

Прогнозирование траекторий динамических систем на длительный период методами машинного обучения

Царев С.П.¹

¹Сибирский федеральный университет,
Красноярск

Конференция «Геометрия, топология и математическая физика»
памяти Сергея Петровича Новикова
2025-06-05

План:

- Мотивация & формализация:
 - ▶ требуется высокая точность (9-10 десятичных знаков);
 - ▶ богатство геометрических методов и методов машинного обучения

План:

- Мотивация & формализация:
 - ▶ требуется высокая точность (9-10 десятичных знаков);
 - ▶ богатство геометрических методов и методов машинного обучения
- В данной лекции:
 - ▶ методы редукции размерности
 - ▶ sparse recovery & compressive sensing

План:

- Мотивация & формализация:
 - ▶ требуется высокая точность (9-10 десятичных знаков);
 - ▶ богатство геометрических методов и методов машинного обучения
- В данной лекции:
 - ▶ методы редукции размерности
 - ▶ sparse recovery & compressive sensing
- <https://forum.mathmeth.ru>
(Лекции по матем. методам в приложении к ГНСС, 2023 г.)

План:

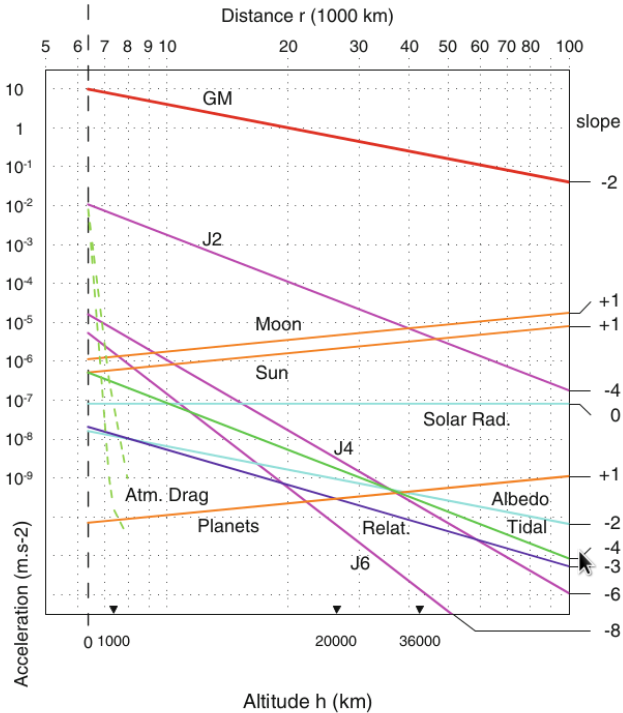
- Мотивация & формализация:
 - ▶ требуется высокая точность (9-10 десятичных знаков);
 - ▶ богатство геометрических методов и методов машинного обучения
- В данной лекции:
 - ▶ методы редукции размерности
 - ▶ sparse recovery & compressive sensing
- <https://forum.mathmeth.ru>
(Лекции по матем. методам в приложении к ГНСС, 2023 г.)
- Вне рамок лекции:
 - ▶ Topological data analysis (TDA)
 - ▶ теория функций *очень* многих переменных и пространств очень большой размерности

