Простые и сложные задачи глобальной оптимизации

О. В. Хамисов

(Институт систем энергетики им. Л.А. Мелентьва СО РАН, Иркутск)

МФТИ, 19 октября 2018

Оптимизация: основные моменты

- дифференцируемость
- структурные свойства

Оптимизация: дифференцируемость

Задача безусловной минимизации

$$f(x) \to \min,$$

$$x \in \mathbb{R}^n.$$

Необходимое условие оптимальности (теорема Ферма)

$$f'(x) = 0 \ \ (\nabla f(x) = 0).$$

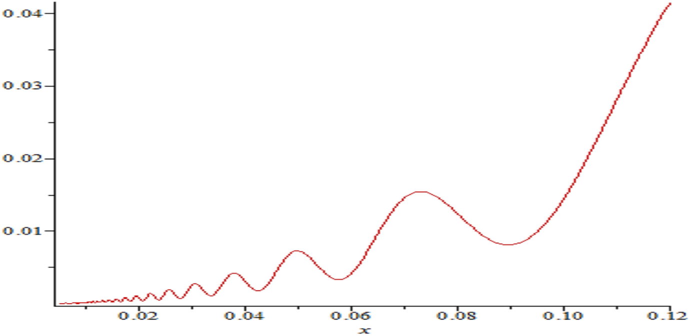
$$\Downarrow$$

Необходима дифференцируемость (=гладкость) f

Оптимизация: дифференцируемость

Пример.

$$f(x) = \begin{cases} 2x^2 + x^2 \sin(\frac{1}{x}), & x \neq 0, \\ 0, & x = 0. \end{cases}$$



$$x^* = 0, \ f(x^*) = 0$$

Оптимизация: дифференцируемость

Пример. Производная

$$f'(x) = \begin{cases} 4x + 2x\sin(\frac{1}{x}) - \cos(\frac{1}{x}), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Две последовательности, сходящиеся к решению,

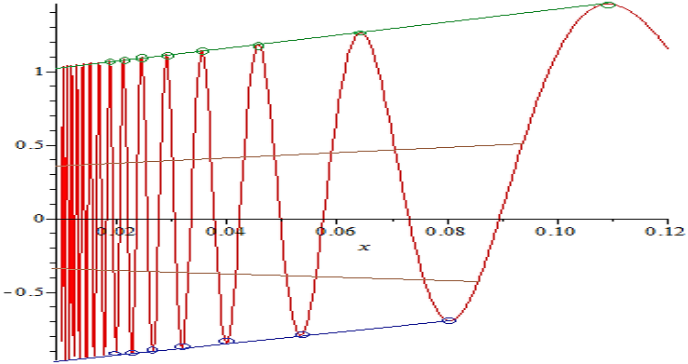
$$x_k = \frac{1}{\pi + 2\pi k} \to 0, \;\; y_k = \frac{1}{2\pi k} \to 0.$$

Тем не менее:

$$f'(x_k) \to 1 \neq 0, \;\; f'(y_k) \to -1 \neq 0!$$

Оптимизация: дифференцируемость

Пример. Геометрическая интерпретация $f'(x)$



$f'(x_k) \to 1$, $f'(y_k) \to -1$

Оптимизация: дифференцируемость

Пример. f дифференцируема, но f' не является непрерывной.

Оптимизация: дифференцируемость

Пример. f дифференцируема, но f' не является непрерывной.

$$\Downarrow$$

С методической точки зрения f должна быть *непрерывно* дифференцируемой, $f \in C^1$.

Оптимизация: дифференцируемость

Пример. f дифференцируема, но f' не является непрерывной.

$$\Downarrow$$

С методической точки зрения f должна быть *непрерывно* дифференцируемой, $f \in C^1$.

$$\Downarrow$$

Любая предельная точка метода наискорейшего спуска для $f \in C^1$ является стационарной (удовлетворяет теореме Ферма)

Оптимизация: дифференцируемость

Задача безусловной минимизации

$$f(x) \to \min,$$

$$x \in \mathbb{R}^n,$$

$f \in C^1$.

Методический подход
Итеративная информация: $f(x^k)$, $f'(x^k)$, $k = 0, 1, \dots$

Гарантия: нахождение стационарной точки x^*: $f'(x^*) = 0$[1].

Вопрос: на сколько сильно отличается значение $f(x^*)$ от (глобально) минимального значения?

[1] Если такая точка существует.

Оптимизация: дифференцируемость

Задача безусловной минимизации

$$f(x) \to \min,$$

$$x \in \mathbb{R}^n,$$

$f \in C^1$.

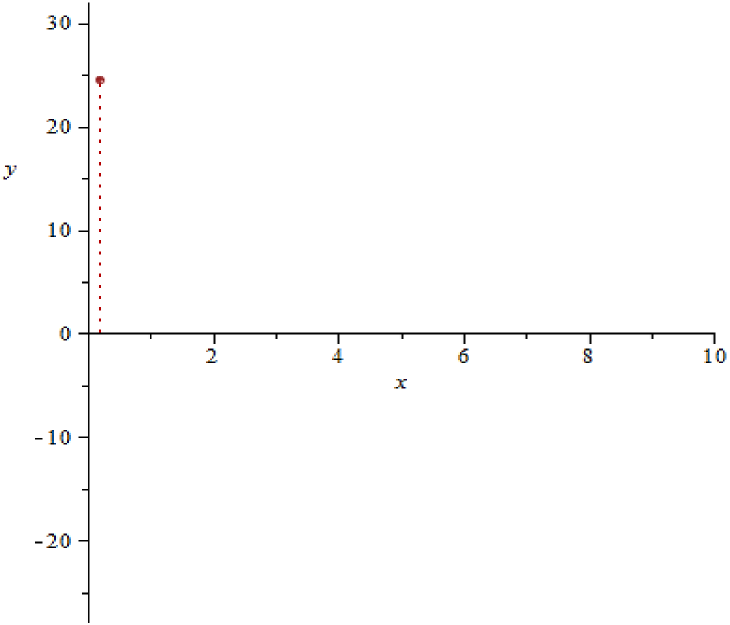
Методический подход
Итеративная информация: $f(x^k)$, $f'(x^k)$, $k = 0, 1, \ldots$

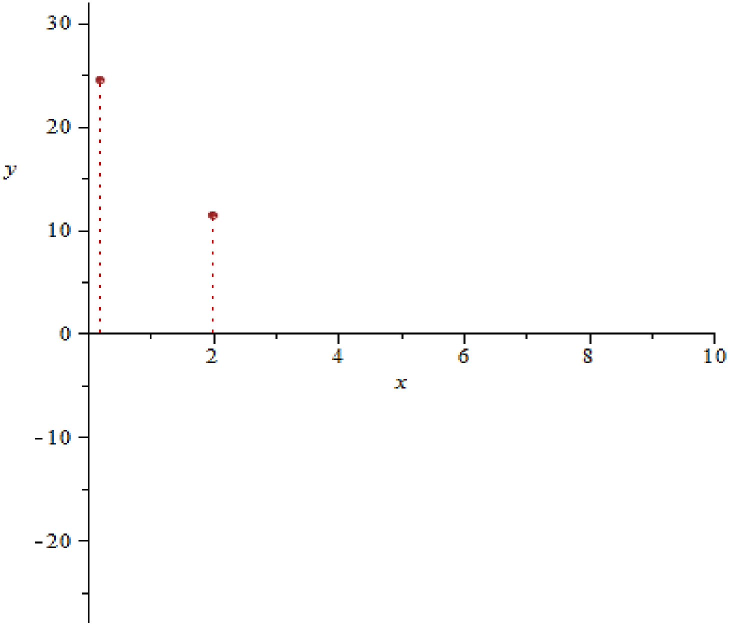
Гарантия: нахождение стационарной точки x^*: $f'(x^*) = 0$[1].

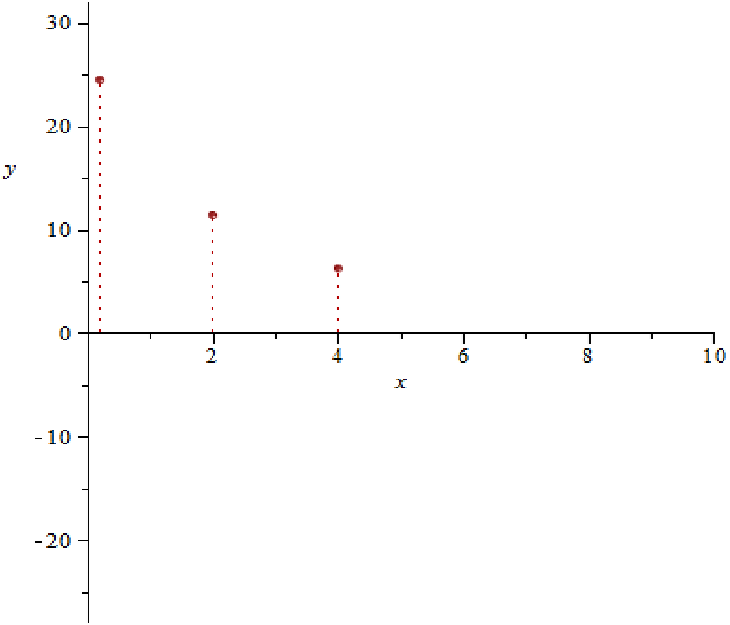
Вопрос: на сколько сильно отличается значение $f(x^*)$ от (глобально) минимального значения?

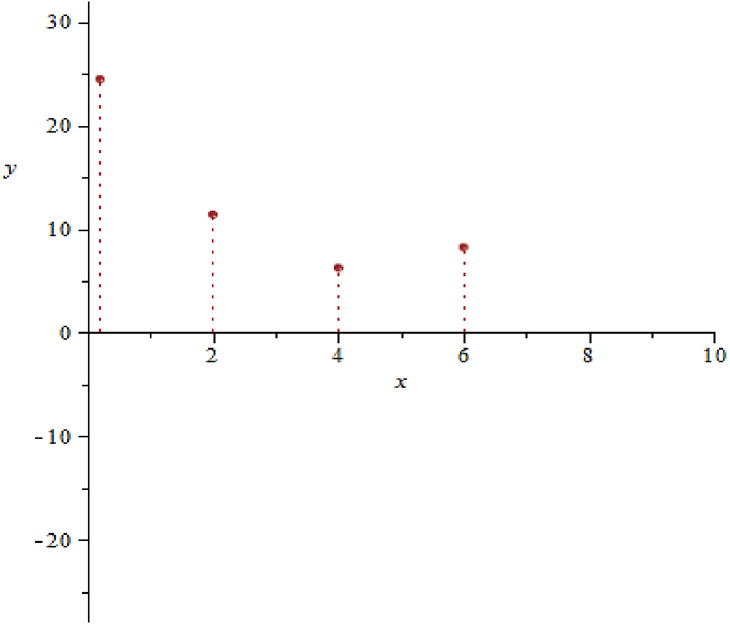
Ответ: на сколько угодно!!!

