Sparse recovery и Compressive sensing в теории и на практике

Царев С.П.[1]

[1]Красноярский Математический Центр,
Сибирский федеральный университет,
Красноярск

Math Models and Integration Methods Seminar
https://mmandim.blogspot.com/
Красноярск, 2024-10-17

План:

- Мотивация & формализация:

План:

- Мотивация & формализация:
  - геометрия конечномерных пространств с (не совсем) нормой;

План:

- Мотивация & формализация:
    - геометрия конечномерных пространств с (не совсем) нормой;
    - внимание со стороны теоретической математики:
      работы Т.Тао, пленарные доклады ММК, премия Гаусса 2018, ...

План:

- Мотивация & формализация:
  - геометрия конечномерных пространств с (не совсем) нормой;
  - внимание со стороны теоретической математики:
    работы Т.Тао, пленарные доклады ММК, премия Гаусса 2018, ...
  - применения в обработке сигналов (томография, удаление шума),

План:

- Мотивация & формализация:
  - геометрия конечномерных пространств с (не совсем) нормой;
  - внимание со стороны теоретической математики:
    работы Т.Тао, пленарные доклады ММК, премия Гаусса 2018, ...
  - применения в обработке сигналов (томография, удаление шума),
- Когда лучше остановиться (теоретически) решая задачу?

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!):

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

В практических приложениях, система должна решаться приближенно:

$$\|\hat{A} \cdot x - b\|_2 < \varepsilon \tag{2}$$

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!):

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

В практических приложениях, система должна решаться приближенно:

$$\|\hat{A} \cdot x - b\|_2 < \varepsilon \tag{2}$$

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

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

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!):

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

В практических приложениях, система должна решаться приближенно:

$$\|\hat{A} \cdot x - b\|_2 < \varepsilon \tag{2}$$

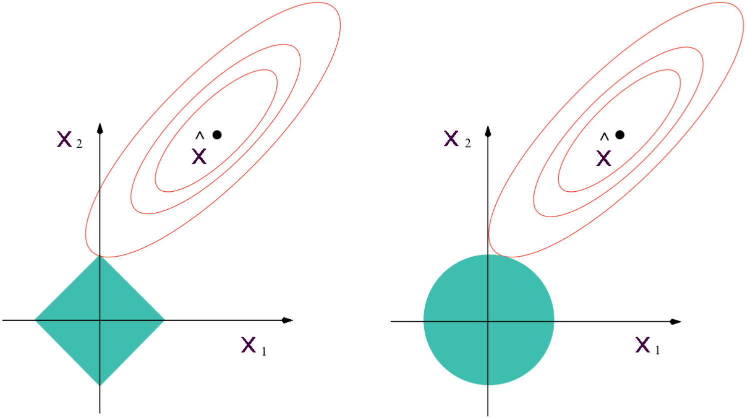
Метод LASSO (1990-е) предлагает минимизировать L_1-норму искомого вектора x:

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

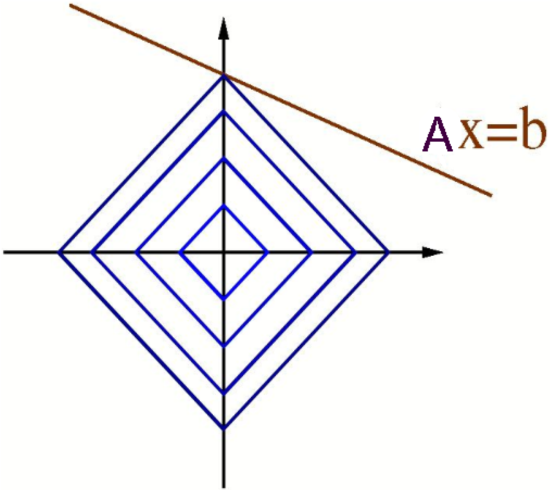
либо объединить оба условия:

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

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 (P0ε) *NP-hard* !!

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

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 (P1ε)

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

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

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

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).