Действующие на спутники силы и их эффект

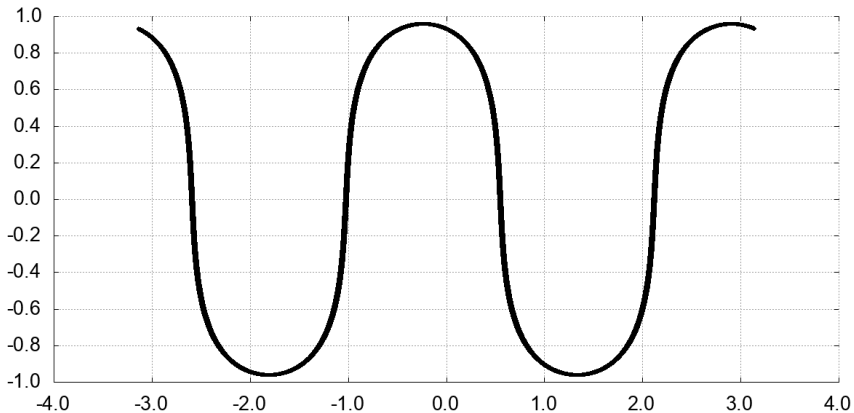
Т а б л и ц а 10.1

Влияние различных факторов на движение НИСЗ

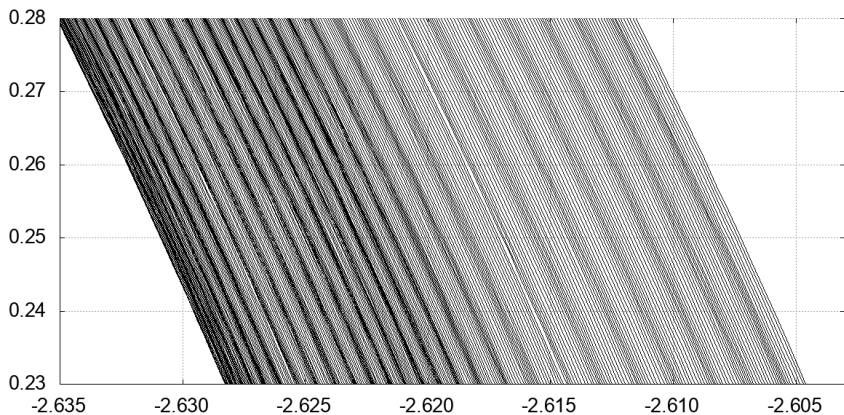
Возмущающие факторы	Максимальное возмущающее ускорение, м/с^2	Максимальное возмущение за 1^h , м
Центральное поле Земли	$5,65 \cdot 10^{-1}$	—
Вторая зональная гармоника	$5,3 \cdot 10^{-6}$	300
Гравитация Луны	$5,5 \cdot 10^{-6}$	40
Гравитация Солнца	$3 \cdot 10^{-6}$	20
Четвертая зональная гармоника	10^{-7}	0,6
Солнечная радиация	10^{-7}	0,6
Гравитационные аномалии	10^{-8}	0,06
Другие силы	10^{-8}	0,06



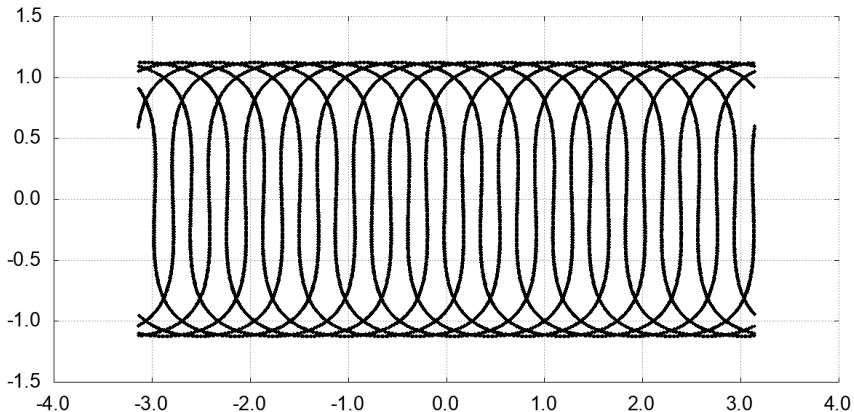
Проекция орбиты G01 (GPS) на поверхность Земли (2013)



Проекция орбиты G01 (GPS) (2013) в увеличенном виде



Проекция орбиты R01 (GLONASS) на поверхность Земли (2013)



Дискретизация траекторий и редукция размерности

Дискретизация траекторий и редукция размерности

Сколько точек надо задать на суточном участке орбиты, чтобы можно было легко восстановить все промежуточные положения?

Дискретизация траекторий и редукция размерности

Сколько точек надо задать на суточном участке орбиты, чтобы можно было легко восстановить все промежуточные положения?

Достаточно задать 96 точек (положения спутника через каждые 15 мин) — промежуточные восстанавливаются с точностью 1 мм полиномиальной интерполяцией по 12 соседним точкам.

Дискретизация траекторий и редукция размерности

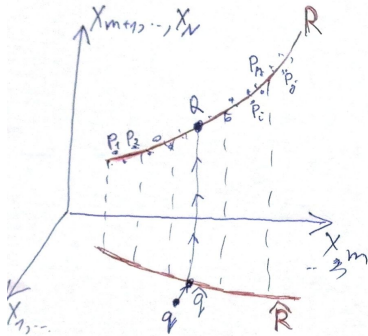
Сколько точек надо задать на суточном участке орбиты, чтобы можно было легко восстановить все промежуточные положения?

