Взвешенные евклидовы проекции на линейное многообразие

Доклад на семинаре «Оптимизация, машинное обучение, исскуственный интеллект»
13 февраля 2025 г.

Зоркальцев Валерий Иванович, zorkal.com

Байкальский государственный университет,
г. Иркутск , vizork@mail.ru.

Иркутск – Санкт-Петербург, 2025


Magic of Baikal © Alexey Trofimov

Рассматриваемые проблемы

Многие прикладные задачи сводятся к геометрической проблеме поиска наименее удалённой от начала координат точки некоторого линейного многообразие Y в R^n при различных определениях понятия «близости». В данном докладе представлены результаты исследования вопросов:

- как различаются решения получаемые при различных определениях понятия «близость»;
- какими свойствами, достоинствами и недостатками обладают решения при разных определениях этого понятия.

Исследования данных вопросов (в течении примерно 40 лет) обобщены в монографии:

В. И. Зоркальцев

Свойства методов аппроксимации (ближайшие к началу координат точки линейных многообразий)

Научный редактор почетный профессор Санкт-Петербургского государственного университета, д-р. физ.-мат. наук В. Н. Малозёмов

Рецензенты:

канд. физ.-мат. наук С. П. Епифанов;

д-р. физ.-мат. наук В. И. Ерохин;

д-р. физ.-мат. наук Л. Д. Попов.

Иркутск, БГУ 2025

Данный доклад - развитие материалов доклада 30 марта 2023 г.

«Наименее удаленные от начала координат точки линейного многообразия»

на семинаре «Оптимизация, машинное обучение, искусственный интеллект».

Важность проведения исследований для множеств линейной структуры

Непосредственное обобщение: Y – 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. Оптимизация методом внутренних точек..
2. Методы использующие линеаризацию.
3. Алгоритмы определения согласованных параметров балансовых моделей ближайших к заданным несогласованным.
4. Регуляризация:

4.1. Псевдорешения при несовместных условиях.

4.2. Поиск решений ближайших к нулевому вектору, в т. ч. для противодействия неединственности, неустойчивости, неограниченности решений.

Линейное многообразие

Множество Y из R^n является линейным многообразием, если оно замкнуто относительно операции взятия аффинных комбинаций, то есть для любых x^1, x^2 из Y при любом вещественном λ вектор

$$y = \lambda x^1 + (1 - \lambda) x^2$$

находится в Y.

Примеры.

1. Сдвиг на вектор $d \in R^n$ образа $n \times m$ матрицы G

$$Y = \{y = d - G\lambda : \lambda \in R^m\}.$$

2. Множество решений системы линейных уравнений: при заданных $b \in R^r$, $r \times n$ матрице A

$$Y = \{y \in R^n : Ay = b\}.$$

Три возможных способа оценки параметров линейной аппроксимации

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

1. Минимизация суммы модулей отклонений

$$\sum h_j \left| y_j \right| \to \min,\ y \in Y.$$

2. Метод наименьших квадратов

$$\sum h_j (y_j)^2 \to \min,\ y \in Y.$$

3. Минимизация максимального отклонения

$$\max_{j=1,...,n} h_j \left| y_j \right| \to \min,\ y \in Y.$$

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.
6. И др.

Некоторые причины введения и варьирования весовых коэффициентов

1.Экзогенные показатели информативности наблюдений при декомпозиции и прогнозировании временных рядов (было в докладе 30. 3. 23).

2.Для учета степени приближения к границам разных ограничений-неравенств в методе внутренних точек (было 30. 3. 23).

3. Для ускорения метода итеративной линеаризации.

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?

Поиск направления улучшения решения методом наименьших квадратов

Вектор s^kможно определить как результат решения задачи:

$$\sum\nolimits_1^n h_j^k \left(s_j\right)^2 \to min, \quad A^k s = -F\left(x^k\right).$$

Здесь

$$A^k = \nabla F\left(x^k\right)$$

матрица $m \times n$ с коэффициентами $a_{i,j}^k, i = 1, \dots, m, j = 1, \dots, n.$

Соотношения компонент вектора весовых коэффициентов h_j^kотражают степень нелинейности зависимостей $F_i, i = 1, \dots, m$ в точке x^k. Например, для дважды дифференцируемых функций

$$h_j^k = 1 + \sum\nolimits_{i=1}^m \sum\nolimits_{l=1}^n (\partial a_{i,l}^k / \partial x_j)^2, \qquad j = 1, \dots, n.$$

Весовые коэффициенты в методе наименьших квадратов не обязательно связывать только с точностью наблюдений.

