Нижегородский государственный университет им. Н.И.Лобачевского

Национальный исследовательский университет

Институт информационных технологий, математики и механики

Системы принятия решений

Vladimir Grishagin

Оценки экстремума

Рассмотрим следующий простой пример. Пусть класс Φ - класс непрерывных функций, т.е. априорно известно, что минимизируемая функция $\varphi(y)$ непрерывна в области Q. Предположим, что вычислены значения функции в конечном числе точек $y^1, y^2, \dots, y^k$. Что после этого можно сказать о координате глобального минимума? Каковы бы ни были точки $y^1, y^2, \dots, y^k$ и значения $z^1, z^2, \dots, z^k$, для любой точки $y^* \in Q$ $(y^* \notin (y^1, \dots, y^k))$ всегда можно построить непрерывную функцию, проходящую через точки $(y^i, z^i), 1 \le i \le k$, т.е. принадлежащую классу $\Phi(\omega_k)$ из (1.11), которая имеет глобальный минимум в точке y^* с любым наперед заданным значением $\varphi^* < \min\limits_{1 \le i \le k} z^i$.

Например, в качестве такой функции можно взять интерполяционный полином k-й степени, проходящий через точки $(y^i, z^i), 1 \le i \le k,$ и точку (y^*, φ^*).

Оценки экстремума

Все сказанное означает, что по результатам конечного числа испытаний никаких выводов о расположении координаты глобального минимума сделать нельзя. Точно так же о величине φ^* глобального минимума можно лишь сказать, что

$$\varphi^* \le \varphi_k^* \tag{1.17}$$

где φ_k^* из (1.15), однако оценить величину

$$\varepsilon_k = \left|\varphi^* - \varphi_k^*\right| \tag{1.18}$$

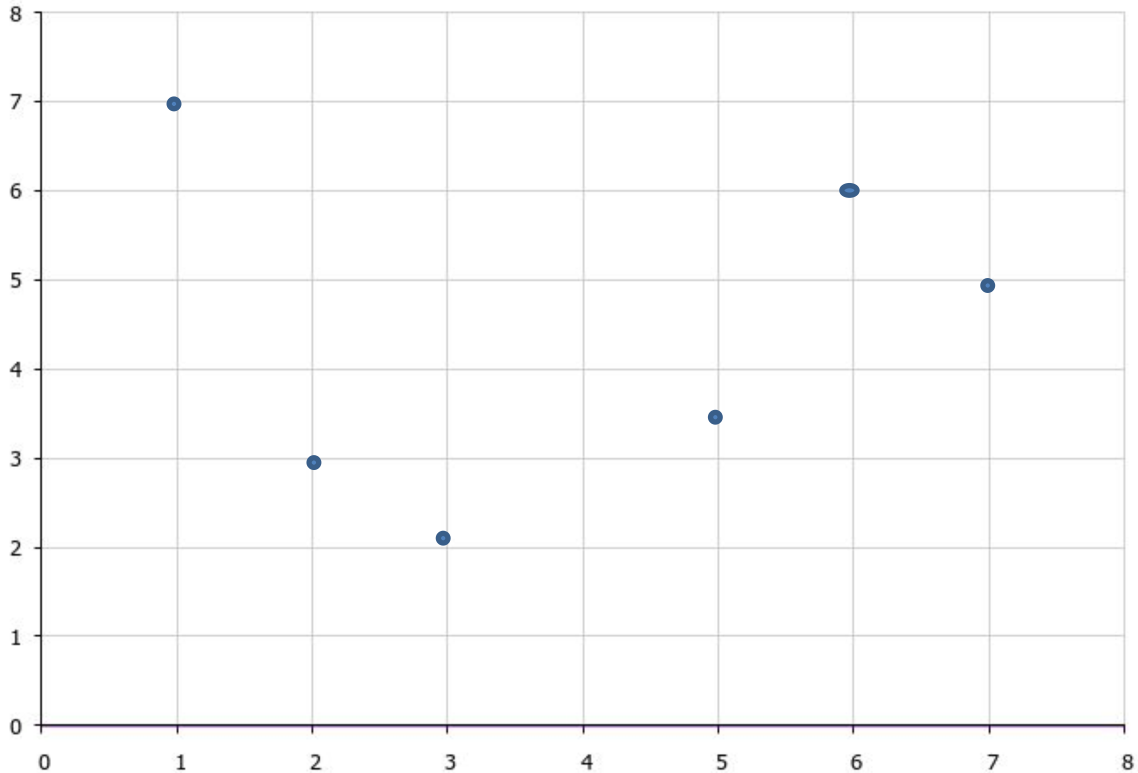
т.е. погрешность решения задачи, невозможно.

Возможность получения оценок экстремума по конечному числу испытаний зависит от свойств класса функций, которому принадлежит минимизируемая функция, или, другими словами, от априорной информации о функции . Фактически известны лишь два широких класса функций, допускающих построение таких оценок: класс одномерных унимодальных функций и класс функций, удовлетворяющих условию Липшица (в общем случае многомерных и многоэкстремальных).

