Ближайшие к нулевому векторы линейного многообразия. Чебышевские аппроксимации могут обходиться без условия Хаара

Зоркальцев Валерий Иванович

Иркутск, 2019

Линейные чебышевские приближения (для функций на отрезке)

Заданы непрерывные на отрезке $[a, b]$:

- линейно независимые функции $v_i = 1,\ldots, m$;
- аппроксимируемая функция w;
- функция весов $h(x) > 0,\ x \in [a, b]$.

Необходимо найти:

$$\overline{\lambda} = \arg\min\left\{\rho(\lambda) : \lambda \in R^m\right\},$$

где $\rho(\lambda) = \sup\left\{h(x)\left|\omega(x) - \sum_{i=1}^{m}\lambda_i v_i(x)\right| : x \in [a, b]\right\}.$

Чебышевской аппроксимацией является функция

$$\overline{\omega}(x) = \sum_{i=1}^{m}\overline{\lambda}_i v_i(x).$$

Пример: аппроксимация полиномами, когда $v_i(x) = (x)^{i-1}$.

Условие Хаара (1918 г.), чебышевские системы

Следующие три условия равносильны:

1. При любом $\lambda \neq 0$ имеет на $[a, b]$ не более $m-1$ нулей функция

$$f_\lambda(x) = \sum_{i=1}^{m} \lambda_i v_i(x).$$

2. При любых различающихся χ^i из $[a, b]$, $i = 1,\ldots, m$ ненулевой определитель матрицы

$$G = [g_{ij} = v_i(x^j)],\ i = 1,\ldots,\ m,\ j - 1,\ldots,\ m.$$

3. Дискретная чебышевская аппроксимация для попарно различающихся более m точек имеет единственное решение.

При выполнении этих (равносильных) условий существует и единственное чебышевское приближение.

Оказывается, можно получать единственное, разумное чебышевское приближение дискретной и непрерывной задач без этих условий!


Magic of Baikal © Alexey Trofimov

Общая постановка проблемы

Многие прикладные задачи сводятся к геометрической проблеме поиска наименее удалённой от начала координат точки линейного многообразия:

$$\sigma(x) \to \min, \; x \in X.$$

Здесь X – линейное многообразие в R^n, $\sigma(x)$ – некоторая характеристика близости вектора x к началу координат. Могут использоваться разные способы задания σ.

Как выбор σ влияет на результат?

Важность проведения исследований для множеств линейной структуры

Непосредственное обобщение: X – полиэдр (множество решений системы линейных неравенств).

1. Многие модели описываются в виде системы линейных уравнений и неравенств.
2. К линейным системам можно прийти в результате преобразований или линеаризации нелинейных задач.
3. Линейные системы – необходимая основа для понимания и решения нелинейных задач.

Приложения в моделировании

1. Задачи оценки параметров зависимостей по имеющимся наблюдениям (линейная регрессия).
2. Выделение и прогнозирование составляющих временных рядов.
3. Поиск допустимых решений максимально приближенных к заданному недопустимому.
4. Экономические модели с минимизацией ущербов от дефицита и избытка продукции.
5. Многокритериальные задачи оптимизации. Мультисубъектная оптимизация (согласование интересов).

Приложения в вычислительных методах

1. Алгоритмы метода внутренних точек в математическом программировании.
2. Методы использующие линеаризацию.
3. Регуляризация:

3.1 Поиск псевдорешений задач с противоречивыми условиями.

3.2 Поиск нормальных (ближайших к началу координат) решений в целях противодействия неединственности и неустойчивости решений.

Решение системы нелинейных уравнений методом линеаризации

Задача. Найти $x \in R^n$, при котором

$$F(x) = 0,$$

где

$$F: R^n \to R^m.$$

Алгоритм. Задан $x^0 \in R^n$. Итеративный переход:

$$x^{k+1} = x^k + \lambda_k s^k, \; k = 0, 1, 2, \ldots$$

Где λ_k – шаг (в частности $\lambda_k = 1$), s^k – направление корректировки, являющиеся решением системы линейных уравнений

$$F(x^k) + \nabla F(x^k) s = 0$$

с минимальной нормой

$$\|s\| \to \min,$$

в целях минимизации погрешности от линеаризации.

Какой нормой лучше пользоваться?

Два примера линейных многообразий и связанные с ним задачи

1. $$X = \{x = d - G\lambda : \lambda \in R^m\}.$$

Оценка параметров линейной регрессии, поиск псевдорешений несовместной СЛАУ.