Теоретические результаты: Approximate Recovery and RIP

Теорема (E.J.Candès (2008))

Suppose that the matrix $\hat{A}$ *is given and* $b = \hat{A}x_0 + e$ *where* $\|e\|_2 \leq \varepsilon$.
If $\delta_{2k} < \sqrt{2} - 1$, *then*

$$\|x^* - x_0\|_2 \leq C_0 k^{-1/2} \sigma_k(x_0)_1 + C_1 \varepsilon,$$

where x^* *is the solution of (P1*ε*) and*

$$\sigma_k(x_0)_1 = min\|x_0 - z\|_1$$

for all k*-sparse* z.

Упражнения

Почему в теоремах выбрано условие на разреженность $2k$:

Теорема

Предположим, что любые $2k$ столбцов $m \times n$ матрицы $\hat{A}$ линейно независимы. Тогда любой k-разреженный сигнал $x \in \mathbb{R}^n$ может быть однозначно восстановлен из $\hat{A}x$.

Упражнения

Почему в теоремах выбрано условие на разреженность $2k$:

Теорема

Предположим, что любые $2k$ столбцов $m \times n$ матрицы $\hat{A}$ линейно независимы. Тогда любой k-разреженный сигнал $x \in \mathbb{R}^n$ может быть однозначно восстановлен из $\hat{A}x$.

Следует отметить, что при решении задачи методом L_1-регуляризации будет (почти) точно найден лишь поднабор ненулевых переменных.

Упражнения

Почему в теоремах выбрано условие на разреженность $2k$:

Теорема

Предположим, что любые $2k$ столбцов $m \times n$ матрицы $\hat{A}$ линейно независимы. Тогда любой k-разреженный сигнал $x \in \mathbb{R}^n$ может быть однозначно восстановлен из $\hat{A}x$.

Следует отметить, что при решении задачи методом L_1-регуляризации будет (почти) точно найден лишь поднабор ненулевых переменных.

Их значения при зашумлении измерений будут, вообще говоря, несколько меньше «оптимальных» значений (эффект shrinkage, отраженный в названии метода LASSO = least absolute shrinkage and selection operator). Однако, поскольку задача отбора ненулевых переменных (selection) решается достаточно точно, можно далее использовать например, метод наименьших квадратов.

Поймите из геометрической картины, почему.

Упражнения

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

Пусть имеется выборка $x_1, \ldots x_n$.

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

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

Упражнения

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

Пусть имеется выборка $x_1, \ldots x_n$.

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

$x_{average2}$ — 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)

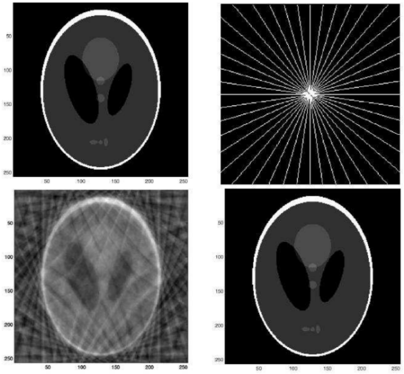


Figure 1: When Fourier coefficients of a testbed medical image known as the Logan–Shepp phantom (top left) are sampled along 22 radial lines in the frequency domain (top right), a naive, "minimal energy" reconstruction setting unobserved Fourier coefficients to 0 is marred by artifacts (bottom left).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

Пример приложения Compressive sensing (томография)