Оценки экстремума для унимодальных функций

Функция $\varphi(y)$ является строго унимодальной на отрезке $[a,b]$, если существует точка $y^* \in [a,b]$ такая, что для всех $a \le y' < y'' < y^*$ выполняется $\varphi(y') > \varphi(y'') > \varphi(y^*)$, а для всех $y^* < y' < y'' \le b$ справедливо неравенство $\varphi(y^*) < \varphi(y') < \varphi(y'')$.

Пример. Унимодальная функция на отрезке *[0,8]*.



Оценки экстремума для унимодальных функций

Пусть теперь в общем случае проведено *k* испытаний в точках $y^1, y^2, ..., y^k \in (a,b)$ и получены значения $z^1, z^2, ..., z^k$. Перенумеруем точки испытаний нижним индексом в порядке возрастания координаты, добавив к ним также концы отрезка поиска *a* и *b*, т.е.

$$a = y_0 < y_1 < y_2 < ... < y_k < y_{k+1} = b \qquad (1.19)$$

Тогда интервалом неопределенности будет интервал (y_{i-1}, y_{i+1}), где номер *i* определяется из условия $y_i = y_k^*$, где y_k^* из (1.16) (в случаях *i=1* и *i=k* интервалами неопределенности будут полуинтервалы $[a, y_2)$ и $(y_{k-1}, b]$ соответственно). Иными словами, для строго унимодальной функции можно построить оценку координаты глобального минимума в виде интервала неопределенности и тем самым оценить погрешность решения задачи (по координате) величиной $max\{y_{i+1} - y_i, y_i - y_{i-1}\}$, ибо

$$\left| y_k^* - y^* \right| < \varepsilon = max\{y_{i+1} - y_i, y_i - y_{i-1}\}$$

Что касается величины глобального минимума, то строгой унимодальности для получения оценки (1.18) недостаточно и требуются более жесткие условия для ее реализуемости.

Оценки экстремума для липшицевых функций

Другим важным классом функций, допускающим построение оценок экстремума по конечному числу испытаний, является класс функций, удовлетворяющих условию Липшица

$$\left|\varphi(y')-\varphi(y'')\right| \le L\left\|y'-y''\right\|, \forall y', y'' \in Q \subseteq R^N \qquad (1.20)$$

где $L>0$ - постоянная величина, называемая константой Липшица.

Что означает неравенство (1.20)? Перепишем его в одномерном случае в виде

$$\frac{\left|\varphi(y')-\varphi(y'')\right|}{\left|y'-y''\right|} \le L, \forall y', y'' \in Q \subseteq R^N$$

Левая часть в последнем неравенстве – это наклон секущей, проходящей через точки $(y',\varphi(y'))$ и $(y'',\varphi(y''))$. Таким образом, класс липшицевых функций (1.20) – это класс функций с ограниченными (константой L) наклонами.

Является ли липшицева функция непрерывной?

Ответ положительный, поскольку согласно (1.20) малому приращению аргумента соответствует малое приращение функции.

Оценки экстремума для липшицевых функций

Является ли непрерывная функция липшицевой?

Не всегда. Пример: $\sqrt{y},\, y \in [0, 1]$. У нее неограниченный наклон около нуля.

Является ли липшицева функция дифференцируемой?

Не всегда. Пример: $|y|,\, y \in [-1, 1]$. Она не дифференцируема в нуле.

Является ли дифференцируемая функция липшицевой?

Да, является с константой $L = \max_{y \in Q} |\varphi'(y)|$.

Очевидно, что если функция является липшицевой с константой L, она будет липшицевой для любой константы $M > L$.

Неравенство (1.20) позволяет построить по результатам испытаний кусочно-линейную функцию, называемую нижней огибающей, или минорантой, которая в точках испытаний совпадает с вычисленными значениями функции $\varphi(y)$, а в остальных точках ограничивает $\varphi(y)$ снизу.

Рассмотрим одномерную задачу

$$\varphi(y) \to \min,\, y \in Q = [a, b] \subset R^1 \qquad (1.21)$$

Предположим, что мы провели *k* испытаний в точках $y^1, y^2, ..., y^k \in (a, b)$, получили значения функции $z^1, z^2, ..., z^k$ в этих точках и перенумеровали точки испытаний вместе с концами отрезка поиска *a* и *b* в соответствии с (1.19).

Оценки экстремума для липшицевых функций

Возьмем некоторую точку $y_i, 1 \le i \le k,$ из множества (1.19). Тогда для липшицевой функции, удовлетворяющей (1.20), для любого $y \in [a,b]$ справедливо неравенство

$$\varphi(y_i) - \varphi(y) \le L|y_i - y|,$$

или

$$\varphi(y) \ge \varphi(y_i) - L|y_i - y|, \qquad (1.22)$$

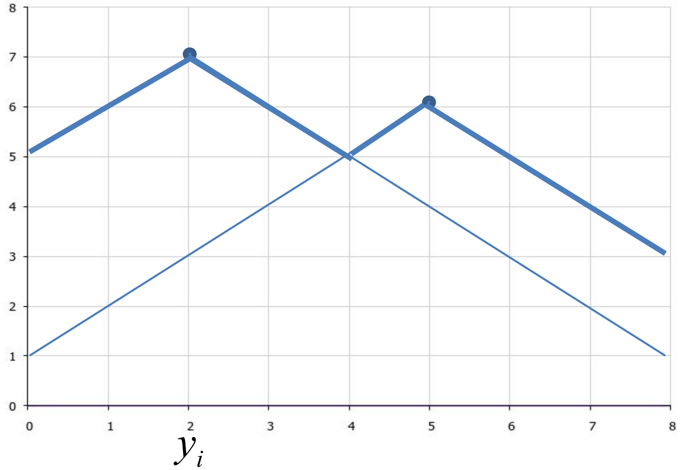
Последнее неравенство при $y \le y_i$, т.е. слева от y_i, имеет вид