Достаточно задать 96 точек (положения спутника через каждые 15 мин) — промежуточные восстанавливаются с точностью 1 мм полиномиальной интерполяцией по 12 соседним точкам.

Простейшая идея редукции размерности:

- задать суточную траекторию как 288-мерный ($288 = 96 \cdot 3$) вектор, соотв. недельный отрезок дает 2016-мерный вектор;
- рассмотреть множество таких «недельных векторов» за длительный период;
- Найти подмногообразие в \mathbb{R}^{2016} меньшей размерности, содержащее полученный набор точек (или очень близкое к нему);
- использовать для прогноза орбит на следующие несколько суток (недель, месяцев).

Прогноз траектории методом уменьшения размерности



X_1, \dots, X_m - координаты точек кагалавного (заданного) участка орбиты

X_{m+1}, \dots, X_N - координаты точек прогнозируемого участка орбиты

P_i - (с координатами X_1, \dots, X_m) - точки "обучающего" набора орбит (кагальн + прогноз из SPЗ данных)

R - редуцированное многообразие размерности $K < m$, приближающее точки P_i

\hat{R} - проекция R на подпр-во (X_1, \dots, X_m)

q - (коорд. q_1, \dots, q_m) - кагальный участок орбиты, которую хотим предсказать (прогнозировать)

\hat{q} - ее проекция на \hat{R}

Q - обратная проекция \hat{q} на R , координаты Q_{m+1}, \dots, Q_N дают прогноз для участка q

Простейшая редукция: метод главных компонент (линейный PCA), суточные траектории

размерность	max, км	СКО, км
10	4.402e-03	4.781e-04
15	5.351e-04	6.344e-05
20	5.712e-05	1.136e-05
25	9.410e-06	2.057e-06
30	4.219e-06	5.871e-07

Простейшая редукция: метод главных компонент (линейный PCA), недельные траектории

размерность	max, км	СКО, км
10	4.872e-02	6.178e-03
15	1.122e-02	1.133e-03
20	1.103-03	1.715e-04
25	2.752-04	4.079e-05
30	7.680-05	1.293e-05

Слабонелинейные подмногообразия,
задание функций (очень) многих переменных
и методы `sparse recovery`

Попробуем задать нелинейное подмногообразие, например полиномами
многих переменных степени 2.

Слабонелинейные подмногобразия, задание функций (очень) многих переменных и методы `sparse recovery`

Попробуем задать нелинейное подмногобразие, например полиномами многих переменных степени 2.

Какое количество коэффициентов при этом необходимо найти — методом машинного обучения?

Слабонелинейные подмногобразия, задание функций (очень) многих переменных и методы *sparse recovery*

Попробуем задать нелинейное подмногобразие, например полиномами многих переменных степени 2.

Какое количество коэффициентов при этом необходимо найти — методом машинного обучения?

У нас НЕТ такого количества данных! (Хотя их немало)

Слабонелинейные подмногобразия, задание функций (очень) многих переменных и методы `sparse recovery`

Попробуем задать нелинейное подмногобразие, например полиномами многих переменных степени 2.

Какое количество коэффициентов при этом необходимо найти — методом машинного обучения?

У нас НЕТ такого количества данных! (Хотя их немало)

Выход: задавать лишь небольшое число ненулевых слагаемых в полиноме многих переменных.

Слабонелинейные подмногобразия, задание функций (очень) многих переменных и методы `sparse recovery`

Попробуем задать нелинейное подмногобразие, например полиномами многих переменных степени 2.

Какое количество коэффициентов при этом необходимо найти — методом машинного обучения?

У нас НЕТ такого количества данных! (Хотя их немало)

Выход: задавать лишь небольшое число ненулевых слагаемых в полиноме многих переменных.

Какие мономы отобрать для наилучшего приближения?

Слабонелинейные подмногочисления, задание функций (очень) многих переменных и методы `sparse recovery`

Попробуем задать нелинейное подмногочисление, например полиномами многих переменных степени 2.

Какое количество коэффициентов при этом необходимо найти — методом машинного обучения?

У нас НЕТ такого количества данных! (Хотя их немало)

Выход: задавать лишь небольшое число ненулевых слагаемых в полиноме многих переменных.

Какие мономы отобрать для наилучшего приближения?

Один из методов: `sparse recovery`.