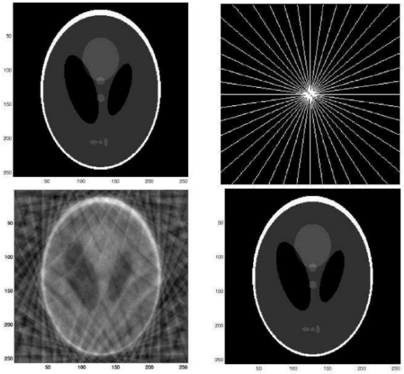
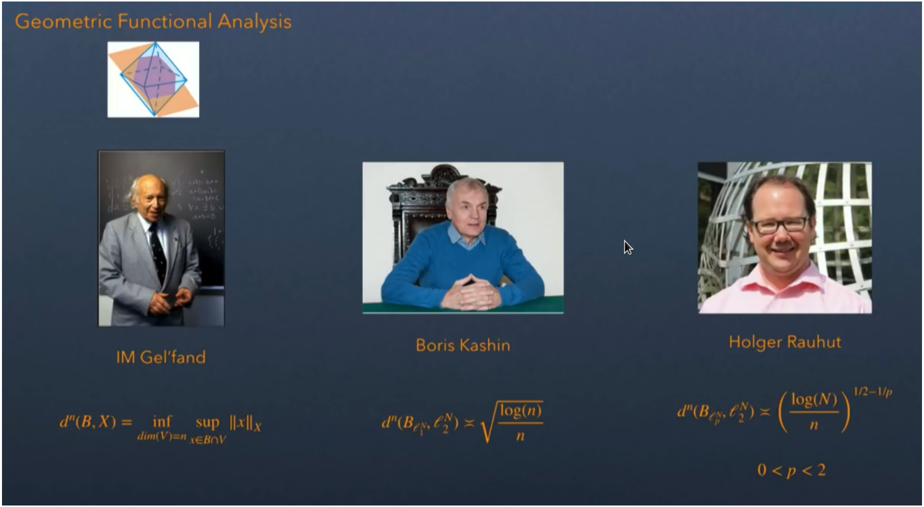


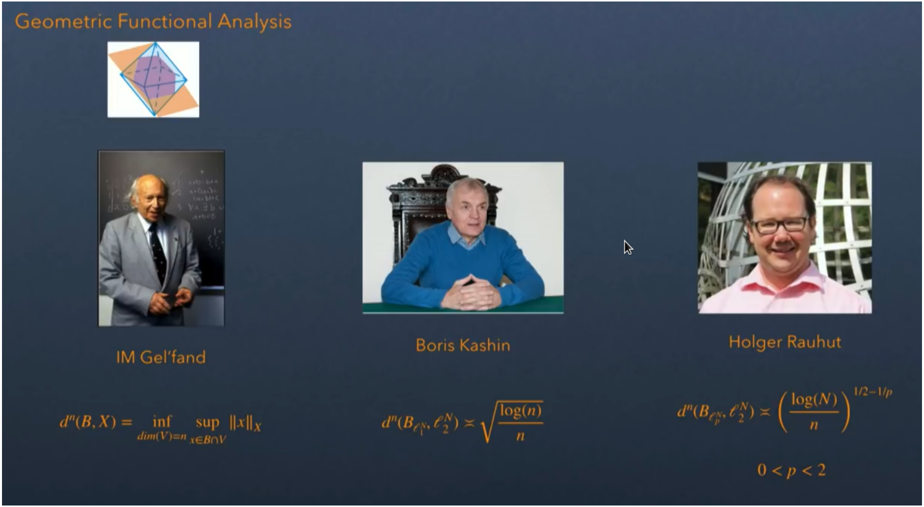
Figure 1: When Fourier coefficients of a testbed medical image known as the Logan–Shepp phantom (top left) are sampled along 22 radial lines in the frequency domain (top right), a naive, "minimal energy" reconstruction setting unobserved Fourier coefficients to 0 is marred by artifacts (bottom left).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

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

Лекция D.Donoho (премия Гаусса 2018)