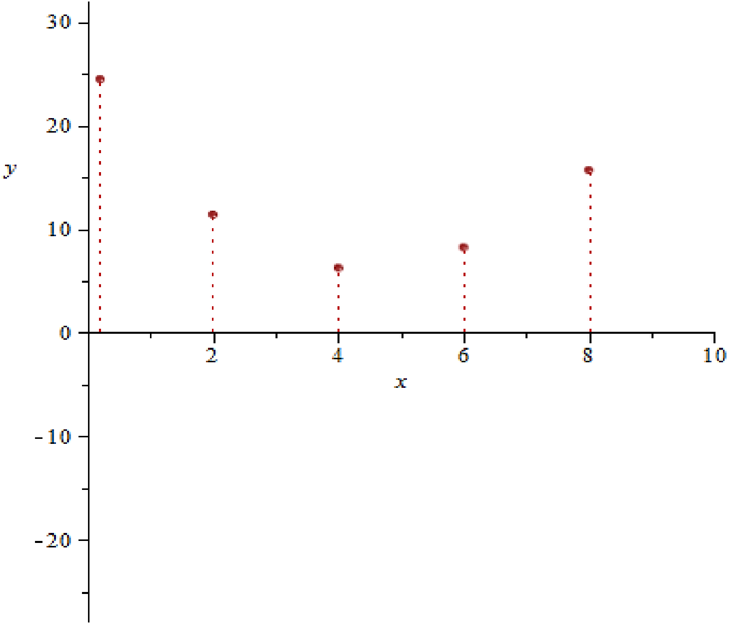
[1] Если такая точка существует.

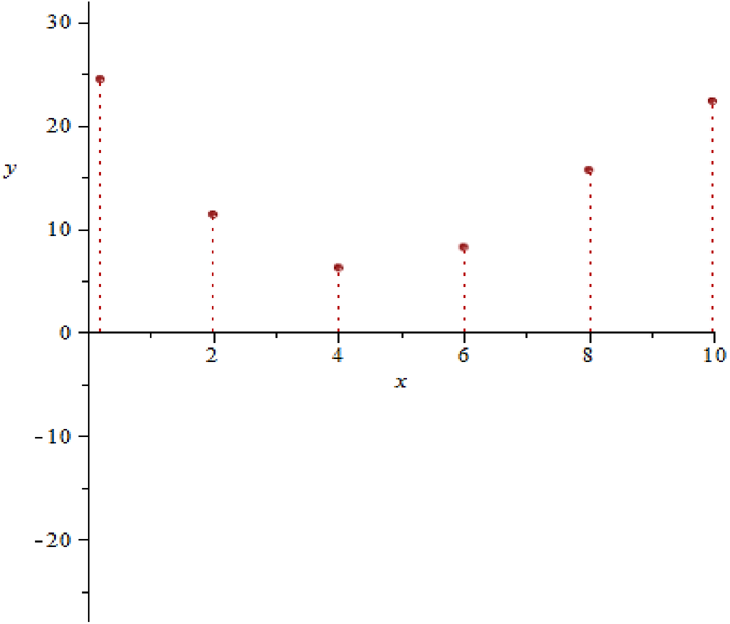

30
20
10
0
-10
-20
y
2
4
6
8
10
x

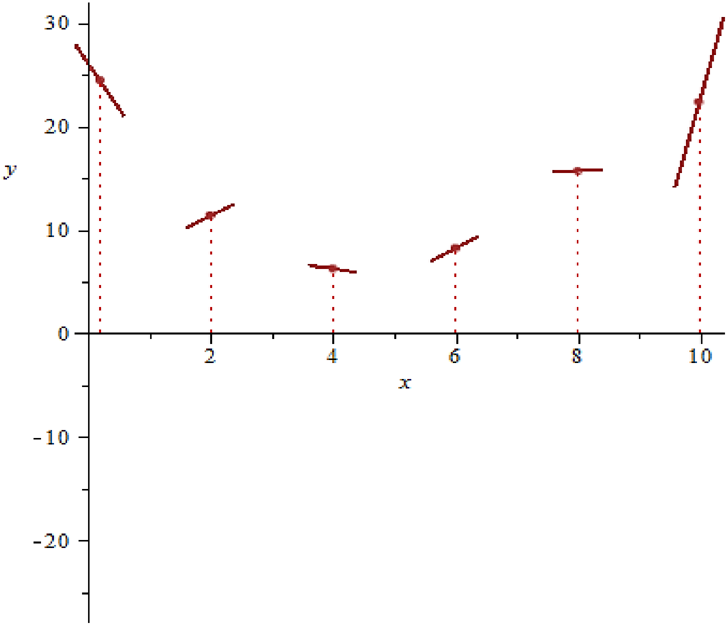

30
20
10
0
-10
-20
y
2
4
6
8
10
x

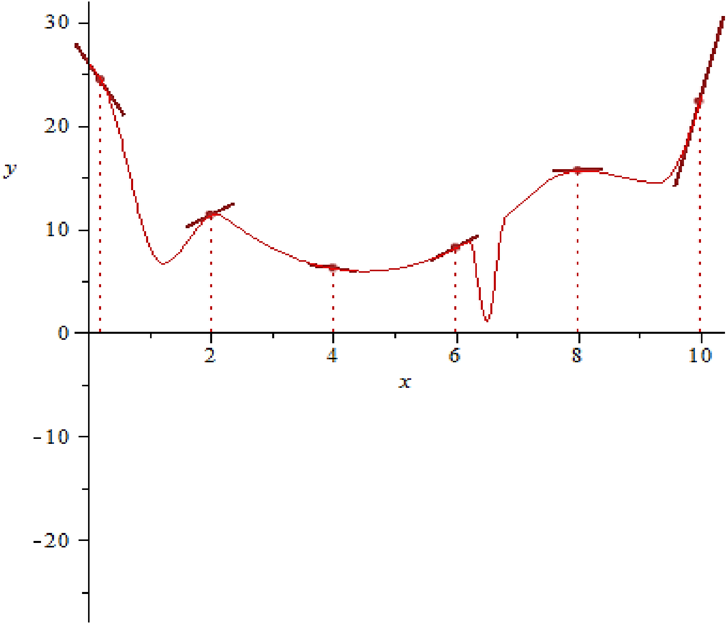

30
20
10
0
-10
-20
y
2
4
6
8
10
x

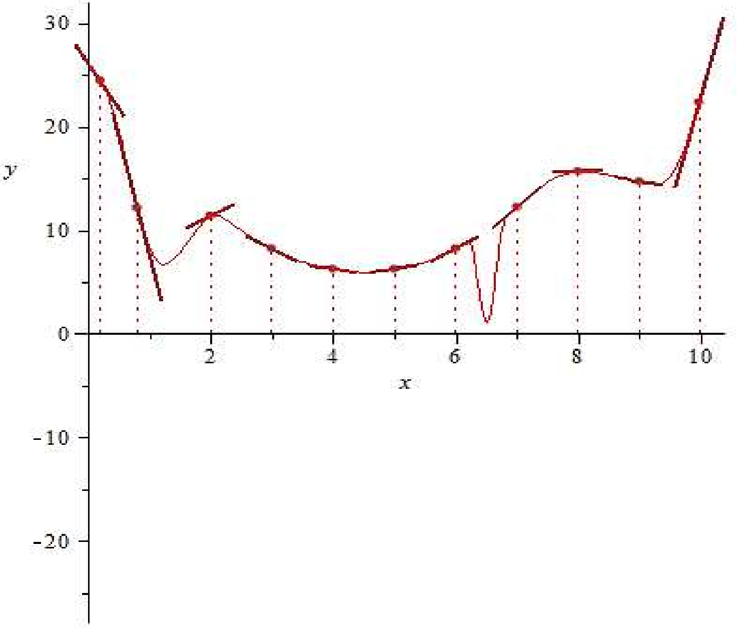

30
20
10
0
-10
-20
y
2
4
6
8
10
x


30
20
10
0
-10
-20
y
2
4
6
8
10
x

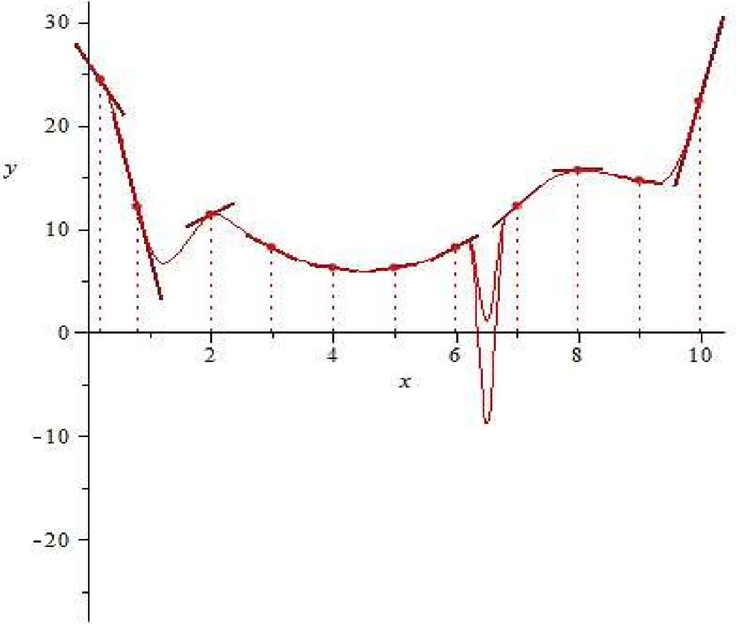

30
20
10
0
-10
-20
y
2
4
6
8
10
x

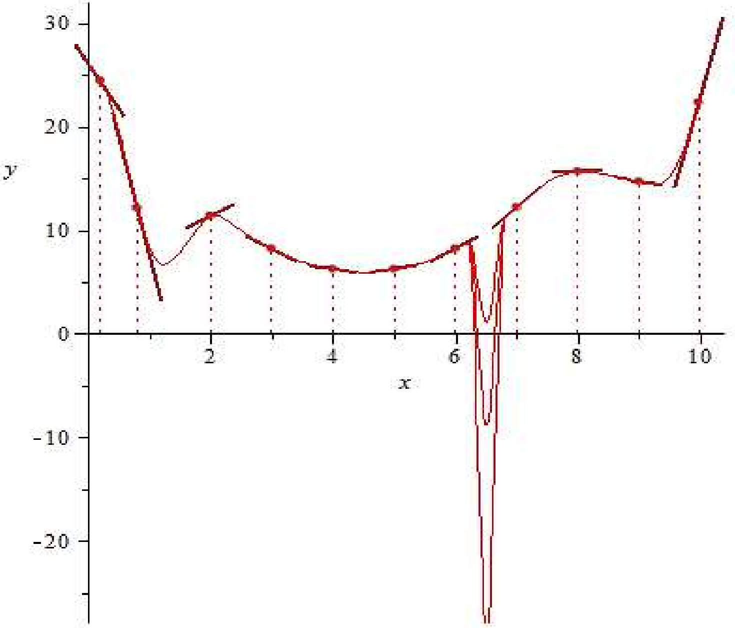

30
20
10
0
-10
-20
y
2
4
6
8
10
x


30
20
10
0
-10
-20
y
2
4
6
8
10
x



30
20
10
0
-10
-20
y
2
4
6
8
10
x


30
20
10
0
-10
-20
y
2
4
6
8
10
x


30
20
10
0
-10
-20
y
2
4
6
8
10
x

Оптимизация: дифференцируемость

Задача безусловной минимизации

$$f(x) \to \min,$$

$$x \in \mathbb{R}^n.$$

Изменится ли ситуация, если предположить $f \in C^k$, $k > 1$?