Весовые коэффициенты в решении системы нелинейных уравнений методом линеаризации

Это пример использования весовых коэффициентов, не являющихся соизмерителями точности наблюдений.

Кроме аппроксимации на базе наименьших квадратов могут использоваться другие постановки в том числе на основе метода наименьших модулей, чебышевская аппроксимация с идейно аналогичными весами.

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 т. ч. более общих постановках (гельдеровские, несимметричные проекции, минимизация штрафных функций, парето – оптимизация). Двойственные задачи.

Используемые обозначения

S - линейное подпространство, параллельное Y,

$S^{\perp}$ - линейное подпространство, ортогональное S.

Для множества $Q \subset R^n$ выражения *co Q*, *cl Q* обозначают выпуклую оболочку Q, замыкание Q.

Если Q – выпуклое множество, то *ri Q* – множество относительно внутренних точек Q *(внутренних относительно минимального линейного многообразия, содержащего Q)*.

Для вектора $x \in R^n$ обозначим множества номеров его компонент с нулевыми, положительными, отрицательными и ненулевыми значениями :

$$J_0(x) = \{j : x_j = 0\},\ J_+(x) = \{j : x_j > 0\},$$

$$J_-(x) = \{j : x_j < 0\},\ J(x) = J_+(x) \cup J_-(x).$$

Первая конкретизация проблемы

Поиск «особых» 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.

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

Нерешенная проблема

- Как назвать векторы из B?
- Это векторы линейного многобразия с минимальным (не сужаемым) носителем,   с максимальным (не расширяемым) набором номеров нулевых компонент.
- Для полиэдра аналогичные вектора определяются сложнее.
- Предложения можно по мейлу: vizork&mail.ru
- Вариант поддержанный С.П.Шарым «Узловые решения»?
-

«Узловые» векторы полиэдра (множества решений систем линейных неравенств)

- Обладают тремя свойствами.

- 1. С несужаемым носителем.
- 2. С несужаемым набором активных ограничений неравенств.
- 3.Среди вышеуказанных только с Парето-минимальными абсолютными значениями компонент.

Уточнение теоремы 1

- Определим носитель линейного многообразия
- $$J(Y) = \{\cup J(y) | y \in Y\}.$$
- Пусть $n1$ количество номеров в $J(Y)$. То есть,
- $$n - n1$$
- - число компонент, имеющих только нулевые значения у всех векторов линейного многообразия.
-
- **Уточненная теорема 1.** *Количество элементов в множестве особых векторов* B *не более, чем* C_{n1}^{m}.

Условие «невырожденности»

Если число векторов в множестве особых (узловых) векторов совпадает с максимально возможным числом $C_{n1}^{m} = \frac{n1!}{m!(n1-m)!}$, то линейное многообразие Y «невырожденное».

Три причины возможной недостижимости этого числа:

1) особые векторы могут иметь более чем m нулевых компонент, несколько особых векторов «сливаются»;

2) два (или больше) множества векторов с нулевыми значениями разных компонент не пересекаются (параллельны) на Y;

3) два (или больше) множества векторов с нулевыми значениями компонент совпадают на Y.

Вторая конкретизация проблемы

Поиск Парето – оптимальных решений многокритериальной задачи:

$$\left|y_j\right| \to \min,\ j = 1,\ldots,n,\quad y \in Y.$$

Множество Парето – оптимальных решений

$$Q = \left\{q \in Y : \neg\exists y \in Y, \left|y_j\right| \le \left|q_j\right|,\ j = 1,\ldots,n, \sum\left|y_j\right| < \sum\left|q_j\right|\right\}.$$

Конструктивные критерии выявления векторов не находящихся и находящихся в Q

- ***Лемма 1***. *Из определения* Q *следует:*
- *вектор* $x \in Y$ ***не находится*** *в* Q, *если и только если существует*
- $s \in S, s \neq 0,$ *такой что*
- $J_{-}(s) \subseteq J_{+}(x), \quad J_{+}(s) \subseteq J_{-}(x)$.
-
- ***Лемма 2.*** *Из теоремы Гордона об альтернативных линейных неравенствах и леммы1 следует:*
- *вектор* $x \in Y$ ***находится*** *в* Q, *если и только если существует* $s \in S^{\perp}$,*такой что*
- $J_{+}(q) \subseteq J_{+}(s), \; J_{-}(q) \subseteq J_{-}(s)$.