Задача сводится к решению недоопределенной линейной системы на большое число неизвестных коэффициентов при мономах, так, чтобы лишь небольшое число неизвестных коэффициентов остались ненулевыми.

Теоретическая постановка задачи (ver. 1)

Простейшая линейная система с числом уравнений, *меньшим*, чем число неизвестных, при условии, что лишь *небольшое* число неизвестных — ненулевые (но неизвестно, какие!):

$$\hat{A} \cdot x = b \tag{1}$$

Теоретическая постановка задачи (ver. 1)

Простейшая линейная система с числом уравнений, *меньшим*, чем число неизвестных, при условии, что лишь *небольшое* число неизвестных — ненулевые (но неизвестно, какие!):

$$\hat{A} \cdot x = b \quad (1)$$

В практических приложениях, система должна решаться приближенно:

$$\|\hat{A} \cdot x - b\|_2 < \varepsilon \quad (2)$$

Теоретическая постановка задачи (ver. 1)

Простейшая линейная система с числом уравнений, *меньшим*, чем число неизвестных, при условии, что лишь *небольшое* число неизвестных — ненулевые (но неизвестно, какие!):

$$\hat{A} \cdot x = b \quad (1)$$

В практических приложениях, система должна решаться приближенно:

$$\|\hat{A} \cdot x - b\|_2 < \varepsilon \quad (2)$$

Метод LASSO (1990-е) предлагает минимизировать L_1 -норму искомого вектора x :

$$\begin{cases} x = \arg \min_x \sum_{i=1}^N |x_i|, & \text{при условии} \\ \|\hat{A} \cdot x - b\|_2 < \varepsilon \end{cases} \quad (3)$$

Теоретическая постановка задачи (ver. 1)

Простейшая линейная система с числом уравнений, *меньшим*, чем число неизвестных, при условии, что лишь *небольшое* число неизвестных — ненулевые (но неизвестно, какие!):

$$\hat{A} \cdot x = b \quad (1)$$

В практических приложениях, система должна решаться приближенно:

$$\|\hat{A} \cdot x - b\|_2 < \varepsilon \quad (2)$$

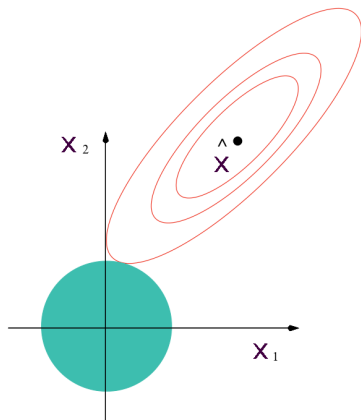
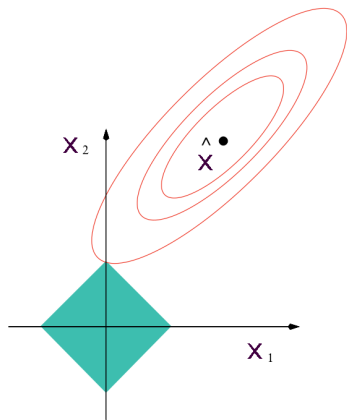
Метод LASSO (1990-е) предлагает минимизировать L_1 -норму искомого вектора x :

$$\begin{cases} x = \arg \min_x \sum_{i=1}^N |x_i|, & \text{при условии} \\ \|\hat{A} \cdot x - b\|_2 < \varepsilon \end{cases} \quad (3)$$

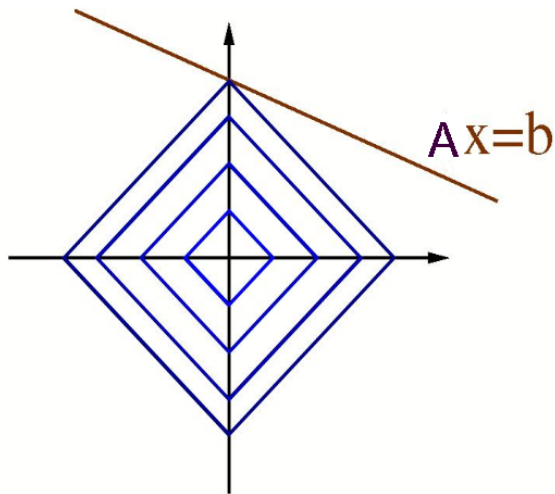
либо объединить оба условия:

$$x = \arg \min_x \left\{ \sum_{i=1}^N |x_i| + \lambda \cdot \|\hat{A} \cdot x - b\|_2 \right\} \quad (4)$$