Ответ: Нет!!!

Теорема Уитни. *Пусть* $A \subset \mathbb{R}^n$ — замкнутое множество. Тогда существует функция $h \in C^\infty$ такая, что

$$A = \{x \in \mathbb{R}^n : h(x) = 0\}.$$

Оптимизация: дифференцируемость

Задача безусловной минимизации

$$f(x) \to \min,$$

$$x \in \mathbb{R}^n.$$

Изменится ли ситуация, если предположить $f \in C^k$, $k > 1$?

Ответ: Нет!!!

Теорема Уитни. *Пусть* $A \subset \mathbb{R}^n$ — замкнутое множество. Тогда существует функция $h \in C^\infty$ такая, что

$$A = \{x \in \mathbb{R}^n : h(x) = 0\}.$$

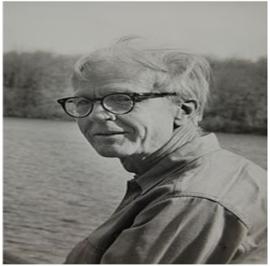
$$\Downarrow$$

$$h^2(x) \to \min .$$

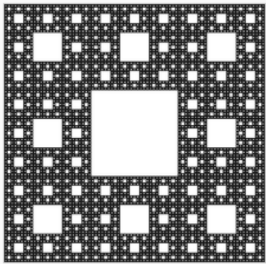
Оптимизация: дифференцируемость

Теорема Уитни. *Пусть* $A \subset \mathbb{R}^n$ — замкнутое множество. Тогда существует функция $h \in C^\infty$ такая, что

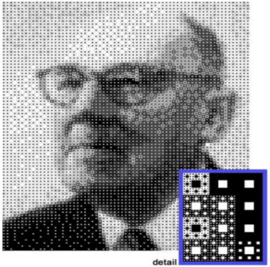
$$A = \{x \in \mathbb{R}^n : h(x) = 0\}.$$



H. Whitney



Ковёр Серпинского



В. Серпинский

Оптимизация: структурируемость

Главный вывод: необходима дополнительная информация (структурные свойства)

Например:

- *линейность*
- *квадратичность*
- *выпуклость*
- ...?

$$\Downarrow$$

Классы (так называемых) хорошо структурированных задач

Оптимизация: структурируемость

Задача безусловной минимизации

$$f(x) \to \min,$$

$$x \in \mathbb{R}^n,$$

$f \in C^2, \exists! x^* : \nabla f(x^*) = 0, \nabla^2 f(x^*) \succ 0$ **, является ли x^* точкой глобального минимума?**

Оптимизация: структурируемость

Задача безусловной минимизации

$$f(x) \to \min,$$

$$x \in \mathbb{R}^n,$$

$f \in C^2, \exists! x^* : \nabla f(x^*) = 0, \nabla^2 f(x^*) \succ 0$ **, является ли** x^* **точкой глобального минимума?**

- **Да, если** $n = 1$.
- **Нет, если** $n > 1$.

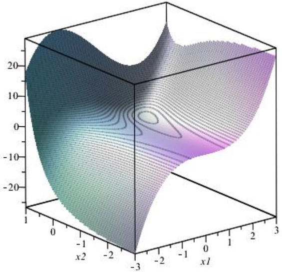
Оптимизация: структурируемость

Пример.

$$f(x_1, x_2) = e^{3x_2} - 3x_1 e^{x_2} + x_1^3,$$

$$x^* = (1, 0), \nabla f(x_1^*, x_2^*) = 0,$$

$$f(x_1, 0) = 1 - 3x_1 + x_1^3.$$



Одноэкстремальные невыпуклые задачи

Свойство выпуклой гладкой функции

$$f(x) - f(y) \geq f'(y)^T(x - y) \ \forall x, y \in \mathbb{R}^n.$$

Замена: $(x - y) \to \eta(x, y)$, $\eta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$

$$\Downarrow$$

$$f(x) - f(y) \geq f'(y)^T \eta(x, y) \ \forall x, y \in \mathbb{R}^n.$$

Тогда!

$f'(y) = 0 \ \Rightarrow \ y$ — точка глобального минимума[2]

f — инвексная функция (*in*variant con*vex*)

[2] Hanson M.A. *On sufficiency of the Kuhn-Tucker conditions.* Journal of Math. Analysis Appl. 80: 545-550, 1981

Одноэкстремальные невыпуклые задачи

- **Как распознать инвексность?**

$$f(x) = g(w(x)) = g(w_1(x), \ldots, w_m(x)),$$

g — выпуклая дифференцируемая функция.

Выпуклость g: $\underbrace{g(w(x))}_{f(x)} - \underbrace{g(w(y))}_{f(y)} \geq \nabla g(w(y))^T (w(x) - w(y))$.

Инвексность: $f(x) - f(y) \geq \nabla f(y)^T \eta(x, y)$, градиент

$$\nabla f(y) = \nabla w(y)^T \nabla g(w(y))$$

Если

$$\nabla w(y) \eta(x, y) = w(x) - w(y),$$

то f — инвексная функция. Данное условие не зависит от g !

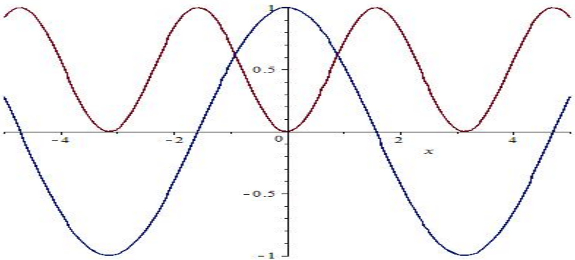
Одноэкстремальные невыпуклые задачи

- **Пример.**

$$f(x) = \sin^2(x), g(w) = w^2, w(x) = \sin(x).$$

Достаточное условие: $w'(x)\eta(x,y) = w(x) - w(y)$

$$\eta(x,y) = \frac{\sin(x) - \sin(y)}{\cos(y)}, \cos(y) \neq 0.$$



Одноэкстремальные невыпуклые задачи

Функция Розенброка

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (x_1 - 1)^2,$$

$$g(w) = w_1^2 + w_2^2, w_1(x) = 10(x_2 - x_1^2), w_2(x) = x_1 - 1.$$

достаточное условие $\nabla w(y)\eta(x, y) = w(x) - w(y)$ $(det(w(y) \neq 0)$:

$$\begin{pmatrix} -20y_1 & 10 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} 10(x_2 - x_1^2) - 10(y_2 - y_1^2) \\ x_1 - y_1 \end{pmatrix}$$

$$\Downarrow$$

$$\eta_1(x, y) = x_1 - y_1,$$

$$\eta_2(x, y) = x_2 - y_2 - (x_1 - y_1)^2.$$

Одноэкстремальные невыпуклые задачи

Теорема (Hanson, 1981). Пусть $f : \mathbb{R}^n \to \mathbb{R}$ — дифференцируемая функция. Тогда f является инвексной функцией тогда и только тогда, когда каждая стационарная точка f есть точка глобального минимума.

$$\Downarrow$$

Следствие. Если f не имеет стационарных точек, то f — инвексная функция.

Одноэкстремальные невыпуклые задачи

ПРИМЕР. $X = \{(x_1, x_2) : x_2 > 0\}$

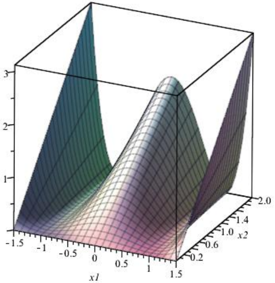
$f(x_1, x_2) = x_2(x_1^2 - 1)^2$

f — invex on X

$\{x : \nabla f(x) = 0\} =$

$= \{(1, x_2), x_2 > 0\} \bigcup$

$\{(-1, x_2), x_2 > 0\}$



Множество стационарных точек не является выпуклым

Одноэкстремальные невыпуклые задачи

Задачи инвексного программирования

$$f_0(x) \to \min,$$

$$f_i(x) \leq 0, \; i = 1, \ldots, m,$$

$$f_i(x) - f_i(y) \geq \nabla f_i(y)^T \eta(x, y), \; i = 0, 1, \ldots, m.$$

η — одна и та же для все f_i!

Теорема. *Пусть* x^* — точка, в которой выполняется какое-либо условие регулярности и необходимые условия первого порядка. Тогда x^* — точка глобального минимума задачи инвексного программирования.

Одноэкстремальные невыпуклые задачи

Пример.

$$f_0(x_1, x_2) = x_1 - \sin(x_2) \to \min,$$

$$f_1(x_1, x_2) = \sin(x_1) - 4\sin(x_2) \leq 0,$$

$$f_2(x_1, x_2) = 2\sin(x_1) + 7\sin(x_2) + x_1 - 6 \leq 0,$$

$$f_3(x_1, x_2) = 2x_1 + 2x_2 - 3 \leq 0,$$

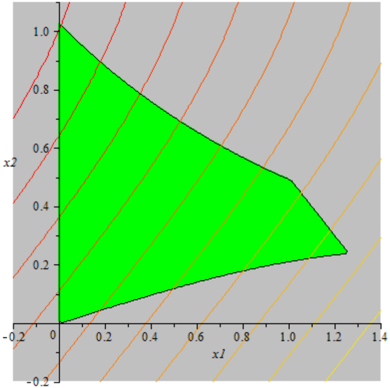
$$f_4(x_1, x_2) = 4x_1^2 + 4x_2^2 - 9 \leq 0,$$

$$f_5(x_1) = -\sin(x_1) \leq 0.$$

$$\eta(x, y) = \begin{pmatrix} \frac{\sin(x_1) - \sin(y_1)}{cos(y_1)} \\ \frac{\sin(x_2) - \sin(y_2)}{cos(y_2)} \end{pmatrix}$$

Одноэкстремальные невыпуклые задачи

Пример.