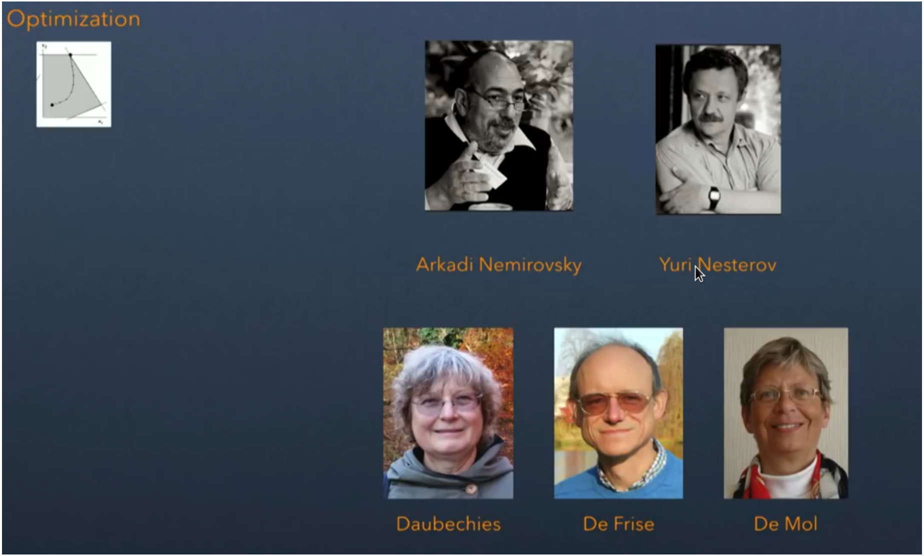


Лекция D.Donoho (премия Гаусса 2018)



Achtung: $0 < p < 2$!!

Лекция D.Donoho (премия Гаусса 2018)



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 $k \le Cm/\log(n/m)$.

Теорема (Candès-Romberg-Tao, 2004)

Пусть $\xi_1, \dots, \xi_m \in \{1, \dots, n\}$ выбраны случайно. Тогда с высокой вероятностью каждый k-разреженный сигнал $f : \{1, \dots, n\} \to \mathbb{C}$ 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 $m \ge 4k$.

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)* $\Phi : \mathbb{R}^N \to \mathbb{R}^n$ *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* $\|x\|^2 = 1$,

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

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 — разрежены (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)?

Пример 2: «dual» to compressive sensing: линейное кодирование с исправлением ошибок.

Сигнал $x \in \mathbb{R}^n$ преобразуется в длинный сигнал $\hat{A} \cdot x \in \mathbb{R}^m$ $(m > n)$ для передачи по сети с шумом. Если части переданного сигнала повреждены, полученные данные имеют вид $b = \hat{A} \cdot x + e$ для некоторого разреженного e, можно точно восстановить x во многих случаях.

Пример 2: «dual» to compressive sensing: линейное кодирование с исправлением ошибок.

Сигнал $x \in \mathbb{R}^n$ преобразуется в длинный сигнал $\hat{A} \cdot x \in \mathbb{R}^m$ $(m > n)$ для передачи по сети с шумом. Если части переданного сигнала повреждены, полученные данные имеют вид $b = \hat{A} \cdot x + e$ для некоторого разреженного e, можно точно восстановить x во многих случаях.

E. Candes, T. Tao. Decoding by linear programming. IEEE transactions on information theory 51(12) (2005), p. 4203–4215.

Обработка сигналов и изображений, сверхразрешение ...

Пример 2: «dual» to compressive sensing: линейное кодирование с исправлением ошибок.

Сигнал $x \in \mathbb{R}^n$ преобразуется в длинный сигнал $\hat{A} \cdot x \in \mathbb{R}^m$ $(m > n)$ для передачи по сети с шумом. Если части переданного сигнала повреждены, полученные данные имеют вид $b = \hat{A} \cdot x + e$ для некоторого разреженного e, можно точно восстановить x во многих случаях.

E. Candes, T. Tao. Decoding by linear programming. IEEE transactions on information theory 51(12) (2005), p. 4203–4215.

Важное направление исследований:

В этом докладе до сих пор предполагалось, что искомый x_i разрежен в «стандартном» базисе. Как правило, приходится сначала *найти* базис, в котором разложение заданного (не разреженного) вектора x_i окажется разреженным. Это скорее задача машинного обучения ...

Обработка сигналов и изображений, сверхразрешение ...

Пример 2: «dual» to compressive sensing: линейное кодирование с исправлением ошибок.

Сигнал $x \in \mathbb{R}^n$ преобразуется в длинный сигнал $\hat{A} \cdot x \in \mathbb{R}^m$ $(m > n)$ для передачи по сети с шумом. Если части переданного сигнала повреждены, полученные данные имеют вид $b = \hat{A} \cdot x + e$ для некоторого разреженного e, можно точно восстановить x во многих случаях.