Геометрическая картина (2×2)
(эффект разреженности для L_1 -нормы)



Геометрическая картина (1×2)
(эффект разреженности для L_1 -нормы)



Определения

Задача (P0) *NP-hard* !!

$$\begin{cases} x = \arg \min_x \{ \|x\|_0 = \#(x_k \neq 0) \}, & \text{при условии} \\ \hat{A} \cdot x = b \end{cases}$$

Задача (P0 ϵ) *NP-hard* !!

$$\begin{cases} x = \arg \min_x \{ \|x\|_0 = \#(x_k \neq 0) \}, & \text{при условии} \\ \|\hat{A} \cdot x - b\|_2 < \epsilon \end{cases}$$

Задача (P1)

$$\begin{cases} x = \arg \min_x \{ \|x\|_1 = \sum_{i=1}^N |x_i| \}, & \text{при условии} \\ \hat{A} \cdot x = b \end{cases}$$

Задача (P1 ϵ)

$$\begin{cases} x = \arg \min_x \{ \|x\|_1 = \sum_{i=1}^N |x_i| \}, & \text{при условии} \\ \|\hat{A} \cdot x - b\|_2 < \epsilon \end{cases}$$

Теоретические результаты: Sparse Recovery and RIP

Restricted Isometry Property of Order k [Candès, Romberg, Tao (2006)]:

Let δ_k be the smallest number such that

$$(1 - \delta_k) \|x\|_2^2 \leq \|\hat{A}x\|_2^2 \leq (1 + \delta_k) \|x\|_2^2$$

for all k -sparse vectors $x \in \mathbb{R}^n$ where $\hat{A} = [a_1 \dots a_n] \in \mathbb{R}^{m \times n}$.

Теорема (E.J.Candès (2008))

If $\delta_{2k} < \sqrt{2} - 1$, then for all k -sparse vectors x such that $\hat{A}x = b$, the solution of (P1) is equal to the solution of (P0).

Упражнения

L_2 vs L_1 в статистике: «средние» значения ($dim = 1$)

Пусть имеется выборка x_1, \dots, x_n .

$$x_{average2} = \arg \min_{x^*} \left\{ \sum_{i=1}^n (x_i - x^*)^2 \right\}$$

$x_{average2}$ — среднее арифметическое, *неустойчива к «выбросам»*

Упражнения

L_2 vs L_1 в статистике: «средние» значения ($dim = 1$)

Пусть имеется выборка x_1, \dots, x_n .

$$x_{average2} = \arg \min_{x^*} \left\{ \sum_{i=1}^n (x_i - x^*)^2 \right\}$$

$x_{average2}$ — среднее арифметическое, *неустойчива к «выбросам»*

$$x_{average1} = \arg \min_{x^*} \left\{ \sum_{i=1}^n |x_i - x^*| \right\}$$

$x_{average1}$ — медиана (почти...), *устойчива к «выбросам»*

Упражнения

L_2 vs L_1 в статистике: «средние» значения ($dim = 1$)

Пусть имеется выборка x_1, \dots, x_n .

$$x_{average2} = \arg \min_{x^*} \left\{ \sum_{i=1}^n (x_i - x^*)^2 \right\}$$

$x_{average2}$ — среднее арифметическое, *неустойчива к «выбросам»*

$$x_{average1} = \arg \min_{x^*} \left\{ \sum_{i=1}^n |x_i - x^*| \right\}$$

$x_{average1}$ — медиана (почти...), *устойчива к «выбросам»*

$x_{average2}$: линейная оценка, $x_{average1}$: *нелинейная*.

Пример приложения Compressive sensing (томография)

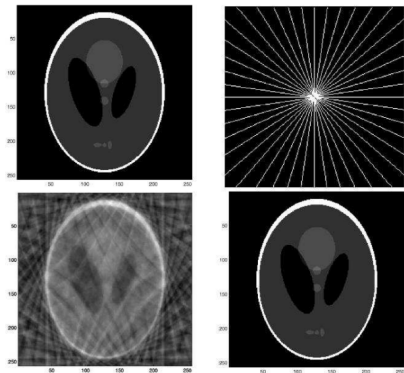


Figure 1: When Fourier coefficients of a testbed medical image known as the Logan-Shepp phantom (top left) are sampled along 22 radial lines in the frequency domain (top right), a naive, “minimal energy” reconstruction setting unobserved Fourier coefficients to 0 is marred by artifacts (bottom left). ℓ_1 -reconstruction (bottom right) is exact.