$$\varphi(y) \ge \varphi(y_i) - Ly_i + Ly,$$

где в правой части стоит возрастающая линейная функция

$$\bar{l}_i(y) = \varphi(y_i) - Ly_i + Ly,\ y \le y_i, \qquad (1.23)$$

значение которой в точке y_i совпадает со значением $\varphi(y_i)$.



Аналогично при $y \ge y_i$ (справа от y_i)

$$\varphi(y) \ge \varphi(y_i) + Ly_i - Ly,$$

т.е. функция $\varphi(y)$ лежит над убывающей линейной функцией

$$\hat{l}_i(y) = \varphi(y_i) + Ly_i - Ly, y \ge y_i. \qquad (1.24)$$

Оценки экстремума для липшицевых функций

Если мы обозначим как

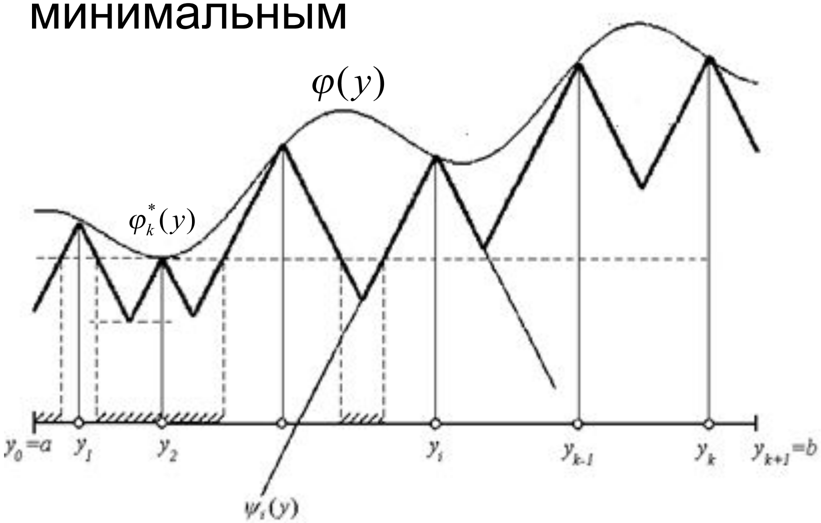
$$\psi_i(y) = \varphi(y_i) - L\left|y_i - y\right|$$

правую часть неравенства (1.22), то в общем случае

$$\varphi(y) \ge \Psi_k(y) = \max\{\psi_i(y), 1 \le i \le k\}, y \in [a,b], \qquad (1.25)$$

где функция $\Psi_k(y)$ - кусочно-линейная миноранта, построенная по результатам испытаний (1.19).

Рассмотрим величину $\Psi_k^* = \min\limits_{y \in [a,b]} \Psi_k(y)$. Тогда очевидно, что величина глобального минимума находится между величинами φ_k^* - минимальным вычисленным значением (1.15) целевой функции – и Ψ_k^* - минимальным значением миноранты. Это значит, что мы можем оценить погрешность решения (1.18) – погрешность по значению – величиной



$$\varepsilon_k = \varphi_k^* - \Psi_k^* \qquad (1.26)$$

а погрешность по координатам оценивается величиной области, в которой $\Psi_k(y) \le \varphi_k^*$.

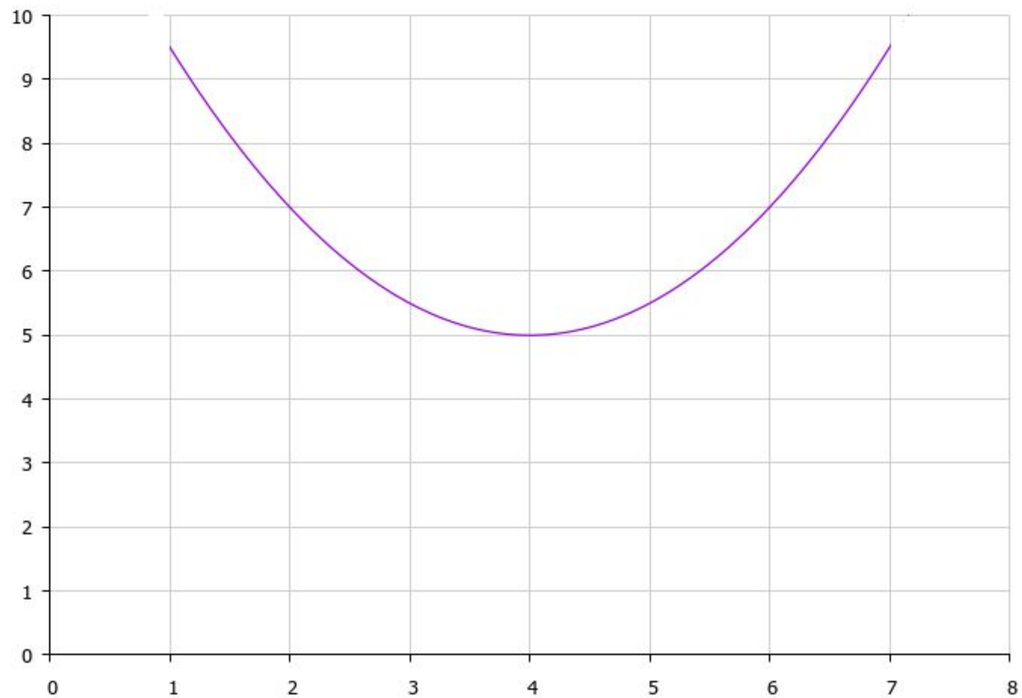
Оценки экстремума для липшицевых функций