2. $$X = \{x \in R^n : Ax = b\}.$$

Поиск нормального решения СЛАУ, поиск допустимого решения максимально приближенного к заданному.

Заданы векторы $d \in R^n$, $b \in R^r$, матрицы A, G размера $r \times n$ и $m \times n$.

Возможные способы оценки параметров линейной регрессии

Пусть h_j – заданные положительные весовые коэффициенты, $j = 1, ..., n.$

1. Минимизация суммы модулей отклонений (октаэдрические нормы)

$$\sum h_j |x_j| \to \min, \; x \in X.$$

2. Метод наименьших квадратов

$$\sum h_j (x_j)^2 \to \min, \; x \in X.$$

3. Минимизация максимального отклонения (чебышевские нормы)

$$\max_{j=1,...,n} h_j |x_j| \to \min, \; x \in X.$$

Возможные причины введения весовых коэффициентов

1. Учёт разной точности наблюдений.
2. Управляемые параметры для достижения желаемых результатов.
3. Соизмерение показателей разной природы.
4. Характеристики «ущербов» от отклонений.
5. Учёт разной информативности наблюдений.

Возникающие вопросы

1. Как соотносятся решения получаемые минимизацией октаэдральной, евклидовой, чебышевской и других возможных норм?
2. Как влияет выбор вектора весовых коэффициентов h на получаемые решения ?
3. Как взаимосвязаны решения, получаемые при иных, в том числе более общих постановок (минимизация штрафных функций, парето – оптимизация) .

Первая конкретизация проблемы

Поиск векторов с минимальным (не сужаемым) носителем составляющих множество

$$B = \{y \in X : \neg\exists x \in X, \quad J(x) \subset J(y)\},$$

где $\subset$ – строгое включение,

$J(x) = \{i : x \neq 0\}$ – носитель вектора $x \in R^n$.

Теорема 1. *Число векторов в* B *конечно, не более* $C_n^m = \dfrac{n!}{m!(n-m)!}$, *где* m *– размерность* X.

Вторая конкретизация проблемы

Найти вектор из множества Парето – оптимальных решений $Q = \{q \in X : \neg\exists y \in X, |y| \le |q|\}$ многокритериальной задачи минимизации абсолютных значений компонент вектора из X.

Здесь $|y|$ – вектор R^n, составленный из абсолютных значений компонент вектора y, неравенство $x \le z$ означает полуположительность вектора z-y (все компоненты неотрицательные и хотя бы одна положительная).

Теорема 2. $B \subseteq Q$.

Теорема 3. $Q \subseteq co\, B$.

Следствие. $co\, Q = co\, B$.

Устройство Парето – оптимальных решений

Пусть q^i, $i=1,...,l$ – максимальный набор векторов Q с нерасширяемыми в Q и различающимися наборами.

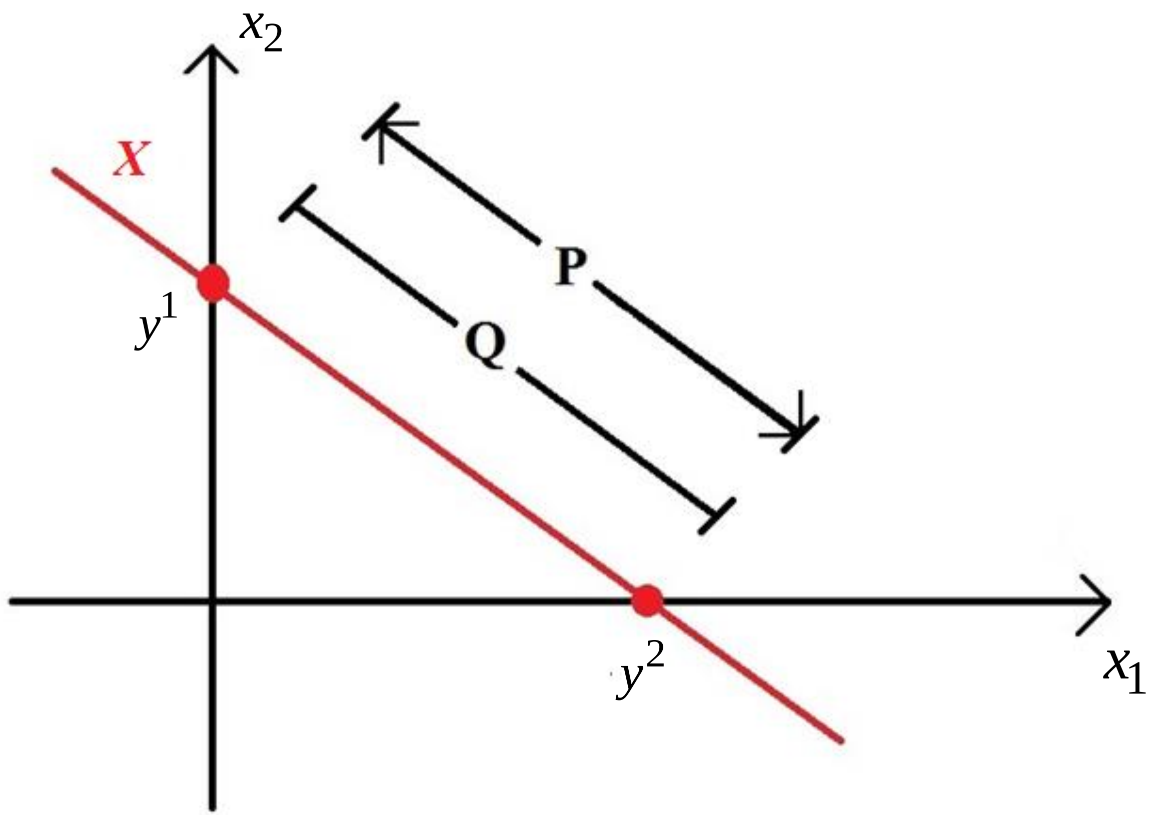
$B^i = B(q^i)$ – особые вектора с суженными или одинаковыми наборами относительно q^i.