Связь Парето – оптимальных и особых решений.

Теорема 2. $B \subseteq Q.$

Теорема 3. $Q \subseteq co\, B.$

Теорема 4 (следствие из теорем 2, 3). $co\, Q = co\, B.$

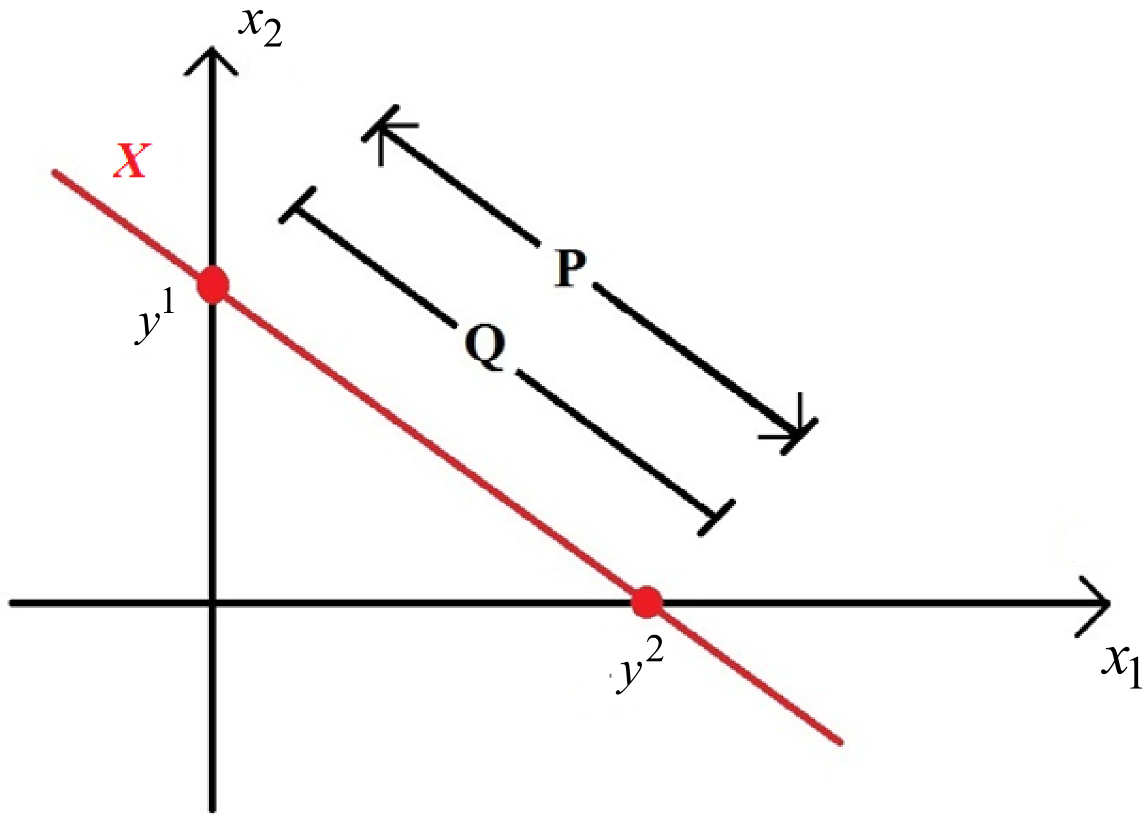

Рис 1. Одномерное линейное многообразие в R^2 *с двумя векторами с минимальным носителем*

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

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

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

Продолжение примера

В рассмотренном примере линейное многообразие - множество решения системы линейных уравнений

$$y_1 + y_2 = 1.$$

Особые векторы: $y^1 = (1,0)^T, \quad y^2 = (0,1)^T.$

Вектор $q^1 = (0.5, 0.5)^T$ находится в Q. Для $q^1, s^1 = (1,1)^T$ выполняются условия леммы 2, в том числе соотношение $s^1 \in S^T$.

Векторы $q^2 = (2, -1)^T$, $q^3 = (-1, 2)^T$ не находятся в Q для них при векторах $s^2 = (-1,1)^T, s^3 = (1, -1)^T$ из S выполняются условия леммы 1.