Рассмотрим пример построения миноранты для конкретной функции

$$\varphi(y) = \frac{(y-4)^2}{2} + 5$$

на отрезке $[1, 7]$.

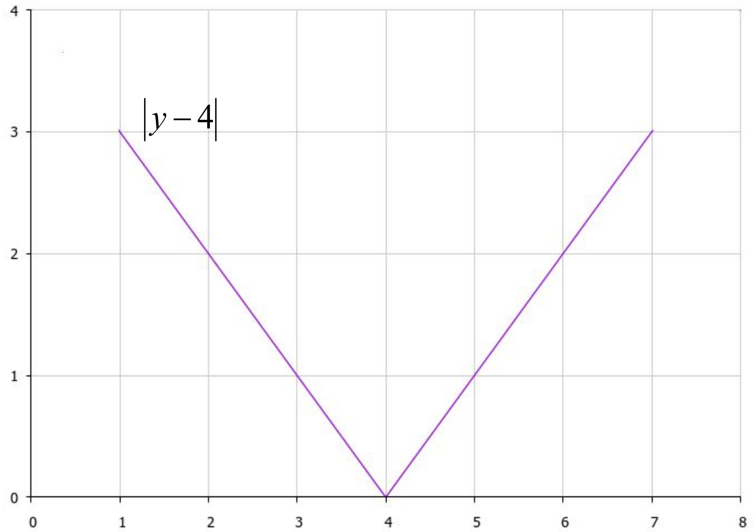
Это квадратичная функция с минимумом $\varphi^* = 5$ в точке $y^* = 4$ и значениями на концах отрезка – в точках 1 и 7 , равными 9.5.



Оценки экстремума для липшицевых функций

Производная этой функции $\varphi'(y) = y - 4$

Поскольку функция $\varphi(y)$ дифференцируема, ее константа Липшица *L* может быть посчитана как максимум на отрезке модуля производной,



Построим последовательно миноранту по 4 точкам:

$$y^1 = 1, y^2 = 7, y^3 = 2, y^4 = 5.$$

Значения функции $\varphi(y)$ в них равны

$$z^1 = \varphi(1) = 9.5, z^2 = \varphi(7) = 9.5, z^3 = \varphi(2) = 7, z^4 = \varphi(5) = 5.5.$$

Оценки экстремума для липшицевых функций

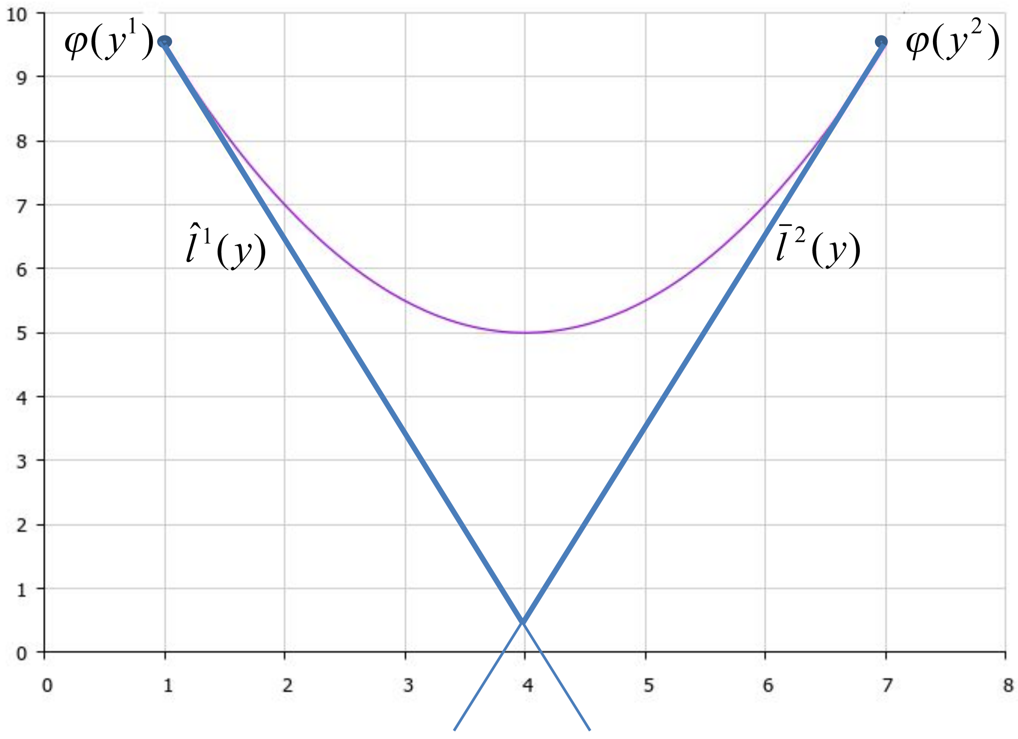
Начнем с точки $y^1 = 1$. Поскольку весь отрезок лежит справа от этой точки, то из двух минорирующих линейных функций (1.23), (1.24) возьмем только последнюю, которая имеет вид