Пример приложения Compressive sensing (томография)

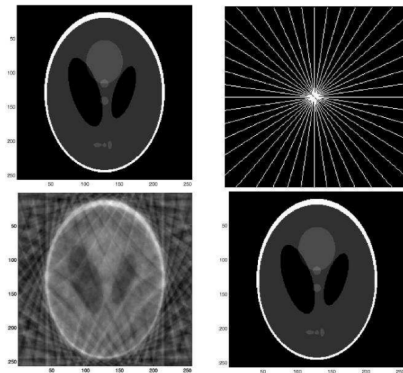
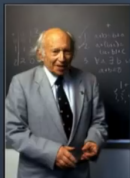


Figure 1: When Fourier coefficients of a testbed medical image known as the Logan-Shepp phantom (top left) are sampled along 22 radial lines in the frequency domain (top right), a naive, “minimal energy” reconstruction setting unobserved Fourier coefficients to 0 is marred by artifacts (bottom left). ℓ_1 -reconstruction (bottom right) is exact.

Размерность вектора данных (количество пикселей): $> 10^5$!!

Лекция D.Донахо (премия Гаусса 2018)

Geometric Functional Analysis



IM Gel'fand



Boris Kashin



Holger Rauhut

$$d^n(B, X) = \inf_{\dim(V)=n} \sup_{x \in B \cap V} \|x\|_X$$

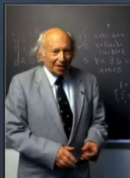
$$d^n(B_{\ell_2^N}, \ell_2^N) \asymp \sqrt{\frac{\log(n)}{n}}$$

$$d^n(B_{\ell_2^N}, \ell_2^N) \asymp \left(\frac{\log(N)}{n} \right)^{1/2-1/p}$$

$$0 < p < 2$$

Лекция D.Донахо (премия Гаусса 2018)

Geometric Functional Analysis



IM Gel'fand



Boris Kashin



Holger Rauhut

$$d^n(B, X) = \inf_{\dim(V)=n} \sup_{x \in B \cap V} \|x\|_X$$

$$d^n(B_{\ell_2^N}, \ell_2^N) \asymp \sqrt{\frac{\log(n)}{n}}$$

$$d^n(B_{\ell_2^N}, \ell_2^N) \asymp \left(\frac{\log(N)}{n} \right)^{1/2-1/p}$$

$$0 < p < 2$$

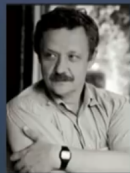
Achtung: $0 < p < 2$!!

Лекция D.Допоно (премия Гаусса 2018)

Optimization



Arkadi Nemirovsky



Yuri Nesterov



Daubechies



De Frise



De Mol

Теоретическая постановка задачи (вер. 2)

Проблема построения CS-матрицы

Для неизвестного многомерного *разреженного* вектора x найти хорошую «матрицу измерений» $\hat{A} \in \mathbb{R}^{m \times n}$ с *малым* m такую, что по имеющимся «измерениям» $\hat{A}x = b$, решая разреженную задачу можно найти x .

Теоретическая постановка задачи (вер. 2)

Проблема построения CS-матрицы

Для неизвестного многомерного разреженного вектора x найти хорошую «матрицу измерений» $\hat{A} \in \mathbb{R}^{m \times n}$ с малым m такую, что по имеющимся «измерениям» $\hat{A}x = b$, решая разреженную задачу можно найти x .

Известные конструкции: случайные матрицы (!!!) дают $k \leq Cm/\log(n/m)$.

Теорема (Candès-Romberg-Тao, 2004)

Пусть $\xi_1, \dots, \xi_m \in \{1, \dots, n\}$ выбраны случайно. Тогда с высокой вероятностью каждый k -разреженный сигнал $f : \{1, \dots, n\} \rightarrow \mathbb{C}$ может быть восстановлен из $\hat{f}(\xi_1), \dots, \hat{f}(\xi_m)$, если $m > Ck \log n$ для некоторой абсолютной константы C .

Численные эксперименты показывают, что на практике большинство k -разреженных сигналов фактически восстанавливаются примерно при $m \geq 4k$.