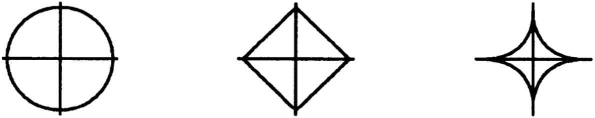
E. Candes, T. Tao. Decoding by linear programming. IEEE transactions on information theory 51(12) (2005), p. 4203–4215.

Важное направление исследований:

В этом докладе до сих пор предполагалось, что искомый x_i разрежен в «стандартном» базисе. Как правило, приходится сначала *найти* базис, в котором разложение заданного (не разреженного) вектора x_i окажется разреженным. Это скорее задача машинного обучения ...

Другой подход: брать избыточные базисы (линейно зависимые наборы N векторов, $N > \dim$), например объединения разных базисов, возможно в которых искомый вектор разрежен.

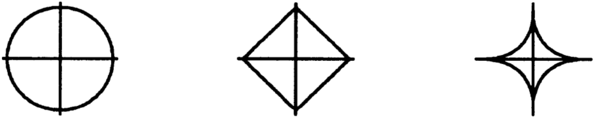
Невыпуклые задачи (ver. $N+1$, $N+2$, ...)



Задача (P_p^ε), $0 < p < 1$

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

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



Задача ($\mathbf{P}_p^\varepsilon$), $0 < p < 1$

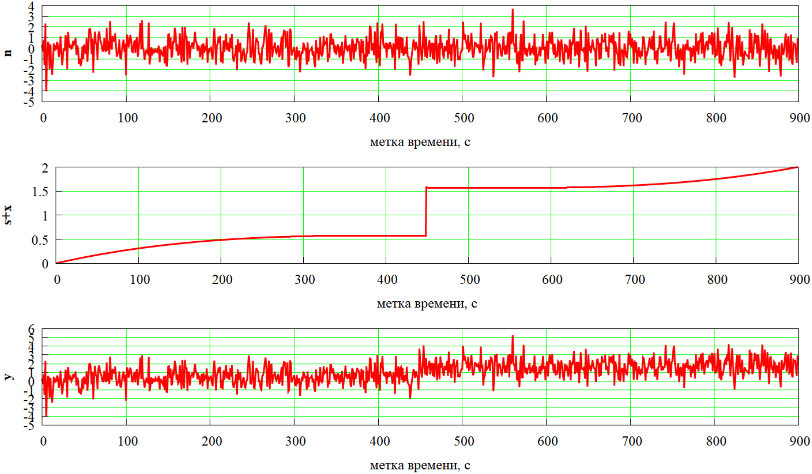
$$\begin{cases} x = \arg\min_x \{\|x\|_p = \sum_{i=1}^{N} |x_i|^p\}, \quad \text{restricted by} \\ \|\hat{A} \cdot x - b\|_2 < \varepsilon \end{cases}$$

Задача ($\mathbf{TV}_p^\varepsilon$), $0 < p \le 1$ для *total variation* (полной вариации)

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

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

Приложения: обнаружение скачков в зашумленных сигналах



Формализация модели и постановка задачи

- $s(t_i) = P(t_i) + x(t_i) + n(t_i)$
  где $s(t_i)$ — сигнал,
  $P(t_i)$ — медленно меняющийся тренд в измерениях;
  $x(t_i)$ — кусочно-постоянная функция скачков в фазовых измерениях;
  $n(t_i)$ — шумовая составляющая измерений;
  t_i — временной ряд, соответствующий моментам времени измерений (измерения идут с одинаковым шагом по времени, однако в таком ряду могут присутствовать пропуски).

Формализация модели и постановка задачи

- $s(t_i) = P(t_i) + x(t_i) + n(t_i)$
  где $s(t_i)$ — сигнал,
  $P(t_i)$ — медленно меняющийся тренд в измерениях;
  $x(t_i)$ — кусочно-постоянная функция скачков в фазовых измерениях;
  $n(t_i)$ — шумовая составляющая измерений;
  t_i — временной ряд, соответствующий моментам времени измерений (измерения идут с одинаковым шагом по времени, однако в таком ряду могут присутствовать пропуски).
- Задача оценки кусочно-постоянной функции методом минимизации полной вариации: $TV_p(x) = \sum_{i=1}^{N-1} |x(t_{i+1}) - x(t_i)|^p$