$$\hat{l}^1(y) = \varphi(y^1) + Ly^1 - Ly = 12.5 - 3y,\ y \geq 1.$$

Следующая точка $y^2 = 7$. Отрезок целиком слева от точки, поэтому берем только линейную миноранту (1.23):

$$\bar{l}^2(y) = \varphi(y^2) - Ly^2 + Ly = 3y - 11.5, y \leq 7.$$

Теперь миноранта (1.25) $\Psi_2(y) = \max\{\hat{l}^1(y), \bar{l}^2(y)\}, y \in [a,b].$



Оценки экстремума для липшицевых функций

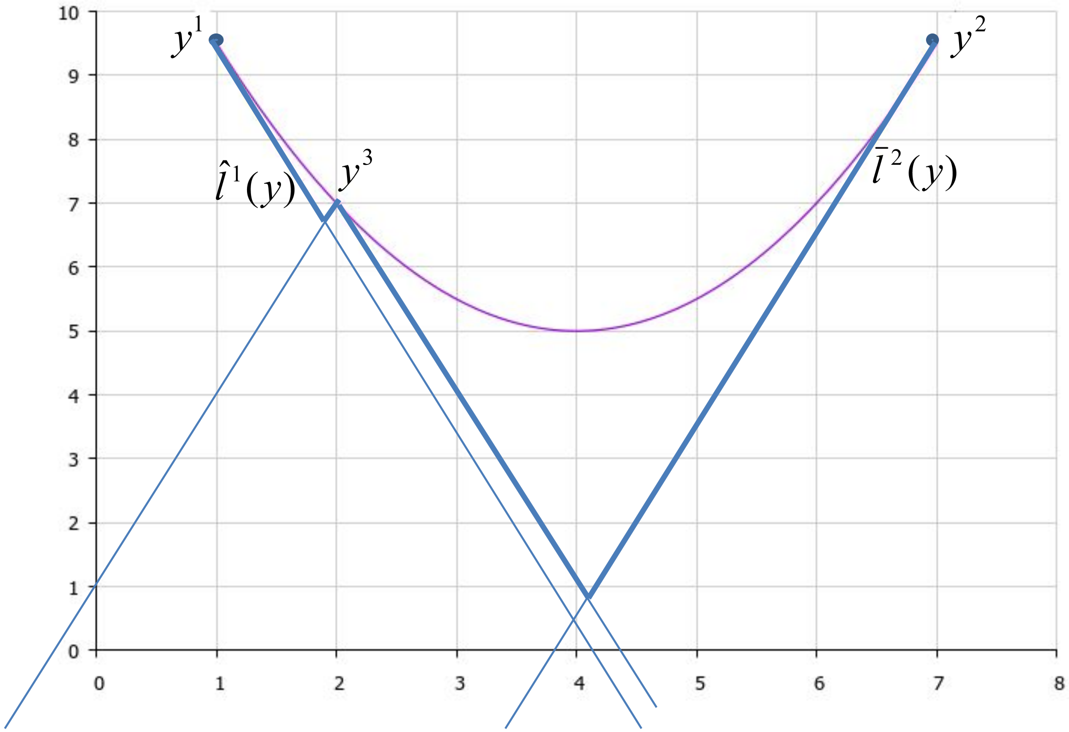
Берем точку $y^3 = 2$. Для нее функция (1.23) имеет вид

$$\bar{l}^3(y) = \varphi(y^3) - Ly^3 + Ly = 1 + 3y, y \le 2,$$

а функция (1.24) -

$$\hat{l}^3(y) = \varphi(y^3) + Ly^3 - Ly = 13 - 3y, y \ge 2.$$

Теперь строим миноранту $\Psi_3(y) = \max\{\hat{l}^1(y), \bar{l}^2(y), \bar{l}^3(y), \hat{l}^3(y)\}, y \in [a,b].$



Оценки экстремума для липшицевых функций

Последняя точка $y^4 = 5$. Функция (1.23) для нее

$$\bar{l}^4(y) = \varphi(y^4) - Ly^4 + Ly = 3y - 9.5, \; y \le 5,$$

а функция (1.24)

