиях:

$$0.1 \leq v_1 \leq 5, \quad 0.1 \leq v_2 \leq 10. \quad (9)$$



На этих рисунках приведены графические представления зависимостей

$$\max_u 2\xi_{1,v_1,v_2}^* + 3\xi_{2,v_1,v_2}^* \quad \text{и} \quad \xi_{2,v_1,v_2}^* \quad \text{от} \quad \{v_1, v_2\}$$

— 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

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

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.

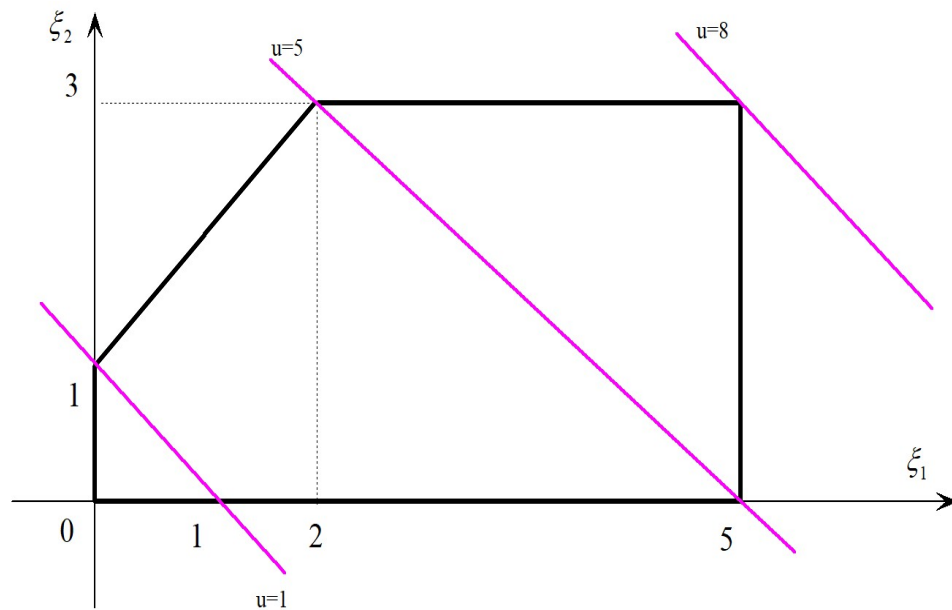
Эти свойства решений задач параметрического программирования могут существенно усложнить процедуру их поиска.

Причины, порождающие подобные свойства можно проиллюстрировать следующим примером.

Рассмотрим следующую задачу параметрического программирования:

$\forall u \in \mathbf{R}$ максимизировать по $x = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \in E^2$ функцию $F = \xi_2$

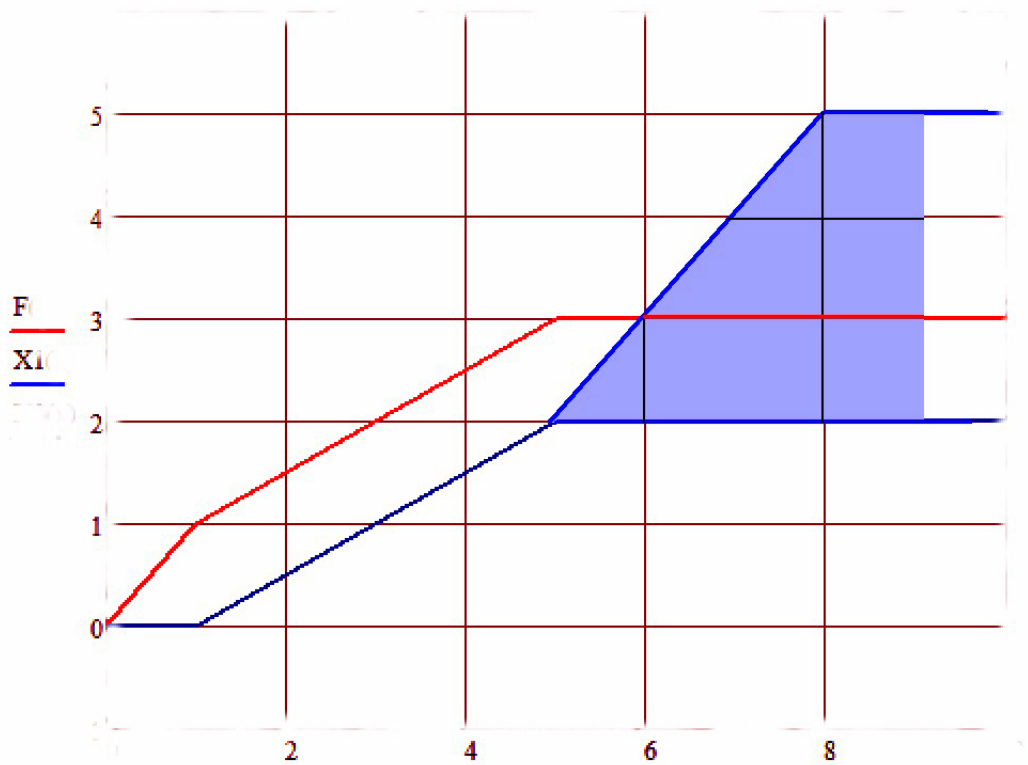
при условиях: $0 \leq \xi_1 \leq 5, \quad 0 \leq \xi_2 \leq 3,$
 $\xi_1 - \xi_2 \leq -1, \quad \xi_1 + \xi_2 \leq u.$



Геометрическая интерпретация задачи параметрического программирования.