$Q^i = co B^i$ – политоп, порождённый B^i.

Множество Q может быть невыпуклым, но оно является объединением конечного числа указанных политопов.

Теорема 4. *Множество Q является:*

1) *замкнутое,* $\mathrm{cl}\, Q = Q$;
2) *связным;*
3) *ограниченным, объединением конечного числа политопов.*



Одномерное линейное многообразие в R^2 *с двумя векторами с минимальным носителем*

$$B = \{y^1, y^2\},\ y^1 = (1, 0),\ y^2 = (0, 1)$$

$$Q = \{q = \lambda y^1 + (1-\lambda) y^2 : \lambda \in [0, 1]\},$$

$$P = \{q = \lambda y^1 + (1-\lambda) y^2 : \lambda \in (0, 1)\}.$$


Magic of Baikal © Alex

Третья конкретизация проблемы

Минимизация штрафной функции: Найти вектор

$$y(f) = \arg\min\{ f(y) : y \in Y\}.$$

Пусть F - множество дифференцируемых функций f от векторов R^n:

1) преобразуемых в результате дифференцируемого возрастающего преобразования в строго выпуклые функции;

2) таких, что для всех $x \in R^n$

$$sign \nabla_j f(x) = sign\, x_j, \quad j = 1, \ldots, n.$$

Теорема 5. *Если* $f \in F$, *то существует и единственный вектор y(f).*

Гёльдеровские и евклидовы нормы

Множеству F принадлежат *гёльдеровские нормы*

$$\rho_p^h(y) = (\sum |h_j y_j|^p)^{1/p}.$$

Заданы: $h \in R_{++}^n$ – вектор весовых коэффициентов, $p \in (1, \infty)$ – степенной коэффициент. Им соответствуют строго выпуклые функции

$$\varphi_p^h(y) = \sum |h_j y_j|^p.$$

При $p = 2$ гёльдеровские нормы являются *евклидовыми нормами*. При штрафной функции

$$f = \varphi_2^h$$

проблема решается методом наименьших квадратов.

Проекции начала координат на линейное многообразие

Определим множество гёльдеровских проекций:
при фиксированном $p > 1$

$$P_p = \{y(\varphi_p^h) : h \in R_+^n\},$$

и при всех $p > 1$

$$P = \bigcup_{p>1} P_p.$$

При p=2 имеем множество евклидовых проекций P_2.

Множество точек минимумов штрафных функций

$$PF = \{ y(f) : \ \ f \in F \}.$$

Теорема 6. $P_p = P = PF, \ \forall p > 1.$

Теорема 7. $PF \subseteq Q.$

Теорема 8. $cl\, PF = Q.$

Следствие. $\boldsymbol{P_2 = P = PF}$, $cl\, P_2 = Q.$

Преимущества метода наименьших квадратов

1. Поиск вектора $x(\varphi_2^h)$ сводится к решению СЛАУ с симметричной неотрицательно определенной матрицей (метод Халецкого, французский военный топограф, погиб в 1918 г. в Сербии).