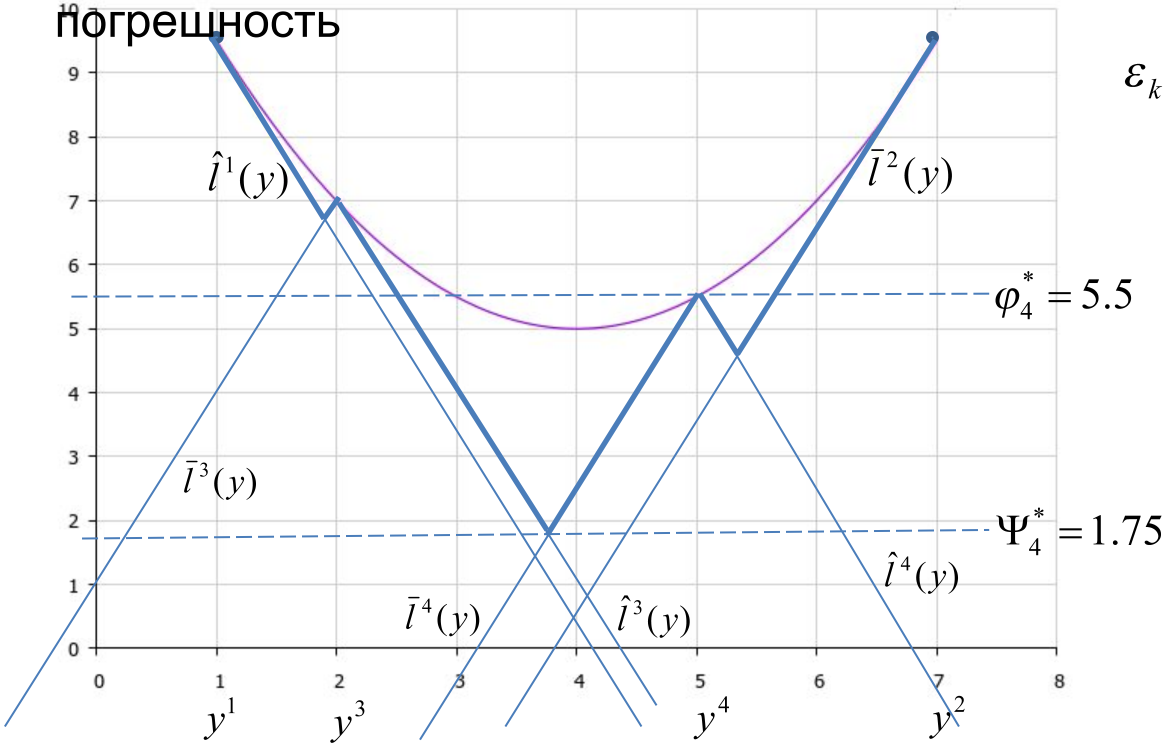
$$\hat{l}^4(y) = \varphi(y^4) + Ly^4 - Ly = 20.5 - 3y, \; y \ge 5.$$

Конструируем миноранту

$$\Psi_4(y) = \max\{\hat{l}^1(y), \bar{l}^2(y), \bar{l}^3(y), \hat{l}^3(y), \bar{l}^4(y), \hat{l}^4(y)\}, \; y \in [a,b].$$

Для глобального минимума φ^* имеем ограничение $\Psi_4^* \le \varphi^* \le \varphi_4^*$, т.е. погрешность

$$\varepsilon_k = \varphi_k^* - \Psi_k^* = 3.75$$



Estimates of extremum for Lipschitzian functions

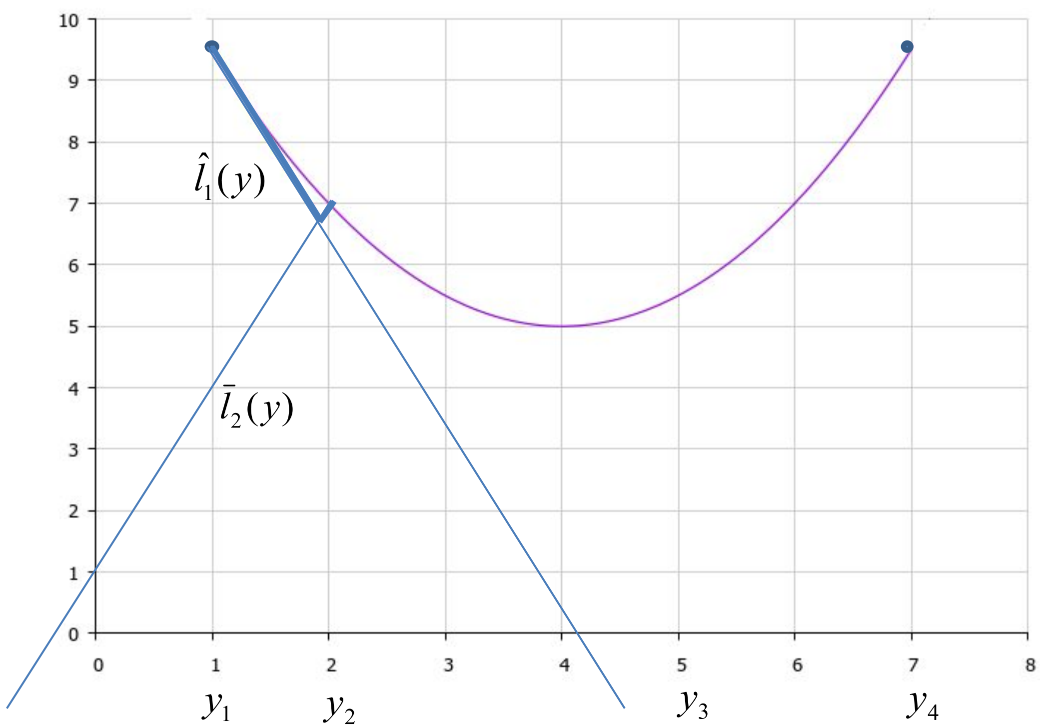
Мы строили миноранту, постепенно уточняя по мере поступления значений. Можно ее строить после того, как все значения получены . С этой целью упорядочим все координаты y^1, y^2, y^3, y^4 по возрастанию, перенумеровав их нижним индексом: $y_1 = 1, y_2 = 2, y_3 = 5, y_4 = 7$, и сопоставим им значения функции $z_1 = \varphi(1) = 9.5, z_2 = \varphi(2) = 7, z_3 = \varphi(5) = 5.5, z_4 = \varphi(7) = 9.5.$