Формализация модели и постановка задачи

- $s(t_i) = P(t_i) + x(t_i) + n(t_i)$
  где $s(t_i)$ — сигнал,
  $P(t_i)$ — медленно меняющийся тренд в измерениях;
  $x(t_i)$ — кусочно-постоянная функция скачков в фазовых измерениях;
  $n(t_i)$ — шумовая составляющая измерений;
  t_i — временной ряд, соответствующий моментам времени измерений (измерения идут с одинаковым шагом по времени, однако в таком ряду могут присутствовать пропуски).
- Задача оценки кусочно-постоянной функции методом минимизации полной вариации: $TV_p(x) = \sum_{i=1}^{N-1} |x(t_{i+1}) - x(t_i)|^p$

$$\begin{cases} \widehat{x} = \arg\min_x TV_p(x) \\ \min_P ||s - P - \widehat{x}||_2 \leq \varepsilon, \quad P = a_0 + a_1 t_i + \dots \end{cases} \tag{5}$$

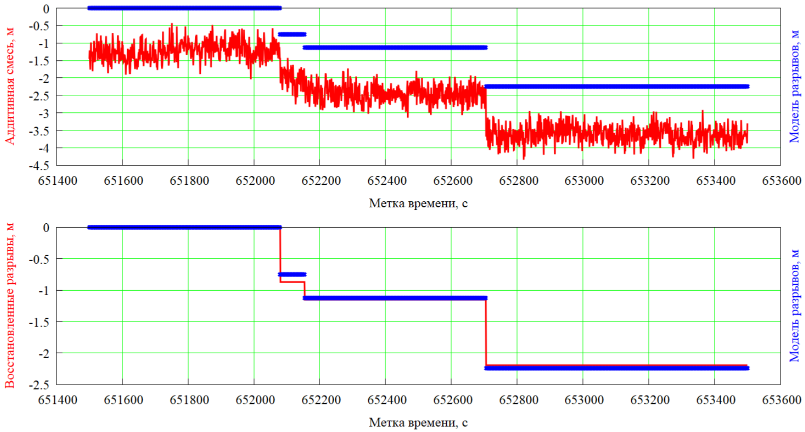
Формализация модели и постановка задачи

- $s(t_i) = P(t_i) + x(t_i) + n(t_i)$
  где $s(t_i)$ — сигнал,
  $P(t_i)$ — медленно меняющийся тренд в измерениях;
  $x(t_i)$ — кусочно-постоянная функция скачков в фазовых измерениях;
  $n(t_i)$ — шумовая составляющая измерений;
  t_i — временной ряд, соответствующий моментам времени измерений (измерения идут с одинаковым шагом по времени, однако в таком ряду могут присутствовать пропуски).
- Задача оценки кусочно-постоянной функции методом минимизации полной вариации: $TV_p(x) = \sum_{i=1}^{N-1} |x(t_{i+1}) - x(t_i)|^p$

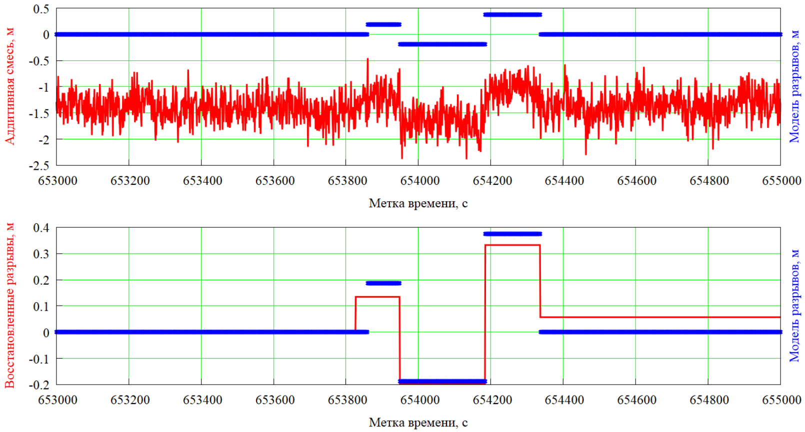
$$\begin{cases} \widehat{x} = \arg\min_x TV_p(x) \\ \min_P ||s - P - \widehat{x}||_2 \leq \varepsilon, \quad P = a_0 + a_1 t_i + \ldots \end{cases} \tag{5}$$