2. **Теорема 9**. *Вектор* $x(\varphi_2^h)$ *является непрерывным отображением вектора* $h>0$.

3. За счет выбора весовых коэффициентов $h>0$ можно получить с требуемой точностью любое решение из любой рассмотренной постановки.

Старый друг – не хуже новых двух!

Октаэдральные и чебышевские нормы

Гёльдеровские нормы при $p \to 1$ и при $p \to \infty$ переходят в недифференцируемые функции – *октаэдраьные и чебышевские нормы*

$$\rho_1^h(y) = \sum h_j \, | \, y_j \, |.$$

$$\rho_\infty^h(y) = \max_j h_j \, | \, y_j \, |.$$

Для любого $h \in R_{++}^n$, любого $y \in R^n$

$$\lim_{p \to 1} \rho_p^h(y) = \rho_1^h(y); \quad \lim_{p \to \infty} \rho_p^h(y) = \rho_\infty^h(y).$$

Октаэдральные проекции

Октаэдральные проекции при данном векторе весовых коэффициентов $h \in R^n_{++}$ могут быть не единственны. Введём множество таких проекций

$$Y_1(h) = Arg\min\{\rho_1^h(y) : y \in Y\}.$$

Это множество является *политопом*. Существует только конечное число различающихся множеств $Y_1(h)$.

Теорема 10. *При любом* $h \in R^n_{++}$

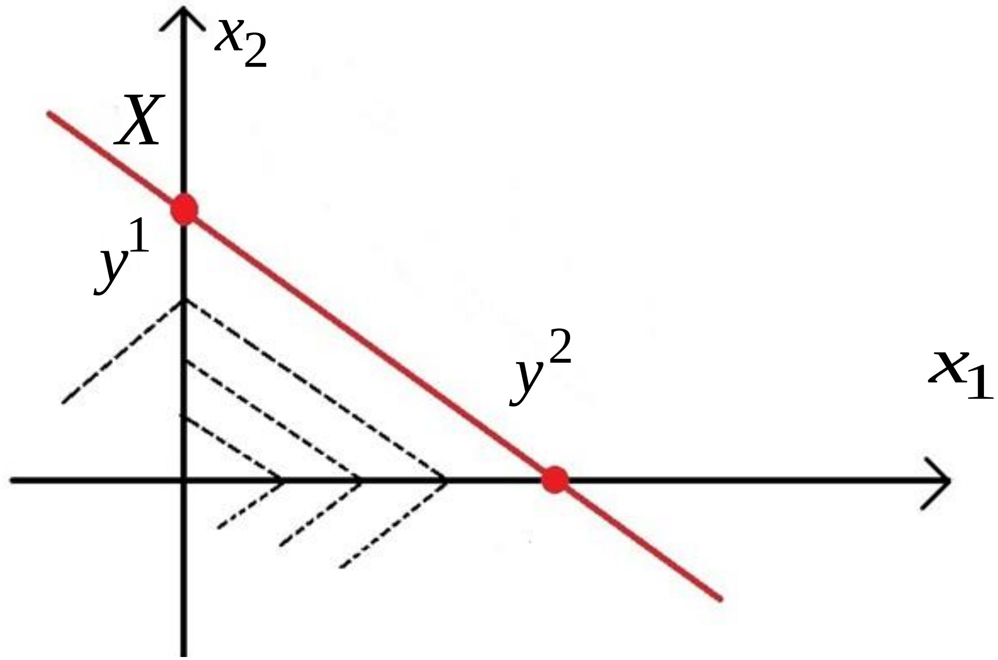
$$B(h) \neq \varnothing,\ Y_1(h) = \mathrm{co}\, B(h),$$

где

$$B(h) = Y_1(h) \cap B.$$

Теорема 11. $P_1 = Q,$ *где* $P_1 = \bigcup_{h \in R^n_+} Y_1(h).$

Октаэдральные проекции начала координат на линейное многообразие



Изображены пунктиром линии уровня функции ρ_1^h.