Тогда для точки y_1 , как и ранее, строится только функция из (1.24)

$$\hat{l}_1(y) = z_1 + Ly_1 - Ly = \hat{l}^1(y) = 12.5 - 3y$$

Для точки $y_2 = 2$ получаем функцию (1.23)

$$\bar{l}_2(y) = z_2 + Ly_2 - Ly = \bar{l}^3(y) = 1 + 3y, y \le 2,$$



Функция $\xi_{1,2}(y) = \max_{y_1 \le y \le y_2} \{\hat{l}_1(y), \bar{l}_2(y)\}$ описывает часть миноранты на отрезке $[y_1, y_2]$.

Оценки экстремума для липшицевых функций

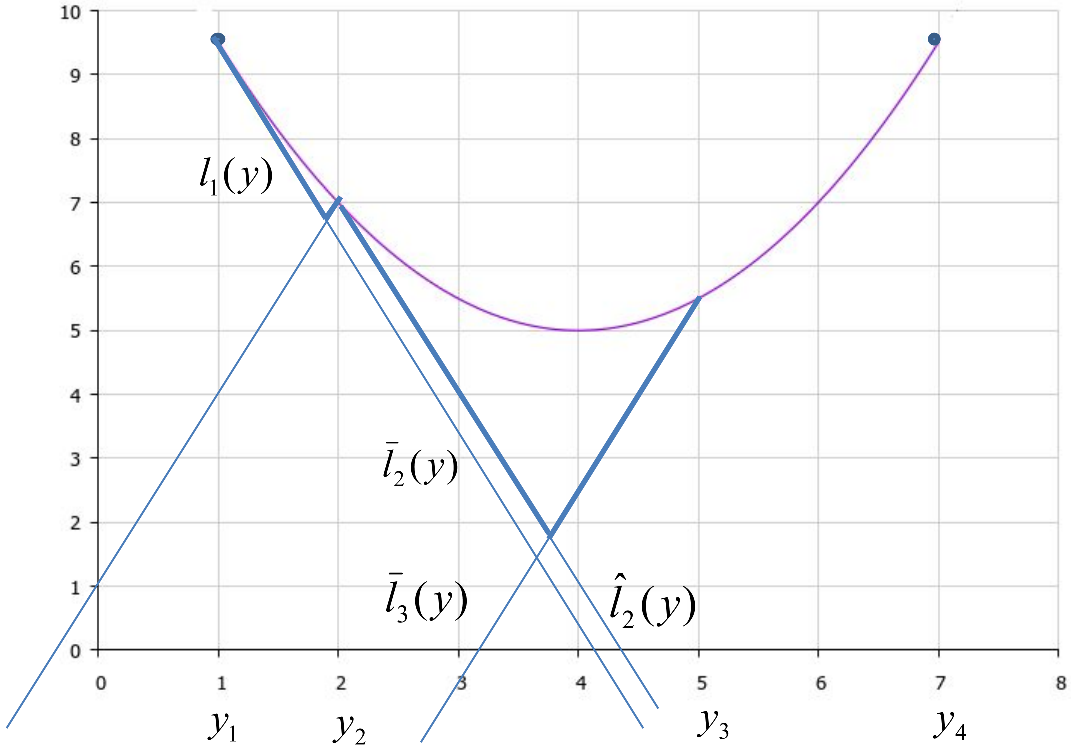
Для точки $y_2 = 2$ получаем функцию (1.24)

$$\hat{l}_2(y) = z_2 - Ly_2 + Ly = \hat{l}^3(y) = 13 - 3y, y \geq 2.$$

Для точки $y_3 = 5$ получаем функцию (1.23)

$$\bar{l}_3(y) = z_3 + Ly_3 - Ly = \bar{l}^4(y) = 3y - 9.5, y \leq 5,$$

Функция $\xi_{2,3}(y) = \max\limits_{y_2 \leq y \leq y_3} \{\hat{l}_2(y), \bar{l}_3(y)\}$ описывает «зубец» миноранты на отрезке $[y_2, y_3]$.