Одноэкстремальные невыпуклые задачи

Пессимистический результат.

Теорема.[3] Если все функции $f_i, i = 1, \ldots, m$ аффинны, то такая задача будет задачей инвексного программирования в том и только том случае, когда целевая функция f_0 есть выпуклая дифференцируемая функция.

[3] Martin D. *The essence of invexity.* JOTA, v. 47, №1, p.65-76, 1982

Одноэкстремальные невыпуклые задачи

Mishra S.K. , Giorgi G. *Invexity and Optimization.* – Springer-Verlag, Berlin, 2008. – 266 p.

- Задачи глобальной оптимизации как правило **NP**-трудны.

$$x^T Qx + c^T x \to \min,$$

$$x \in X = \{x \in \mathbb{R}^n : Ax \leq b\},$$

Q имеет хотя бы одно отрицательное собственное число.
Пример 1. (“простая” **NP**-трудная задача)[4]

$$x_1 - x_2^2 \to \min,$$
$$x \in X = \{x \in \mathbb{R}^n : Ax \leq b\}.$$

[4] P.M. Pardalos, S.V. Vavasis *Quadratic programming with one negative eigenvalue is* **NP**-*hard* // JOGO. – 1991. – Vol.1, №1 .– P.15-22

- Задачи глобальной оптимизации как правило **NP**-трудны.

$$x^T Qx + c^T x \to \min,$$

$$x \in X = \{x \in \mathbb{R}^n : Ax \leq b\},$$

Q имеет хотя бы одно отрицательное собственное число.
Пример 2. ("простая" **NP**-трудная задача)[5]

$$x_1 x_2 \to \min,$$
$$x \in X = \{x \in \mathbb{R}^n : Ax \leq b\}.$$

[5] T. Matsui **NP**-*hardness of linear multiplicative programming and related problems* // JOGO. – 1996. – Vol.9, №2. – P.113-119

Пример 3. (“простая” **NP**-трудная задача)[6]

$$x_1 - \frac{1}{x_2} \to \min,$$
$$x \in X = \{x \in \mathbb{R}^n : Ax \le b,\ x_2 > 0\}.$$

Пример 4. (“простая” **NP**-трудная задача)[6]

$$\frac{1}{x_1} + \frac{1}{x_2} \to \max,$$
$$x \in X = \{x \in \mathbb{R}^n : Ax \le b,\ x_1 > 0,\ x_2 > 0\}.$$

[6] T. Matsui **NP**-*hardness of linear multiplicative programming and related problems* // JOGO. – 1996. – Vol.9, №2. – P.113-119

Пример 5. ("стандартная" **NP**-трудная задача)[7]

$$x^T Qx + c^T x \to \min,$$

$$x \in S = \{x \in \mathbb{R}^n_+ : \sum_{j=1}^{n} x_j = 1\},$$

Q имеет хотя бы одно отрицательное собственное число.

Пример 6. ("стандартная" **NP**-трудная задача)[7]

$$x^T Qx + c^T x \to \min,$$

$$x \in \Pi = \{x \in \mathbb{R}^n : 0 \le x_j \le 1,\ j = 1, \dots, n\},$$

$Q \ne I$ имеет хотя бы одно отрицательное собственное число.

[7] S. Vavasis *Nonlinear optimization. Complexity issues* Oxford University Press, 1991.

- Минимизация квадратичной функции на шаре

$$
\begin{aligned}
x^T Qx + c^T x &\to \min, \\
\|x\| &\leq r.
\end{aligned}
\tag{1}
$$

Пусть $U : U^T QU = diag(\lambda_1, \ldots, \lambda_n),\ p = U^T c$. Тогда (1) эквивалентна задаче *выпуклого* программирования[8]:

$$
\begin{aligned}
\sum_{j=1}^{n} \lambda_j y_j - 2|p_j|\sqrt{y_j} &\to \min, \\
\sum_{j=1}^{n} y_j \leq r,\ y_j \geq 0,\ j &= 1, \ldots, n.
\end{aligned}
\tag{2}
$$

Соотношения: $x_j^* = sign(p_j)\sqrt{y_j^*}, j = 1, \ldots, n.$

[8] A. Ben-Tal, M. Teboulle *Hidden convexity in some nonconvex quadratically constrained quadratic programming problems* // Mathematical Programming. – 1996. – Vol.72. – P.51-63

- Минимизация квадратичной функции при одном квадратичном ограничении

$$
\begin{aligned}
q_0(x) &= x^T A x + c^T x \to \min, \\
q_1(x) &= x^T B x + p^T x + r \le 0.
\end{aligned} \tag{3}
$$

Теорема I[9]. *Пусть* $B \neq 0$ *и* $-\infty \le \inf\limits_{x \in \mathbb{R}^n} q_1(x) < 0$. *Точка* $\overline{x} : q_1(\overline{x}) \le 0$ *будет точкой глобального минимума тогда и только тогда, когда существует* $\overline{\lambda} \ge 0$*:*

- $2(A + \overline{\lambda} B)\overline{x} + c + \overline{\lambda} p = 0$;
- $\overline{\lambda} q_1(\overline{x}) = 0$;
- $A + \overline{\lambda} B$ *положительно полуопределена.*

[9] J.J. More *Generalization of the trust region problem* // Optim. Meth. and Software. – 1993. – Vol.2. – P.189-209

- Минимизация квадратичной функции при двух квадратичных ограничениях

$$
\begin{aligned}
q_0(x) &= x^T A x + c^T x \to \min, \\
q_1(x) &= x^T B x + p^T x + r \le 0, \\
q_2(x) &= x^T Q x + d^T x + s \le 0.
\end{aligned}
\tag{4}
$$

Существует разрыв между необходимыми и достаточными условиями глобальной оптимальности.

Теорема II (достаточные условия)[10]. *Пусть $\overline{x}$: $q_1(\overline{x}) \le 0, q_2(\overline{x}) \le 0$ и существуют $\overline{\lambda}_1 \ge 0$ и $\overline{\lambda}_2 \ge 0$:*

- $\nabla q_0(\overline{}) + \overline{\lambda}_1 \nabla q_1(\overline{x}) + \overline{\lambda}_2 \nabla q_2(\overline{x}) = 0$;
- $\overline{\lambda}_1 q_1(\overline{x}) = \overline{\lambda}_2 q_2(\overline{x}) = 0$;
- $A + \overline{\lambda}_1 B + \overline{\lambda}_2 Q$ *положительно полуопределена.*

Тогда $\overline{x}$ – глобальный минимум задачи (4).

Теорема III (необходимые условия)[10]. *Пусть $\overline{x}$ – глобальный минимум в задаче (4), $q_1(\overline{x}) = 0, q_2(\overline{x}) = 0$ и градиенты $\nabla q_1(\overline{x})$, $\nabla q_2(\overline{x})$ линейно независимы. Тогда существуют единственные $\overline{\lambda}_1 \ge 0$ и $\overline{\lambda}_2 \ge 0$:*

- $\nabla q_0(\overline{x}) + \overline{\lambda}_1 \nabla q_1(\overline{x}) + \overline{\lambda}_2 \nabla q_2(\overline{x}) = 0$;
- $A + \overline{\lambda}_1 B + \overline{\lambda}_2 Q$ *имеет не более одного отрицательного собственного числа.*

[10] J.M. Peng, Y. Yuan *Optimality conditions for the minimization of a quadratic with two quadratic constraints* // SIAM J. Optimization. – 1997. – Vol.7,№3. – P.579-594

- В общем случае задачи глобальной оптимизации не поддаются эффективному решению.
- Тем не менее, некоторые (специальные) задачи глобальной (многоэкстремальной) оптимизации обладают свойством "*скрытой выпуклости*".
- Термин "*скрытая выпуклость*" введен В.М. Тихомировым в 1989 г.
- В настоящее время многоэкстремальные задачи, обладающие свойством "*скрытой выпуклости*", поддаются эффективному решению. Для этих задач
  - найдены критерии глобальной оптимальности;
  - характерно отсутствие разрыва классической (лагранжевой) двойственности;

- И в задаче с одним квадратичным ограничением и в задаче с двумя квадратичными ограничениями *разрыв двойственности равен нулю*!
- Возможная схема в $\mathbb{R}^n$

$$\begin{gathered}\text{Условия оптимальности}\\ \Uparrow \\ \text{Теоремы об альтернативе}\\ \Uparrow \\ \text{Теоремы об отделимости}\end{gathered}$$

Классическая двойственность: f^* - оптимальное значение прямой задачи

$$f_0(x) \to \min,$$

$$f_i(x) \leqslant 0, \; i = 1, \dots, m,$$

$$x \in X \subset \mathbb{R}^n.$$

Θ^* - оптимальное значение двойственной задачи

$$\Theta^* = \sup_{\lambda \geqslant 0} \min_{x \in X} \{ f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) \}.$$

Слабая теорема двойственности: $\Theta^* \leqslant f^*$.
Разрыв двойственности (duality gap): $f^* - \Theta^* \geqslant 0$.

Теорема (Фана)[11]. *Пусть X – выпуклое множество в $\mathbb{R}^n$, $f_1, \dots, f_k$ – выпуклые функции на X. Предположим, что система*

$$f_i(x) < 0, i = 1, \dots, k$$

не имеет решения на X. Тогда существуют $\lambda_1 \geq 0, \dots, \lambda_k \geq 0$, не равные нулю одновременно:

$$\sum_{i=1}^{k} \lambda_j f_i(x) \geq 0 \; \forall x \in X.$$

Доказательство основано на разделении плоскостью *выпуклых* множеств в "пространстве образов"

$$U_1 = \{u \in \mathbb{R}^k : f_i(x) \leq u_i \text{ при некотором } x \in X\}$$

$$U_2 = \{u \in \mathbb{R}^k : u_i < 0, i = 1, \dots, k\}.$$

[11] А.Г. Сухарев, А.В. Тимохов, В.В. Фёдоров *Курс методов оптимизации.*– М.: Наука, 1996. – 328 с.