Если $h_1 = h_2$, *то* $Y(\bar{h}) = [y^1, y^2]$,

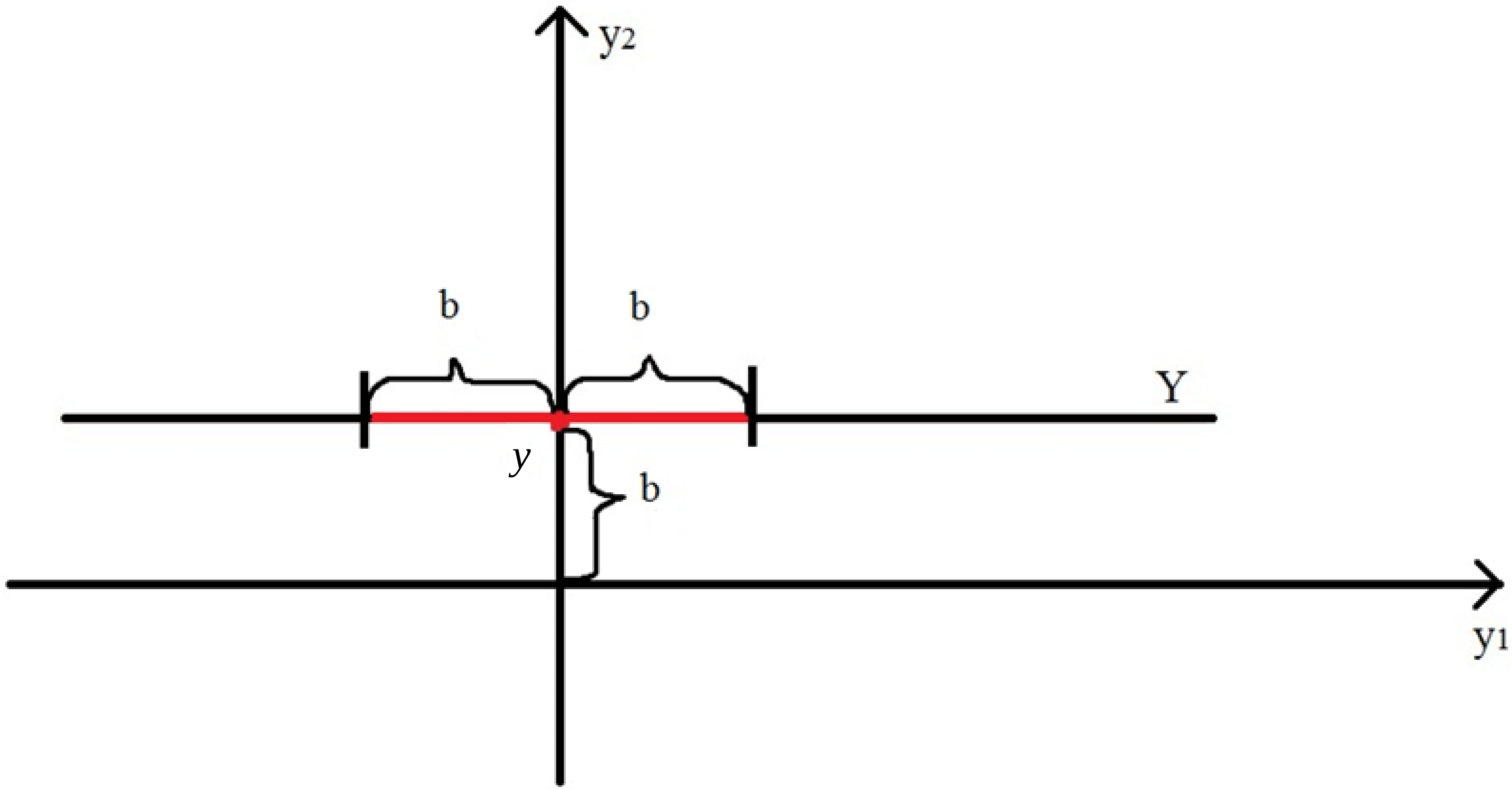
если $\dfrac{h_1}{h_2} > 1$, *то* $Y(h) = \{y^2\}$, *если* $\dfrac{h_1}{h_2} < 1$, *то* $Y(h) = \{y^1\}$.


gic of Baikal © Alexey Trofimov

Чебышевские проекции

Задача минимизации чебышевской нормы на линейное многообразие может иметь неединственное решение, среди которых могут быть явно неудовлетворительные по содержательным соображениям, не находящиеся в Q.

Предложен алгоритм, основанный на вычислениях относительно внутренних точек оптимальных решений последовательности задач линейного программирования, всегда приводящий к однозначному решению $x(\rho_\infty^h)$, находящемуся в Q.



Чебышевские проекции на линейное многообразие $x_2 = b.$

До использования лексикографической оптимизации – весь выделенный интервал длины *2b*.