Решение этой задачи представляется такими зависимостями

u	F_u^*	ξ_{1u}^*	ξ_{2u}^*
$(-\infty, 0)$	не суц.	не суц	не суц
$[0, 1)$	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 $u \in \Omega \subseteq E'$ и любых положительных τ функцией $\bar{x}(\tau, u)$, такой что $\lim_{\tau \rightarrow +0} \bar{x}(\tau, u) = x_u^*$,

В качестве аппроксимирующей зависимости $\bar{x}(\tau, u)$ 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

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

где вспомогательная функция метода гладких штрафных функций выбирается следующего вида:

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

где

$$\lim_{\tau \rightarrow +0} P(\tau, s) = \begin{cases} +\infty, & s > 0, \\ 0, & s < 0 \end{cases}$$

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

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

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

$$\text{grad}_x A(\tau, \bar{x}, u) = o. \quad (14)$$

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 $\lim_{\tau \rightarrow +0} \bar{x}(\tau, u) = x_u^*$.
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 $0 \leq \xi \leq u \quad \forall u \in \mathbf{R}$.

Построим вспомогательную функцию, используя штрафную функцию:

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

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

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

Нетрудно также показать, что имеют место предельные равенства:

$$\lim_{\tau \rightarrow +0} \bar{x}(\tau, u) = u \quad \text{при } u > 0 \quad \text{и} \quad \lim_{\tau \rightarrow +0} \bar{x}(\tau, u) = \frac{u}{2} \quad \text{при } u < 0.$$

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

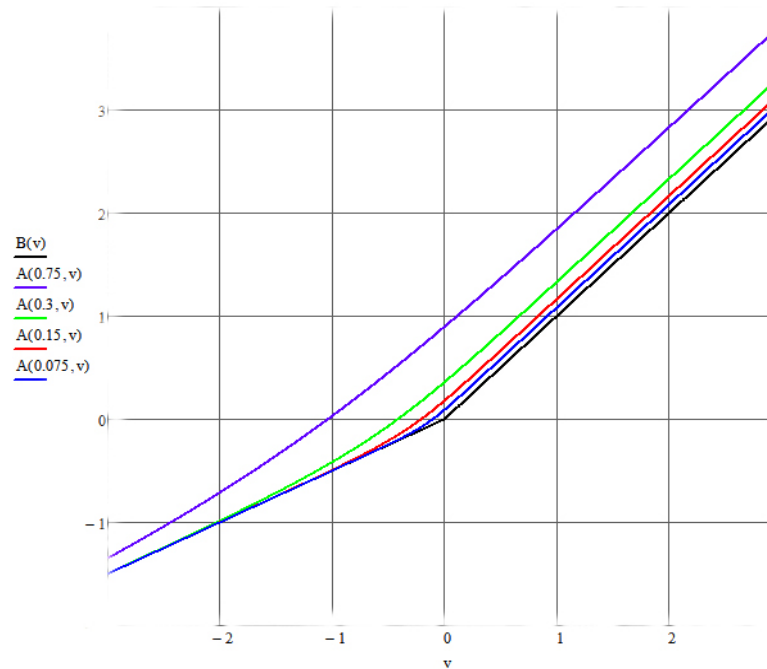


Схема решения задач параметрического программирования

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

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

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

Как было отмечено ранее, основным затруднением при реализации процедуры вида

$$u_{k+1} = u_k + \sigma_k w_k, k = 0, 1, 2, \dots \quad (17)$$

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

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

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

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

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

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

а в силу (18) окончательно получаем

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

Формула для компонент матрицы Гессе, будет несколько более сложной:

$$E''_{v_j v_t} = \frac{\partial^2 A}{\partial v_j \partial v_t} + \sum_{i=1}^n \frac{\partial^2 A}{\partial \xi_i \partial v_j} \frac{\partial \bar{\xi}_i}{\partial v_t} \quad \forall j, t = [1, l]. \quad (20)$$

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

$$\sum_{i=1}^n \frac{\partial^2 A}{\partial \xi_i \partial \xi_j} \frac{\partial \bar{\xi}_i}{\partial v_t} = - \frac{\partial^2 A}{\partial \xi_j \partial v_t} \quad \forall j, t = [1, l], \quad (21)$$

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

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

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