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,* $cl\, Q = Q$;
2) *связным;*
3) *ограниченным, объединением конечного числа политопов,* $Q = \{\cup\, Q^i \;|\, i = 1, \dots, l\}$.

Строение линейных многообразий

Пусть $q^i, i = 1, \dots, L$ максимальный набор векторов Y с не расширяемыми в Y носителем и различающихся наборами положительные и отрицательных компонент.

Лемма 3. $L \leq 2^{n1}$. *Равенство будет только в том случае, когда* Y *линейное подпространство.*

Лемма 4.$J\left(q^i\right) = J(Y),\ i = 1, \dots, L.$

Лемма 5. *Среди векторов* $q^i, i = 1, \dots, L$ *существует* l *векторов находящихся в* Q.

Лемма 6. *Число* $L - l$ *четное.*

Далее считаем, что первые l векторов q^{i}находятся в Q.

Строение линейных многообразий

Введем выпуклые полиэдры, порождаемые векторами q^i:

$$G^i = \{y \in Y:\ J_+(y) \subseteq J_+(q^i).\ J_-(y) \subseteq J_-(q^i)\},$$

$i = 1, \ldots L.$

Справедливы утверждения:

1. $Y = \bigcup_1^L G^i;$

2. $ri\ G^i = \{y \in Y: J_+(y) = J_+(q^i),\ J_-(y) = J_-(q^i)\},$
$i = 1, \ldots L;$

3. $B^i \neq \emptyset$, где $B^i = B \cap G^i, \quad i = 1, \ldots L;$

Строение линейных многообразий

4. *Первые* l *множеств ограниченные (политопы)*

$$G^i = Q^i = co\, B^i, i = 1, \dots, l;$$

5. *Остальные множества неограниченные,*

$$R^t = G^i = coB^i + C^i,\ t = 1, \dots, L - l, \quad i = t + l,$$

где $C^i = \left\{s \in S\colon\ J_+(s) \subseteq J_+\left(q^i\right).\ J_-(s) \subseteq J_-\left(q^i\right)\right\}$

-конус рецессивных направлений в множестве G^i*;*

6. *У каждого неограниченного множества есть его «антипод»: для любого* $i \in \{l + 1, \dots, L\}$ *существует* $j \in \{l + 1, \dots, L\}$*, для которых*

$$J_+\left(q^i\right) = J_-(q^j), \qquad J_-\left(q^i\right) = J_+\left(q^j\right).$$

Отсюда четность числа неограниченных множеств, множеств не входящих в Q.