После лексикографической оптимизации остаётся точка

$$y = x(\rho_{\infty}^{h}).$$

Предлагаемый алгоритм вычисления чебышевской проекции

Этапы $t=1,2,...,T.$ Пусть α^t, x^t оптимальные решения *с минимальным (не сужаемым) набором активных ограничений* задачи ЛП:

$$\alpha \to \min, \; x \in X,$$
$$-\alpha \le h_i x_j \le \alpha, \; i \in J^t,$$
$$x_j = \bar{x}_j, \; j \in R^t,$$

где

$$R^1 = \emptyset, \;\; J^1 = \{ j = 1,...,n\};$$
$$I^t = \left\{ j : h_j \left| x_j^t \right| = \alpha^t \right\}; \;\; R^{t+1} = R^t \bigcup I^t; \;\; J^{t+1} = J^t / I^t.$$

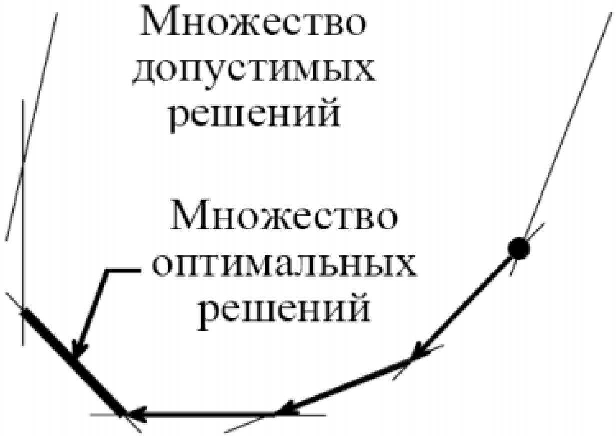
Положим $\bar{x}_j = x_j^t; \;\; j \in I^t.$

На всех этапах $I^t \neq \emptyset$. Поэтому будет этап $t=T$, когда $J^{t+1} = \emptyset$. Полагаем

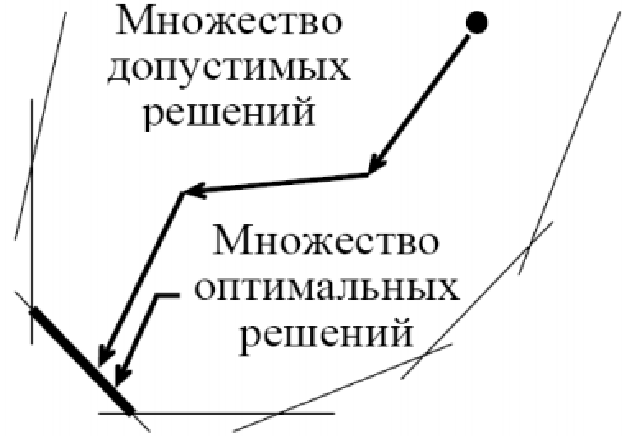
$$x(\rho_\infty^{\mathrm{h}}) = x^T.$$

Различия двух типов алгоритмов

1. Симплекс-метод улучшает решение на границе полиэдра векторов удовлетворяющих условиям – неравенствам. Приводят к граничным оптимальным решениям.
2. Внутренние точки улучшают решения по траекториям внутри полиэдра. Приводят к относительно внутренним точкам оптимальных решений с минимальными наборами активных ограничений.