Оценки экстремума для липшицевых функций

Наконец, строим функцию (1.24) для точки $y_3 = 5$

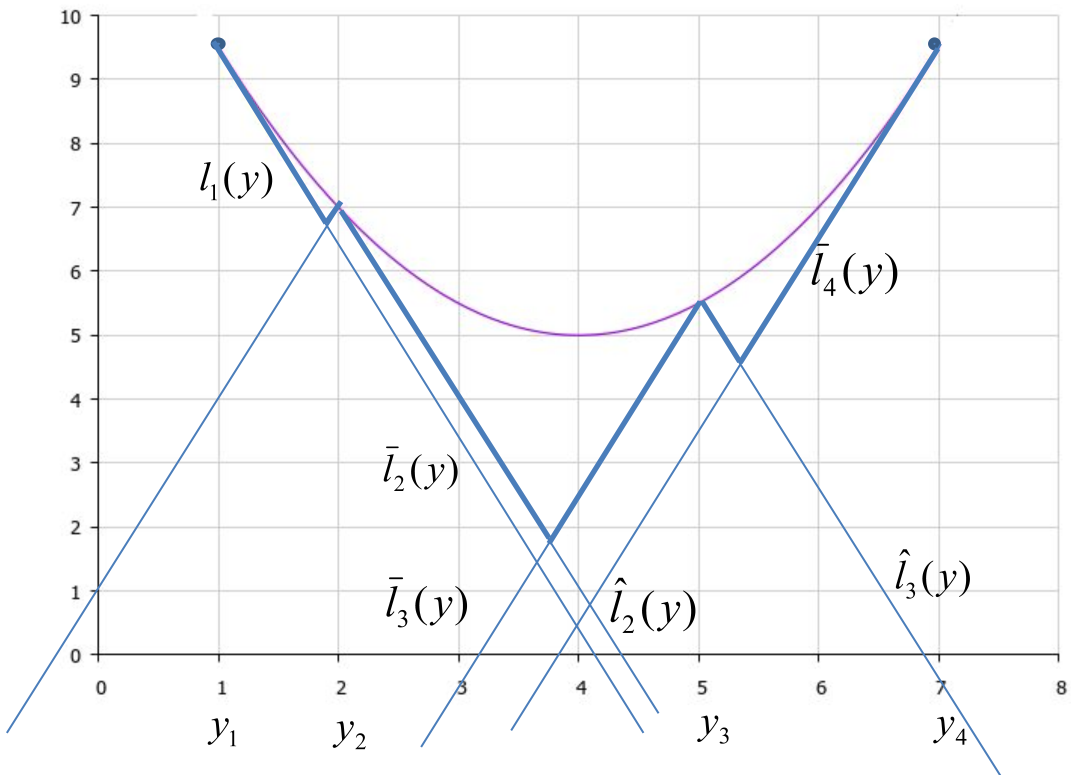
$$\hat{l}_3(y) = z_3 - Ly_3 + Ly = \hat{l}^4(y) = 20.5 - 3y,\ y \ge 5,$$

и функцию (1.23)

$$\bar{l}_4(y) = z^4 - Ly^4 + Ly = \bar{l}^2(y) = 3y - 11.5,\ y \le 7.$$

для точки $y_4 = 7$ чтобы получить последнюю часть миноранты на отрезке $[y_3, y_4]$.

$$\xi_{3,4}(y) = \max_{y_3 \le y \le y_4} \{\hat{l}_3(y), \bar{l}_4(y)\}$$



Оценки экстремума для липшицевых функций

Рассмотрим две соседних точки $y_{i-1} < y_i$ из ряда (1.19) со значениями $z_{i-1} = \varphi(y_{i-1}),\ z_i = \varphi(y_i)$. Зубец миноранты между ними описывается функцией

$$\xi_{i-1,i}(y) = \max_{y_{i-1} \le y \le y_i} \{\hat{l}_{i-1}(y), \bar{l}_i(y)\}$$

где

$$\hat{l}_{i-1}(y) = z_{i-1} + Ly_{i-1} - Ly,\ y \ge y_{i-1},$$

$$\bar{l}_i(y) = z_i - Ly_i + Ly,\ y \le y_i,$$

Точка $\tilde{y}_i$ пересечения этих двух функций является точкой минимума функции $\xi_{i-1,i}(y)$ на отрезке $[y_{i-1}, y_i]$.

Ее легко найти из равенства

$$z_{i-1} + Ly_{i-1} - Ly = z_i - Ly_i + Ly,$$

откуда

$$\tilde{y}_i = \frac{1}{2}(y_{i-1} + y_i) - \frac{z_i - z_{i-1}}{2L}. \qquad (1.26)$$

Значением миноранты в точке $\tilde{y}_i$ является величина $\bar{l}_i(\tilde{y}_i)$, равная

$$\tilde{z}_i = \frac{1}{2}(z_{i-1} + z_i) - \frac{L}{2}(y_i - y_{i-1}). \qquad (1.27)$$

Построение миноранты

Задание

Построить для константы Липшица $L = 15$ миноранту функции $y^3 - 5y^2 + 6y + 2$ на интервале [0,4] по точкам испытаний

$$y_1 = 0, y_2 = 1, y_3 = 2, y_4 = 4$$

и указать оценку величины глобального минимума $\varepsilon_k = \varphi_k^* - \Psi_k^* = 3.75$