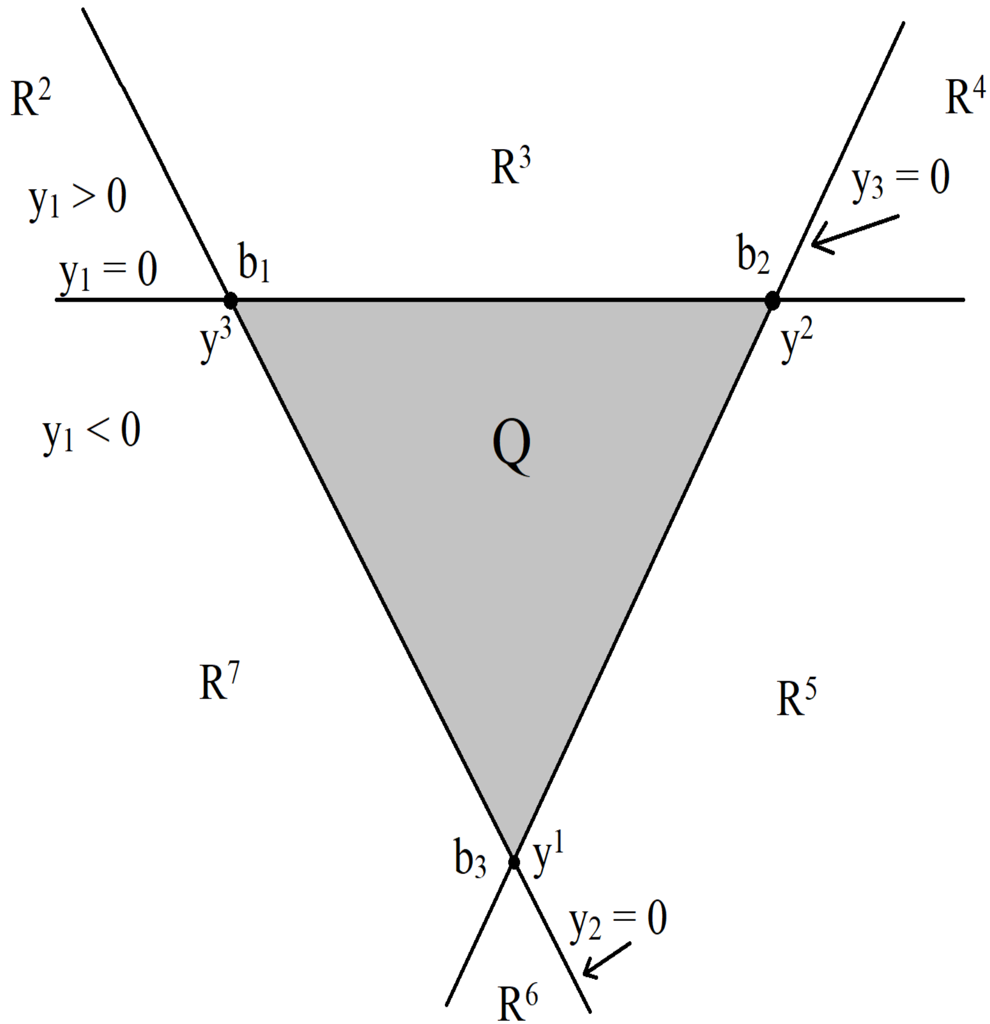

Рис. 3. Двумерное линейное многообразие в R^3 с тремя особыми векторами. Линейного многообразия на плоскости многообразия. Прямые линии – точки, где одна из компонент вектора многообразия равна нулю. Множество евклидовых проекций – относительная внутренность треугольника

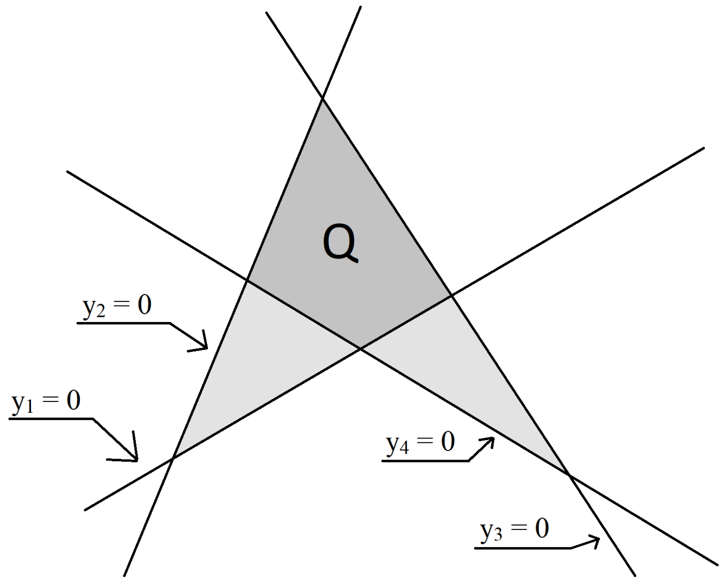

Рис. 3. Двумерное линейное многообразие в четырехмерном пространстве. Прямые линии состоят из точек с нулевым значением одной из компонент. Множество Q не выпуклое