Задача

$$f_0(x) \to \min,$$

$$f_i(x) \le 0, i = 1, \dots, m,$$

$$x \in X$$

называется задачей *“с выпуклым образом”*[12] если $\forall x^1, x^2 \in X$, $\forall \alpha \in [0,1]$ $\exists x_\alpha \in X$:

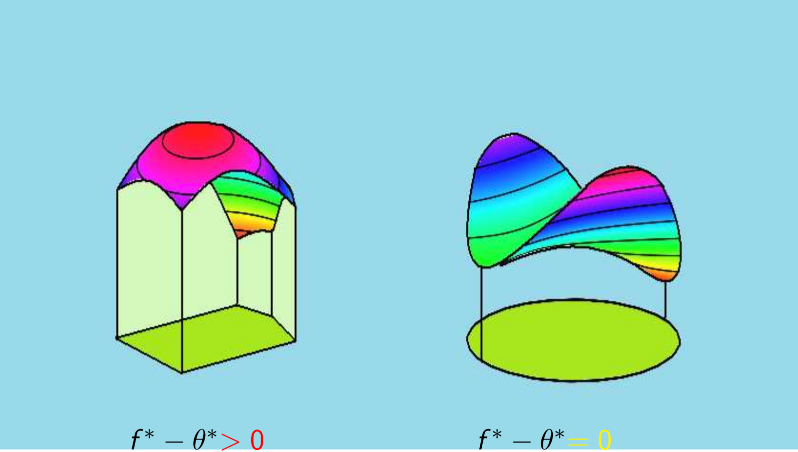
$$f_i(x_\alpha) \le (1-\alpha) f_i(x^1) + \alpha f_i(x^2), \;\; i = 0, \dots, m.$$

Зарубежный аналог: convexlike – выпуклоподобная

Выпуклоподобные задачи и только они допускают разделение плоскостью “образа задачи” с отрицательным ортантом[13].

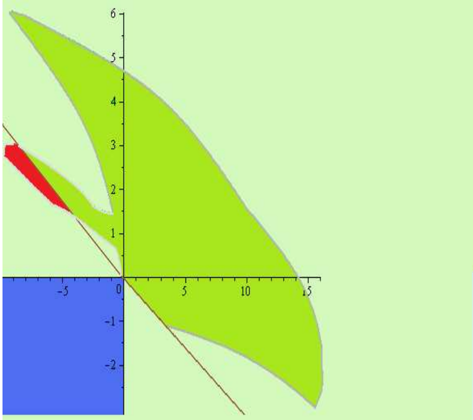
[12] В.М. Тихомиров *Теория экстремума и экстремальные задачи классического анализа* // Итоги науки и техники. Тематические обзоры. – М.: ВИНИТИ, 1998. – Т.60. – С.188-258

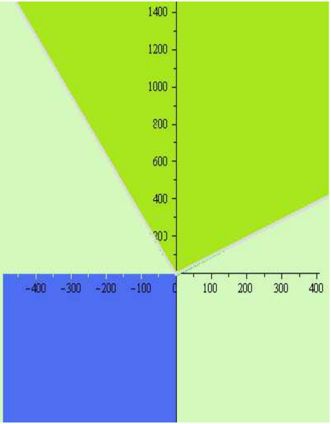
[13] F. Gianessi *Constrained Optimization and Image Space Analysis. Volume 1: Separation of Sets and Optimality Conditions.* - Springer, 2005. - 396 p.



f^* – оптимальное значение прямой задачи

θ^* – оптимальное значение двойственной задачи


6
5
4
3
2
1
0
-1
-2
-5
5
10
15


1400
1200
1000
800
600
400
-400
-300
-200
-100
0
100
200
300
400

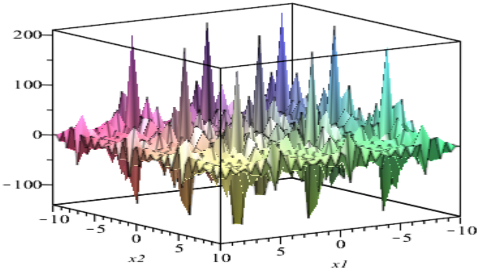
- Работы Якубовича В.А. и Фрадкова А.Л.
- Polyak B.T. *Convexity of quadratic transformations and its use in control and optimization.* - JOTA, 1998, vol. 99, No. 3, pp. 553-583
- Поляк Б.Т. *Локальное программирование.* - ЖВМиМФ, 2001, т.41, №9, с. 1324-1331

B&B

Why B&B?

Example 1. Shubert function

$$f(x_1, x_2) = \left(\sum_{i=1}^{5} i\cos\left[(i+1)x_1 + i\right]\right) \cdot \left(\sum_{i=1}^{5} i\cos\left[(i+1)x_2 + i\right]\right),$$

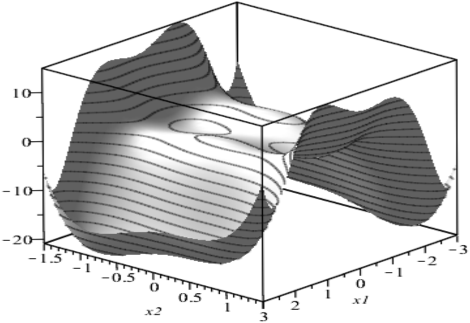


760 local minima, 18 global minima

Why B&B?

Example 2.

$$f(x_1, x_2) = 4x_1^2 - 2.1x_1^4 + \frac{x_1^6}{6} + x_1x_2 - 4x_2^2 + 4x_2^4,$$
$$-3 \leqslant x_1 \leqslant 3, \; -1.5 \leqslant x_2 \leqslant 1.5.$$

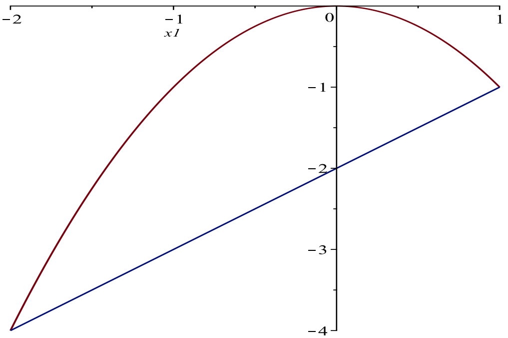


6 local minima, 2 global minima

Convex envelopes

- Univariate concave function

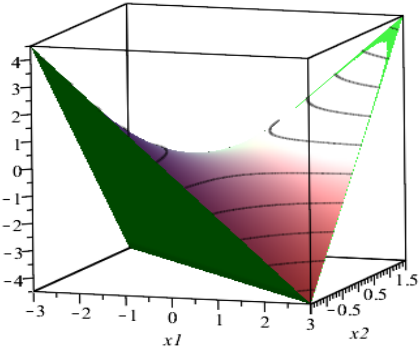
$$-ax_1^2 \geqslant l_1(x) = -a(\overline{x}_1 + \underline{x_1})x + a\overline{x}_1\underline{x}_1, \quad \underline{x}_1 \leqslant x_1 \leqslant \overline{x}_1.$$



Convex envelopes

- Bilinear function

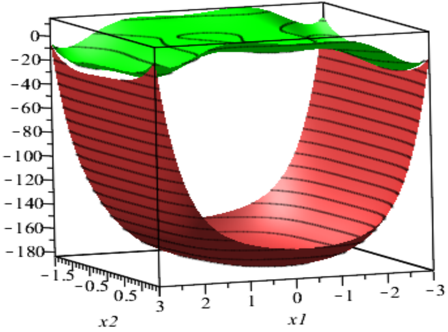
$$x_1x_2 \geqslant b(x_1, x_2) = \max(\underline{x}_2 x_1 + \underline{x}_1 x_2 - \underline{x}_1\underline{x}_2,\ \overline{x}_2 x_1 + \overline{x}_1 x_2 - \overline{x}_1\overline{x}_2)$$



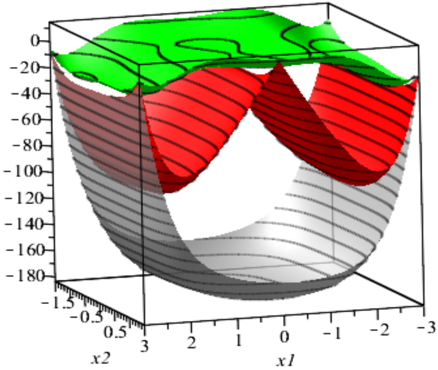
Bounding

$$f(x_1, x_2) = 4x_1^2 - 2.1x_1^4 + \frac{x_1^6}{6} + x_1 x_2 - 4x_2^2 + 4x_2^4 \geqslant$$

$$\geqslant 4x_1^2 + \frac{x_1^6}{6} + 4x_2^4 + l_1(x_1) + l_2(x_2) + b(x_1, x_2).$$



Bounding

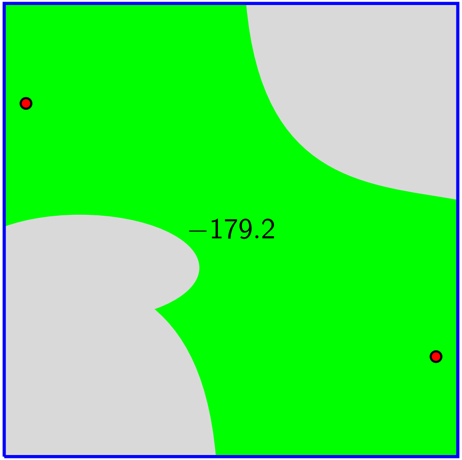


B&B

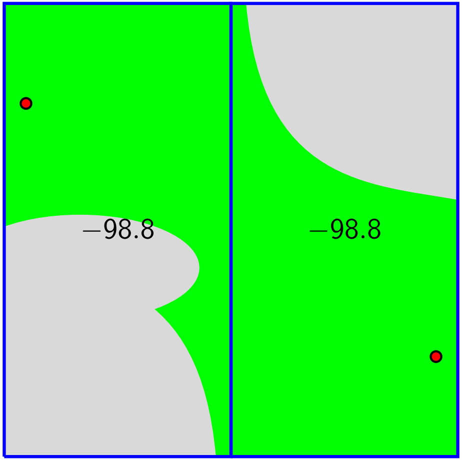
$$
\begin{gathered}
4x_1^2 - 2.1x_1^4 + \frac{x_1^6}{6} + x_1x_2 - 4x_2^2 + 4x_2^4 \to \min, \\
x_1x_2 \leqslant 2, \\
0.5(x_1+2)^2 + 1.8(x_2+0.5)^2 \geqslant 1, \\
-3 \leqslant x_1 \leqslant 3, \ -1.5 \leqslant x_2 \leqslant 1.5.
\end{gathered}
$$