Теоретическая постановка задачи (вер. 2)

Пример «хорошей» случайной CS-матрицы:

Лемма (Лемма Джонсона-Линденштрауса (вер. 1))

Даны $\varepsilon \in (0, 1)$, множество X из m точек в \mathbb{R}^N и целое число n , такое что $n > C \ln m / \varepsilon^2$, тогда существует линейное отображение (случайная ортогональная проекция) $\Phi : \mathbb{R}^N \rightarrow \mathbb{R}^n$ такое, что

$$(1 - \varepsilon) \|u - v\|_2 \leq \|\Phi(u) - \Phi(v)\|_2 \leq (1 + \varepsilon) \|u - v\|_2, \quad \forall u, v \in X.$$

Лемма (Лемма Джонсона-Линденштрауса (вер. 2))

Для любого целого числа $d > 0$ и любых $0 < \varepsilon, \delta < 1/2$ существует распределение вероятностей на $k \times d$ действительных матрицах для $k = \Theta(\varepsilon^{-2} \log(1/\delta))$ такое, что для любого $x \in \mathbb{R}^d$ с $\|x\|_2^2 = 1$,

$$\text{Prob}_S[|\|Sx\|_2^2 - 1| > \varepsilon] < \delta.$$

Насколько хороши случайные матрицы?

Мысленный эксперимент: при заданном m попробуем максимально увеличить n , выбирая случайные векторы. Два таких столбца практически ортогональны: корреляция (= скал. произведение) мала. Редкие случайно коррелирующие столбцы отбрасываем.

Насколько хороши случайные матрицы?

Мысленный эксперимент: при заданном m попробуем максимально увеличить n , выбирая случайные векторы. Два таких столбца практически ортогональны: корреляция (= скал. произведение) мала. Редкие случайно коррелирующие столбцы отбрасываем.

С точки зрения (наивной) статистики, такой процесс наращивания n бесконечен ... Что противоречит (наивной) геометрии: количество почти ортогональных векторов в \mathbb{R}^m конечно.

Насколько хороши случайные матрицы?

Мысленный эксперимент: при заданном m попробуем максимально увеличить n , выбирая случайные векторы. Два таких столбца практически ортогональны: корреляция (= скал. произведение) мала. Редкие случайно коррелирующие столбцы отбрасываем.

С точки зрения (наивной) статистики, такой процесс наращивания n бесконечен ... Что противоречит (наивной) геометрии: количество почти ортогональных векторов в \mathbb{R}^m конечно.

Статистика плохо работает при *очень больших* выборках?

Или есть пояснение причины этого парадокса?

Насколько хороши случайные матрицы?

Мысленный эксперимент: при заданном m попробуем максимально увеличить n , выбирая случайные векторы. Два таких столбца практически ортогональны: корреляция (= скал. произведение) мала. Редкие случайно коррелирующие столбцы отбрасываем.

С точки зрения (наивной) статистики, такой процесс наращивания n бесконечен ... Что противоречит (наивной) геометрии: количество почти ортогональных векторов в \mathbb{R}^m конечно.

Статистика плохо работает при *очень больших* выборках?

Или есть пояснение причины этого парадокса?

А насколько велико количество почти ортогональных векторов в \mathbb{R}^m ?

Насколько хороши случайные матрицы?

Мысленный эксперимент: при заданном m попробуем максимально увеличить n , выбирая случайные векторы. Два таких столбца практически ортогональны: корреляция (= скал. произведение) мала. Редкие случайно коррелирующие столбцы отбрасываем.

С точки зрения (наивной) статистики, такой процесс наращивания n бесконечен ... Что противоречит (наивной) геометрии: количество почти ортогональных векторов в \mathbb{R}^m конечно.

Статистика плохо работает при *очень больших* выборках?

Или есть пояснение причины этого парадокса?

А насколько велико количество почти ортогональных векторов в \mathbb{R}^m ?

Парадоксальный недавний результат:

Количество векторов единичной длины в \mathbb{R}^m , скал. произведения которых по модулю не более заданного малого δ , экспоненциально растет при увеличении m :

Kainen, Paul C., and Věra Kůrková.

Quasiorthogonal dimension of Euclidean spaces. Applied math letters, 6(3) (1993), p. 7–10.

Обработка сигналов и изображений, сверхразрешение и т. д. (версия 3, 4, ..., N)

Актуальная проблема: *Детерминированное* построение CS-матриц: поперечники Гельфанда, ...