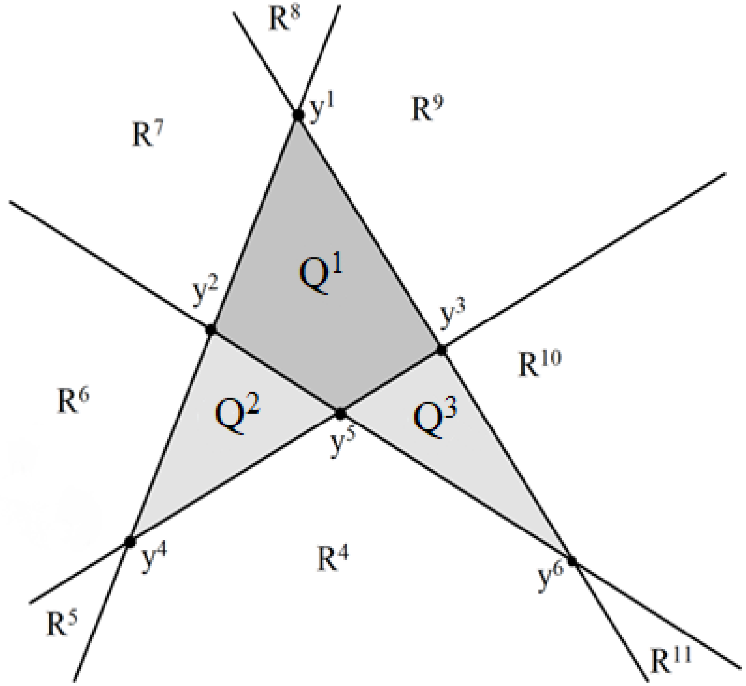

Рис. 4. Двумерное линейное многообразие в четырехмерном пространстве. Пересечения двух прямых соответствуют опорным векторам y^i, $i = 1, \ldots ,6$. Три политопа Q^1, $i = 1, 2, 3$, образуют Q.


Magic of Baikal © Alexe

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

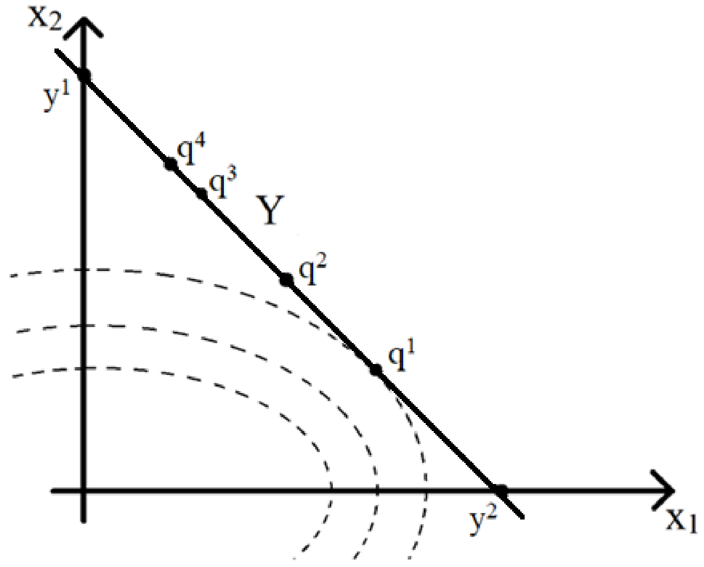

Рис. 2. Евклидовы проекции начала координат на линейное многообразие. Изменения проекций при увеличении соотношения первого ко второму весовых коэффициентов.

Преимущества метода наименьших квадратов

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

2. **Теорема 9 (анонсируется).** *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, метод наименьших квадратов – не хуже новых двух!

Некоторые приложения результатов

1.Множества P_2и Qблизки:

$$ri\,Q \subseteq P_2 \subseteq Q,$$

$$cl\,ri\,Q = clP_2 = Q.$$

Этим объясняется большая роль изучения свойств множества Q (векторов линейного многообразия с Парето-минимальными абсолютными значениями компонент) в изучении свойств множества евклидовых проекций P_2.

Любое решение из Qможно получить или точно или с любой требуемой степенью точности в виде евклидовой проекции начала координат на линейное многообразие за счет выбора весовых коэффициентов.

Некоторые приложения результатов

2. Евклидовы проекции за счет выбора весовых коэффициентов могут заменить задачи минимизации на линейном многообразии штрафных функций из широкого класса, в том числе любую гёльдеровскую проекцию:

$$P_2 = P = PF.$$

3.Множество евклидовых проекций ограниченно. Возможные вариации значений Вкомпонент векторов евклидовых проекций можно определить на основе расчета вариаций компонент конечного числа особых (узловых) векторов B:

$$co\ cl\ P_2 = \ co\ B..$$


gic of Baikal © Alexey Trofimov

Не рассмотрены (оставлены на будущее)

1. Октаэдральные и чебышевские проекции.
2. Несимметричные нормы и проекции.
3. Двойственные задачи проектирования.
4. Применения в методе внутренних точек.
5. Применение в проблеме выделения составляющих временных рядов.
6. Все для ближайших к началу координат точек полиэдров.

Спаси бо за!