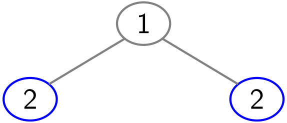
$f^* = -20.98$

1

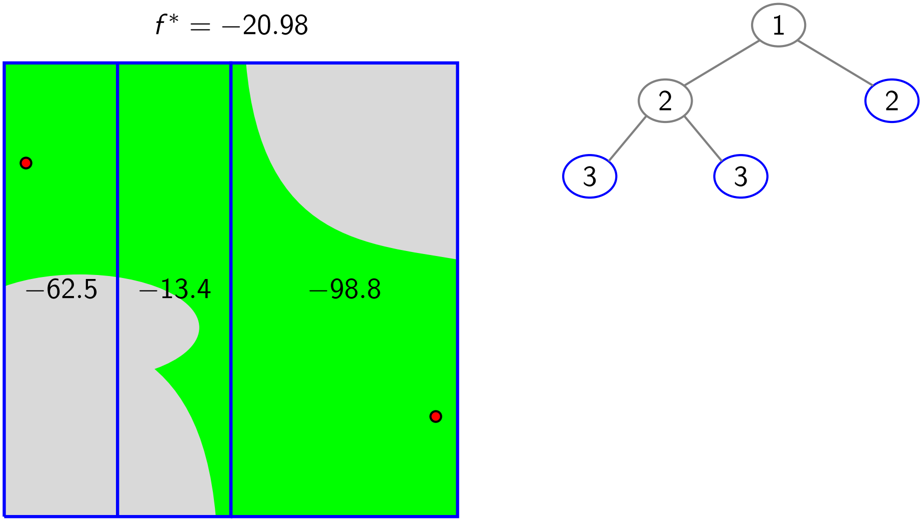


$$f^* = -20.98$$

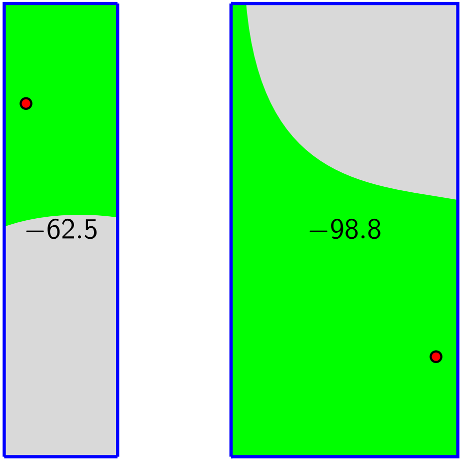


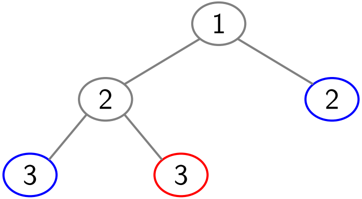


$$f^* = -20.98$$

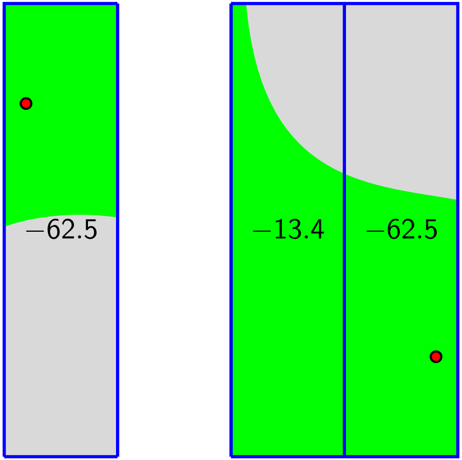


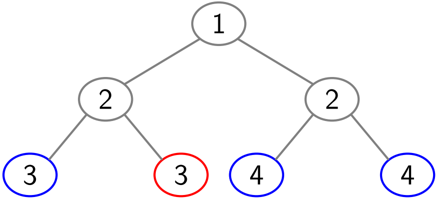
$f^* = -20.98$



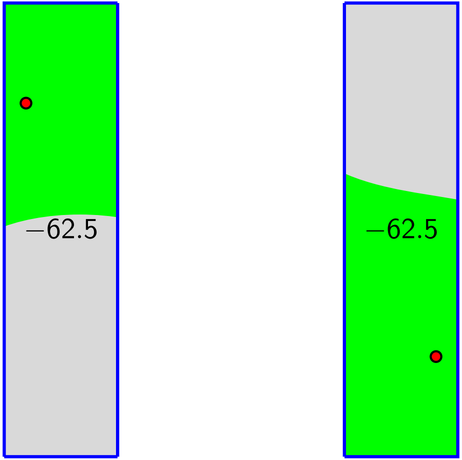


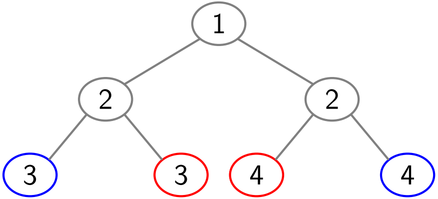
$$f^* = -20.98$$



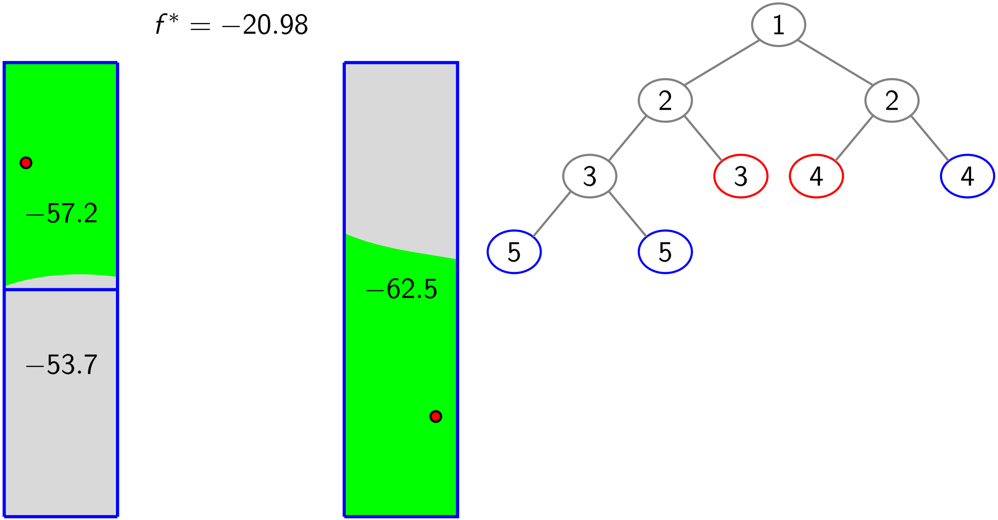


$f^* = -20.98$

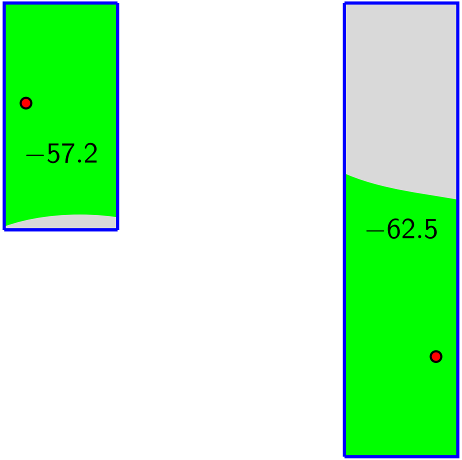


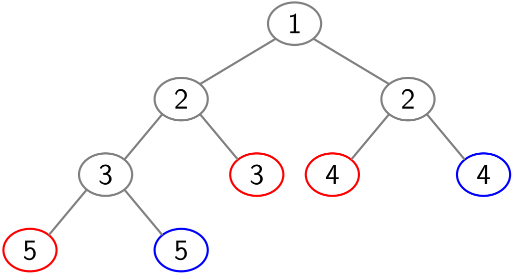


$f^* = -20.98$



$f^* = -20.98$





B&B

- Main points
  - Lower bounding
  - Partition
  - Upper bounding
  - Selection rule
  - Fathoming
- Global continuous optimization put strong assumption on B&B methods in order to guarantee convergence.

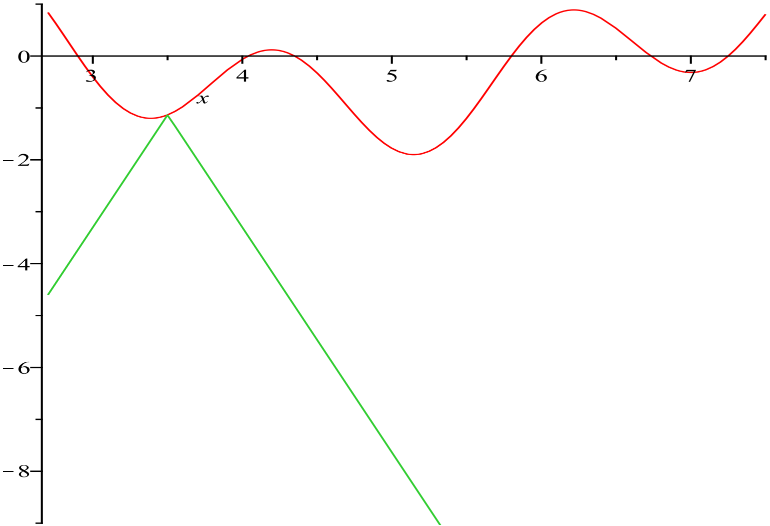
LIPSCHITZ OPTIMIZATION

$$|f(x) - f(y)| \leq L\|x - y\|$$

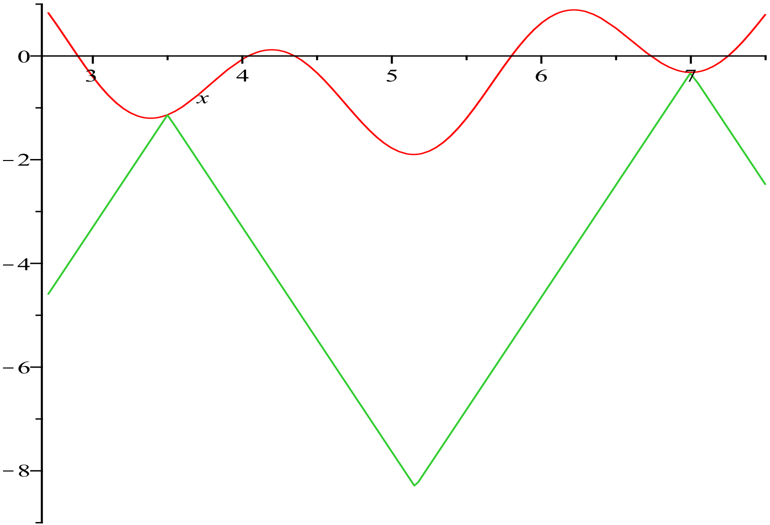
$$\Downarrow$$

$$\underbrace{f(y) - L\|x - y\|}_{\text{support minorant}} \leq f(x) \leq \underbrace{f(y) + L\|x - y\|}_{\text{support majorant}}$$