Траектория симплекс-метода



Траектория методов внутренних точек

Свойства чебышевских проекции

Пусть

$$P_\infty = \{ x(\rho_\infty^h) : h \in R_{++}^n \}.$$

Теорема 12. $P_\infty \subseteq Q.$

Теорема 13. $P_\infty = P_2.$

Теорема 14. *Для любого* $h \in R_{++}^n$ *при* $p \to \infty$

$$x(\rho_p^h) \to x(\rho_\infty^h).$$

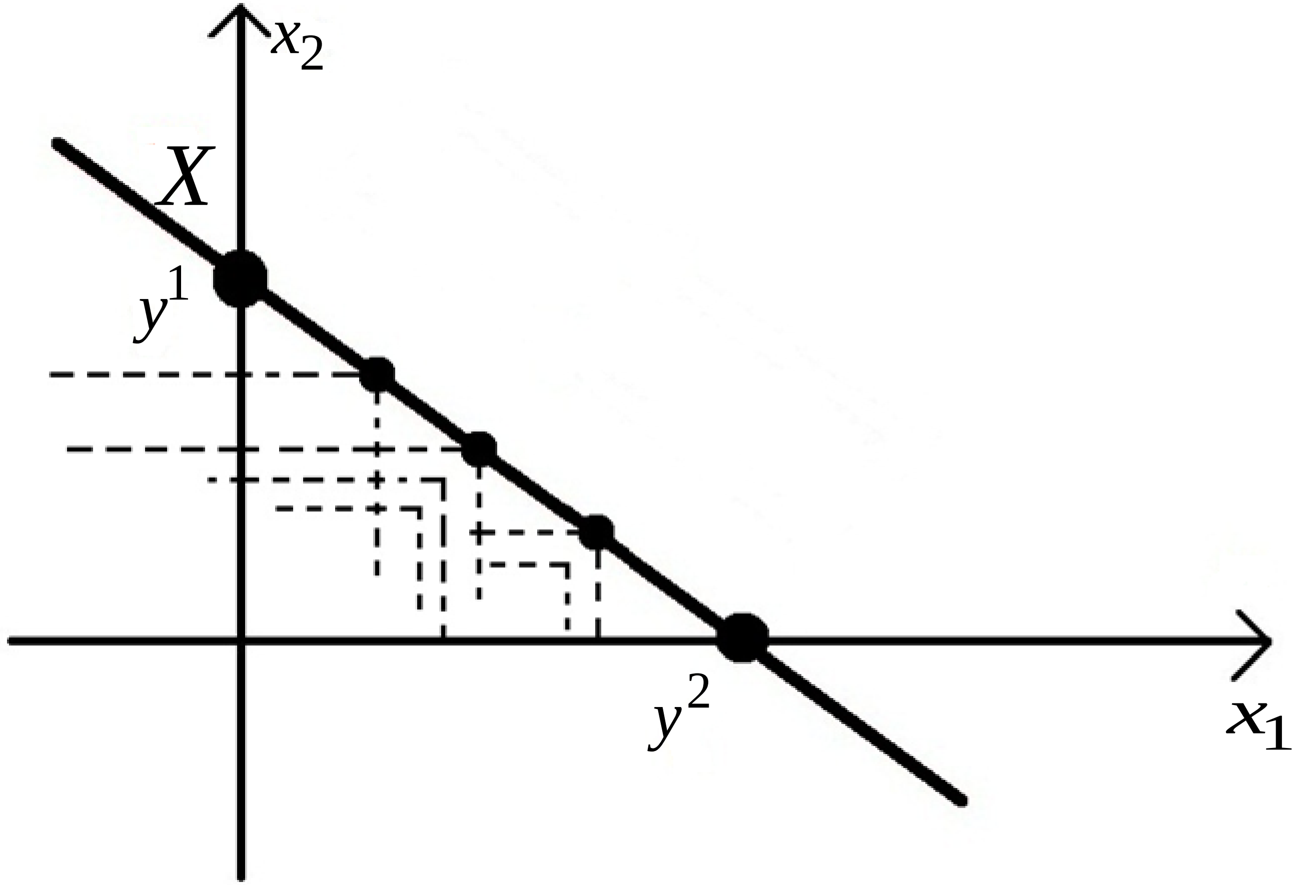
Гипотеза. $x(\rho_\infty^h)$ – *непрерывная вектор функция от* h.

НА ОСНОВЕ ТЕОРЕМЫ 14

$$x(\rho_p^h) \to x(\rho_\infty^h) \text{ при } p \to \infty$$

можно:

1) дать другое (равносильное) определение однозначной, разумной чебышевской проекции;
2) предложить иной алгоритм вычисления чебышевской проекции – на основе итеративного уточнения гельдеровской проекции с итеративным повышением степенного коэффициента p.



Чебышевские проекции начала координат на линейное многообразие при различных весовых коэффициентах занимают открытый интервал (y^1, y^2).

Пунктиром изображены линии уровня функции ρ_∞^h при разных h.