Обработка сигналов и изображений, сверхразрешение и т. д. (версия 3, 4, ..., N)

Актуальная проблема: *Детерминированное* построение CS-матриц: поперечники Гельфанда, ...

«Преодоление предела Найквиста-Котельникова», сверхразрешение и повышение резкости изображений, ...

Обработка сигналов и изображений, сверхразрешение и т. д. (версия 3, 4, ..., N)

Актуальная проблема: *Детерминированное* построение CS-матриц: поперечники Гельфанда, ...

«Преодоление предела Найквиста-Котельникова», сверхразрешение и повышение резкости изображений, ...

Пример: the twelve coins puzzle:

имея двенадцать монет, одна из которых фальшивая (и, следовательно, тяжелее или легче других), можно определить фальшивую монету всего за три взвешивания, взвешивая монеты в надлежащим образом выбранных партиях.

Ключевым моментом является то, что поддельные монеты — разрезаны (в ряду настоящих)!

Обработка сигналов и изображений, сверхразрешение и т. д. (версия 3, 4, ..., N)

Актуальная проблема: *Детерминированное* построение CS-матриц: поперечники Гельфанда, ...

«Преодоление предела Найквиста-Котельникова», сверхразрешение и повышение резкости изображений, ...

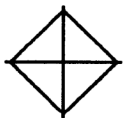
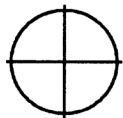
Пример: the twelve coins puzzle:

имея двенадцать монет, одна из которых фальшивая (и, следовательно, тяжелее или легче других), можно определить фальшивую монету всего за три взвешивания, взвешивая монеты в надлежащим образом выбранных партиях.

Ключевым моментом является то, что поддельные монеты — разрежены (в ряду настоящих)!

Как найти «хорошую матрицу взвешиваний» ($a_{ij} \in \{0, 1\}$) с небольшим m и большим n . (ОТК на массовом производстве)?

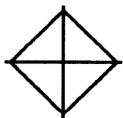
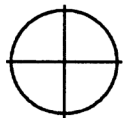
Невыпуклые задачи (ver. $N+1, N+2, \dots$)



Задача (P_ρ^ε) , $0 < \rho < 1$

$$\begin{cases} x = \arg \min_x \{ \|x\|_\rho = \sum_{i=1}^N |x_i|^\rho \}, & \text{restricted by} \\ \|\hat{A} \cdot x - b\|_2 < \varepsilon \end{cases}$$

Невыпуклые задачи (ver. $N+1, N+2, \dots$)



Задача ($\mathbf{P}_\rho^\varepsilon$), $0 < \rho < 1$







$$\begin{cases} x = \arg \min_x \{ \|x\|_\rho = \sum_{i=1}^N |x_i|^\rho \}, & \text{restricted by} \\ \|\hat{A} \cdot x - b\|_2 < \varepsilon \end{cases}$$

Задача ($\mathbf{TV}_\rho^\varepsilon$), $0 < \rho \leq 1$ для *total variation* (полной вариации)

$$\begin{cases} x = \arg \min_x \{ TV_\rho(x) = \sum_{i=1}^{N-1} |x(t_{i+1}) - x(t_i)|^\rho \}, & \text{restricted by} \\ \|\hat{A} \cdot x - b\|_2 < \varepsilon \end{cases}$$

Эта задача очень популярна при сегментации изображений, поиске и сглаживании резких границ и т. д.

References

-  Rudin LI, Osher S, Fatemi E. *Nonlinear total variation based noise removal algorithms* // Physica D: nonlinear phenomena. 1992 Nov 1;60(1-4):259-68.
-  Hastie, T., Tibshirani, R., Wainwright, M. *Statistical learning with sparsity: The Lasso and generalizations*. (2015).
-  Donoho D.L. *Compressed sensing* // IEEE Trans Inform Theory, v. 52, 1289–1306. 2006.
-  Candes E.J., Romberg J.K., Tao T. *Stable signal recovery from incomplete and inaccurate measurements* // Comm Pure Appl Math, 2006. v. 59. No. 8. 1207–1223.
-  Candes EJ, Fernandez-Granda C. *Super-resolution from noisy data* // J Fourier Anal Appl. 2013, 1229–54
-  А.С. Пустошилов, С.П. Царев *Обнаружение разрывов в фазовых измерениях одночастотных навигационных приемников при различной нестабильности опорных генераторов* // Ural Radio Engineering Journal. 2021. No 5(2). С. 144–161.