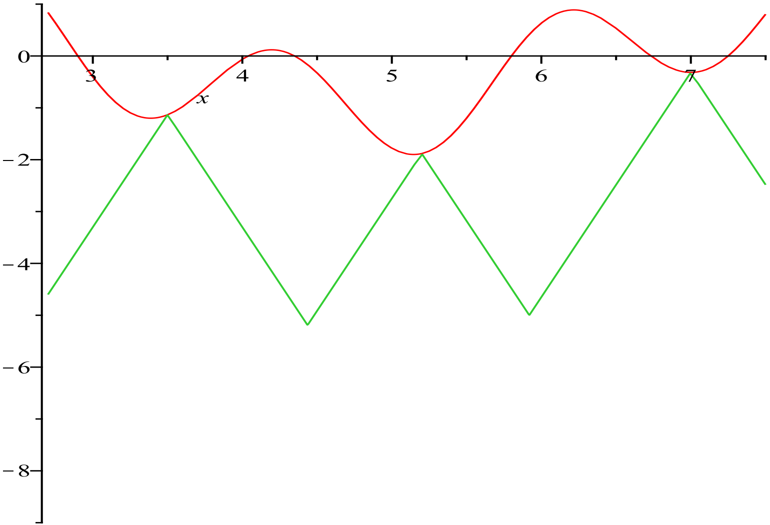
Lipschitz optimization



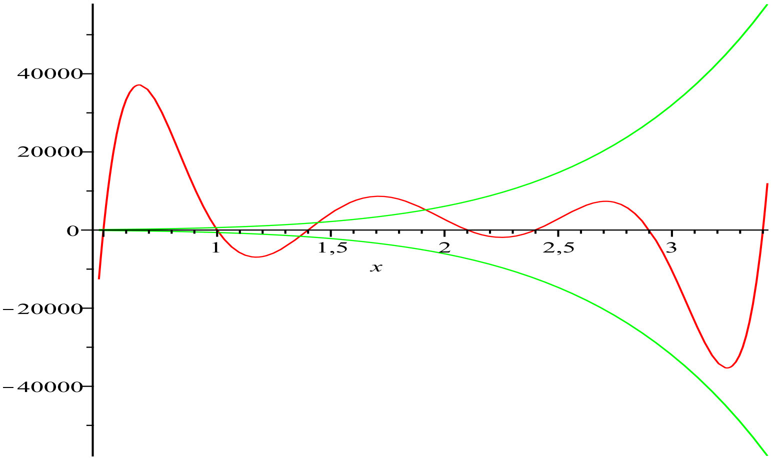
Lipschitz optimization



Lipschitz optimization

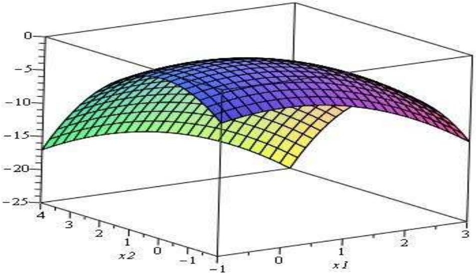


D.C. OPTIMIZATION ($f(x) = p(x) - q(x)$)



1. Reduction to Concave Programming Problem:

$$p(x) - q(x) \to \min,$$
$$x \in X.$$
$$\Downarrow$$
$$z - q(x) \to \min,$$
$$p(x) \le z,\ x \in X.$$

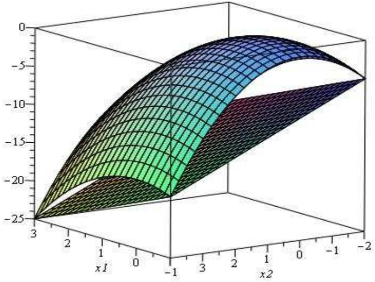


2. Convex minorants:

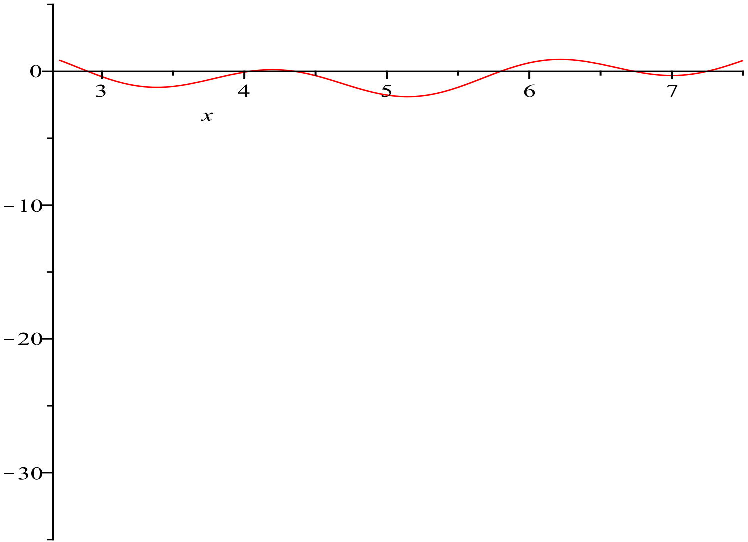
$$p(x) - q(x) \geq \underbrace{p(x) - l(x)}_{\text{convex function}},$$

$$-q(x) \geq l(x), \; l(x) = a^T x + b, \; x \in S,$$

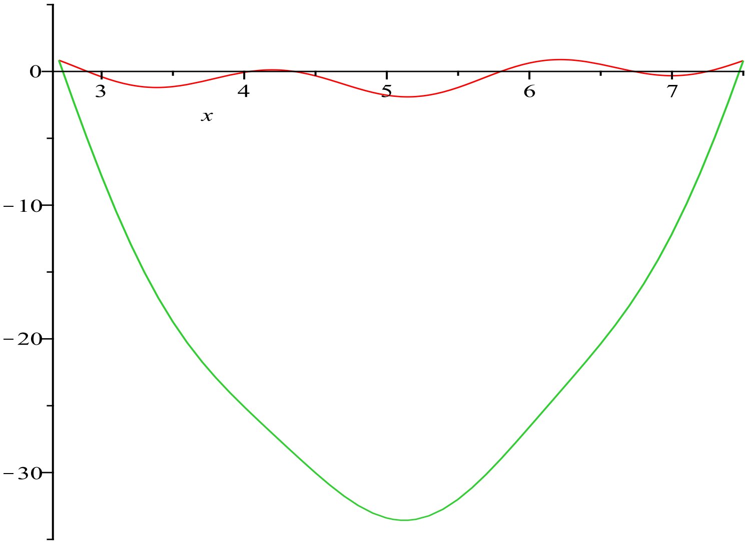
S is a simplex or parallelepiped.



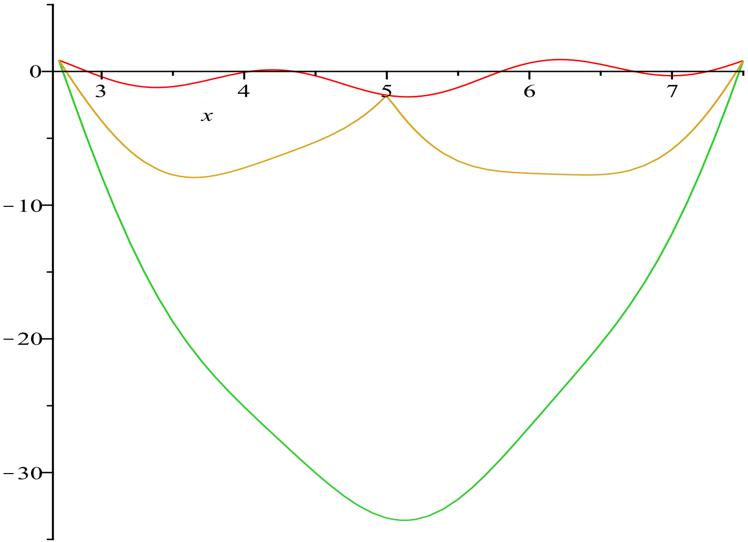
B&B with convex majorants



B&B with convex majorants



B&B with convex majorants



- Horst R., Tuy H. *Global Optimization. (Deterministic Approaches)*, 3rd, revised and enlarged edition. – Berlin: Springer-Verlag, 1996.
- Horst R., Pardalos P.M., Thoai N.V. *Introduction to Global Optimization.* – Dordrecht: Kluwer Academic Publishers, 1995.

- Not very often used underestimator[14]:

Assumption A1

$$\|\nabla^2 f(x) - \nabla^2 f(y)\| \leq N\|x - y\|$$

then

$$\left| f(x) - f(y) - \nabla f(y)^T(x - y) - \frac{1}{2}(x - y)^T \nabla^2 f(y)(x - y) \right| \leq \frac{N}{6}\|x - y\|^3$$

and 2-minorant

$$\psi_2(x, y) = f(y) + \nabla f(y)^T(x - y) + \frac{1}{2}(x - y)^T \nabla^2 f(y)(x - y) - \frac{N}{6}\|x - y\|^3$$

[14] Khamisov O.V. *Optimization with quadratic support functions in nonconvex smooth optimization.* — Numerical Computations: Theory and Algorithms (NUMTA–2016), AIP Conf. Proc. 1776, 050010-1–050010-4, **2016**; doi: 10.1063/1.4965331

- Modification: from 2-minorant to q-minorant

Determine $d : \|x-y\| \leq d \ \forall x, y \in X$, then

$$\psi_2(x,y) = f(y) + \nabla f(y)^T(x-y) + \frac{1}{2}(x-y)^T\nabla^2 f(y)(x-y) - \frac{N}{6}\|x-y\|^3 \geq$$

$$\geq f(y) + \nabla f(y)^T(x-y) + \frac{1}{2}(x-y)^T\nabla^2 f(y)(x-y) - \frac{N\cdot d}{6}\|x-y\|^2 =$$

$$\geq f(y) + \nabla f(y)^T(x-y) + \frac{1}{2}(x-y)^T[\nabla^2 f(y) - \frac{N\cdot d}{3}\cdot I](x-y) = \psi_q(x,y),$$

I — unit matrix.

- Main property of q-minorant:

$$\psi_q(x,y) = f(y) + \nabla f(y)^T(x-y) + \frac{1}{2}(x-y)^T[\nabla^2 f(y) - \frac{N \cdot d}{3} \cdot I](x-y)$$

Let $y \in X$ be a local minimum:

1. $f(y)^T(x-y) \geq 0 \; \forall x \in X$;
2. $\nabla^2 f(y) \succ 0$.

If d is small enough then

$$\nabla^2 f(y) - \frac{N \cdot d}{3} \cdot I \succcurlyeq 0$$

and $\psi_q(x,y)$ is convex in x.