Двумерное линейное многообразие

Задано в виде множества решений СЛАУ из двух уравнений с четырьмя переменными:

$$\begin{aligned} x_1 + x_2 + x_3 \quad\quad &= 6, \\ x_3 + x_4 &= 8. \end{aligned}$$

При $h_j = 1,\ j = 1, \ldots, \mathrm{n}$ решения с минимальной чебышевской нормой составляют интервал

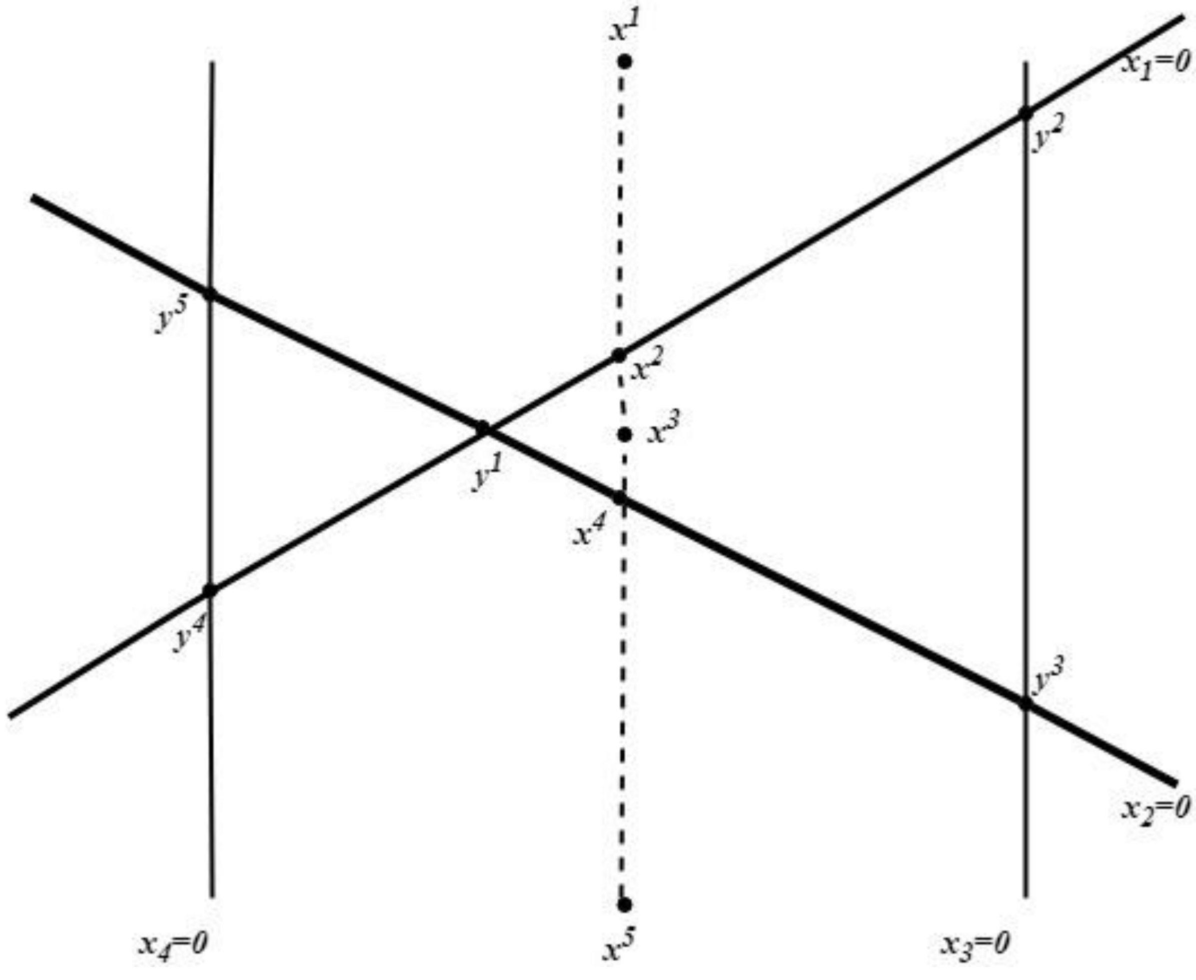
$$x = \lambda x^1 + (1-\lambda) x^5 \text{ при } \lambda \in [0, 1],$$

где

$$x^1 = (-2, 4, 4, 4),\ x^5 = (4, -2, 4, 4).$$

Чебышевской проекцией будет при $\lambda = 0{,}5 \quad x(\rho_\infty^h) = (1, 1, 4, 4).$

Множество векторов с парето-минимальным абсолютными значениями компонент Q и векторы линейного многообразия с минимальным носителем $y^i, i=1,...,5$ в рассматриваемом примере



Предложенные два равносильные определения однозначной гельдеровской проекции для дискретных задач позволяют:

1) решать во всех случаях задачу линейной чебышевской аппроксимации;
2) расширяют возможности для решения задач нелинейной чебышевской аппроксимации!

Спаси бо за внимание!