- Main property of q-minorant:

— As soon as $d \searrow 0$ $\psi_q(x, y)$ transforms from concave to convex in x function

— If

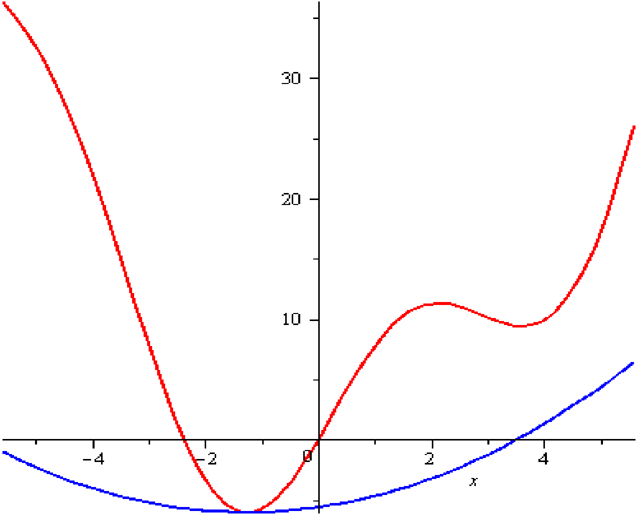
$$\nabla^2 f(y) - \frac{N \cdot d}{3} \cdot I \succcurlyeq 0$$

then y is global minimum over

$$S_d = \{x \in \mathbb{R}^n : \|x - y\| \leq d\}.$$

— Sufficient d can be calculated

$$d = \frac{3 \cdot \lambda_{min}(\nabla^2 f(y))}{N}.$$


30
20
10
−4
−2
0
2
4
x

Simplifying the objective (a trick)[15]

[15] Khamisov O.V. *Objective Function Decomposition in Global Optimization.* in: R. Battiti, D.E. Kvasov, Y.D. Sergeev (Eds.) LION 2017, LNCS 10556, pp. 338-344, **2017**, https://doi.org/10/1007/978-3-319-69404-7_28

Statement of the problem:

$$\min g(x), x \in X, \tag{5}$$

where g is a continuous composite function

$$g(x) = F(f_1(x), \dots, f_p(x)),$$

$F : \mathbb{R}^p \to \mathbb{R},\ f_i : X \to \mathbb{R}$ are continuous functions, $X \subset \mathbb{R}^n$ is a compact set.

Obvious reduction:

$$\min F(y),$$

$$y_i = f_i(x),\ i = 1, \dots, p,\ x \in X.$$

For the reduced problem

$$\min F(y), \tag{6}$$

$$y_i = f_i(x),\ i = 1, \ldots, p, \tag{7}$$

$$x \in X. \tag{8}$$

we assume

- F is "less complicated" (or $p < n$);
- Equations (7) are computationally tractable in x for fixed y .

Reduced problem:

$$\min F(y),$$

$$y_i = f_i(x),\ i = 1, \dots, p,\ x \in X.$$

For $f(x) = (f_1(x), \dots, f_p(x))$ define

$$Y = \{y \in \mathbb{R}^p : y = f(x) \text{ for some } x \in X\}, \tag{9}$$

$$\underline{y}_i \leqslant \underline{f}_i = \min_{x \in X} f_i(x),\ \max_{x \in X} f_i(x) = \overline{f}_i \leqslant \overline{y}_i,\ i = 1, \dots, p. \tag{10}$$

Problem in y

$$\min F(y), \tag{11}$$

$$\underline{y}_i \leqslant y_i \leqslant \overline{y}_i,\ i = 1, \dots, p, \tag{12}$$

$$y \in Y. \tag{13}$$

Inclusion (13) - Induced Constraint.

Reduced problem:

$$\min F(y),$$

$$y_i = f_i(x),\ i = 1, \ldots, p,\ x \in X.$$

Agreed decomposition

$$g(x) = F(f_1(x^1), \ldots, f_p(x^p)), \tag{14}$$

$$x^i \in X^i \subset R^{n_i},\ i = 1, \ldots, p, \tag{15}$$

$$X^1 \times \ldots \times X^p = X. \tag{16}$$

$$\Downarrow$$

$$\min F(y),$$

$$\underline{f}_i \leq y_i \leq \overline{f}_i,\ i = 1, \ldots, p,$$

$$y_i = f_i(x^i),\ x^i \in X^i,\ i = 1, \ldots, p.$$

Reduced agreed problem:

$$\min F(y), \tag{17}$$

$$\underline{f}_i \le y_i \le \overline{f}_i,\ i = 1, \ldots, p, \tag{18}$$

$$y_i = f_i(x^i),\ x^i \in X^i,\ i = 1, \ldots, p. \tag{19}$$

Three stage approach

I. Bounding stage. Determine $\underline{f}_i$ and $\overline{f}_i$ through (10).

II. Optimality stage. Solve the problem (17)-(18) and find optimal point y^*.

III. Feasibility stage. Solve the feasibility problem (19) for $y = y^*$ and obtain (all) optimal point(s) x^* for the initial problem.

Example 1.

$$g(x) = -(x_1 - 1)(x_1 + 2)(x_2 + 1)(x_2 - 2)x_3^2,$$
$$X = [-2, 2] \times [-2, 2] \times [-2, 2].$$

Let

$$F(y_1, y_2, y_3) = -y_1 y_2 y_3,$$
$$f_1(x_1) = (x_1 - 1)(x_1 + 2),$$
$$f_2(x_2) = (x_2 + 1)(x_2 - 2),$$
$$f_3(x_3) = x_3^2,$$
$$X^i = [-2, 2], \; i = 1, 2, 3.$$

I. Bounding stage. Each function f_i is convex, so

$$\underline{f}_1 = \underline{f}_2 = -2.25,\ \underline{f}_3 = 0,\ \overline{f}_1 = \overline{f}_2 = \overline{f}_3 = 4.$$

II. Optimality stage.

$$\min\{F(y) = -y_1 y_2 y_3\},$$
$$-2.25 \leqslant y_1 \leqslant 4,\ -2.25 \leqslant y_2 \leqslant 4,\ 0 \leqslant y_3 \leqslant 4,$$

$y^* = (4, 4, 4)$, $F(y^*) = -64$.

III. Feasibility stage.

$$(x_1 - 1)(x_1 + 2) = 4,\ -2 \leqslant x_1 \leqslant 2,\ \Rightarrow x_1^* = 2,$$
$$(x_2 + 1)(x_2 - 2) = 4,\ -2 \leqslant x_2 \leqslant 2,\ \Rightarrow x_2^* = -2,$$
$$x_3^2 = 4,\ -2 \leqslant x_3 \leqslant 2,\ \Rightarrow x_3^* = 2 \vee -2.$$

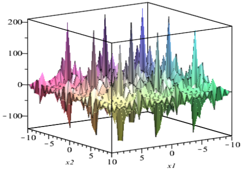
Two global minimum points: $(2, -2, 2)$, $(2, -2, -2)$.

Example 2. (Schubert function).

$$g(x_1, x_2) = f_1(x_1) \cdot f_2(x_2) =$$

$$= \left(\sum_{i=1}^{5} i \cos\left[(i+1)x_1 + i\right] \right) \cdot \left(\sum_{i=1}^{5} i \cos\left[(i+1)x_2 + i\right] \right),$$

$$X = X_1 \times X_2, \; X_1 = X_2 = [-10, 10].$$



It is obvious that
$F(y_1, y_2) = y_1 y_2,$
$y_1 = f_1(x_1), \; y_2 = f_2(x_2).$

Additional assumption: we can efficiently solve (explicit) univariate global optimization problems.

I. Bounding stage.

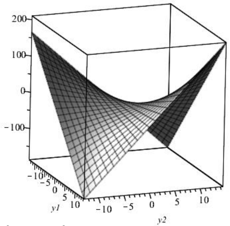
$$\underline{f}_1 = \underline{f}_2 = -12.87088549, \ \overline{f}_1 = \overline{f}_2 = 14.50800793.$$

II. Optimality stage.

$\min\{F(y_1, y_2) = y_1 y_2\}$,

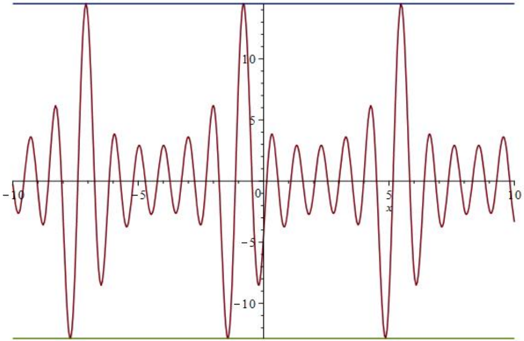
$-12.87088549 \leqslant y_i \leqslant 14.50800793$,

$i = 1, 2.$



Two global minimum points $y^{*,1} = (y_1^{*,1}, y_2^{*,1}) = (-12.87088549, 14.50800793)$,
$y^{*,2} = (y_2^{*,1}, y_1^{*,1}) = (14.50800793, -12.87088549)$,
$F(y^{*,1}) = F(y^{*,2}) = -186.7309088.$

III. Feasibility stage. For every $i = 1, 2$ each equation $y_j^{*,i} = f_j(x_j)$, $j = 1, 2$ has three solutions in x_j. Hence, for every $i = 1, 2$ we have 9 solutions that gives 18 global minimum points in total.



In this way we can fight against multiplicity of global minima.

Example 3. (Goldstein-Price function), (using linear transformation).

$$g(x_1,x_2)=\left(1+(x_1+x_2+1)^2\left(19-14x_1+3x_1^2-14x_2+6x_1x_2+3x_2^2\right)\right)\times$$

$$\times\left(30+(2x_1-3x_2)^2\left(18-32x_1+12x_1^2+48x_2-36x_1x_2+27x_2^2\right)\right),$$

$X=[-2,2]\times[-2,2]$.

By linear transformation of variables, g can be rewritten:

$$g(x_1,x_2)=\left(1+(x_1+x_2+1)^2\left(3(x_1+x_2+1)^2-20(x_1+x_2+1)+36\right)\right)\times$$

$$\times\left(30+(2x_1-3x_2)^2\left(3(2x_1-3x_2)-16(2x_1-3x_2)+18\right)\right).$$

Then

$$y_1=x_1+x_2+1,\ y_2=2x_1-3x_2. \tag{20}$$

For fixed (y_1,y_2) system (20) always has a unique solution in (x_1,x_2).

Example 3. (Goldstein-Price function), (using linear transformation).

$$F(y_1, y_2) = \left(1 + y_1^2\left(3y_1^2 - 20y_1 + 36\right)\right) \cdot \left(30 + y_2^2\left(3y_2^2 - 16y_2 + 18\right)\right)$$

— product of two univariate functions.

I. Bounding stage.

$$\underline{y}_1 = -3, \underline{y}_2 = -10, \overline{y}_1 = 5, \overline{y}_2 = 10.$$

II. Optimality stage.

$$\min F(y_1, y_2),$$

$$-3 \leqslant y_1 \leqslant 5, \ -10 \leqslant y_2 \leqslant 10.$$

Global minimum $y^* = (0, 3)$.

III. Feasibility stage.

$$x_1 + x_2 + 1 = 0 = y_1^*,$$
$$2x_1 - 3x_2 = 3 = y_2^*,$$

Solution $x^* = (0, -1)$.

Thank you for attention