- Метод решения: модификация [Selesnik2012]

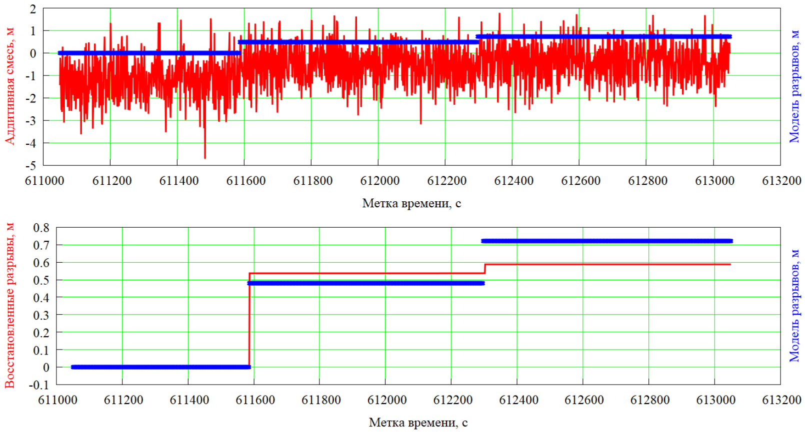
Численные эксперименты (1)



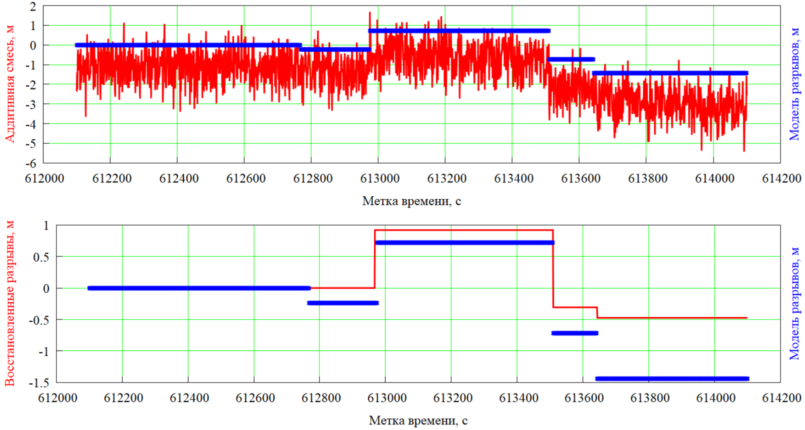
Численные эксперименты (2)



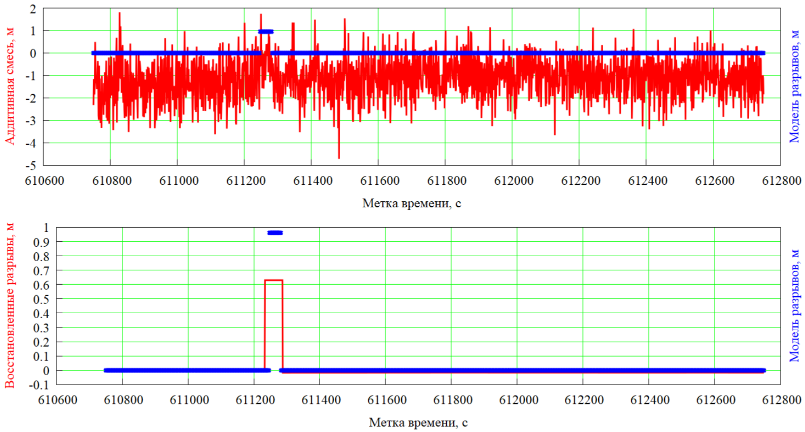
Численные эксперименты (3)



Численные эксперименты (4)



Численные эксперименты (5)



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.
- Tibshirani R. *Regression shrinkage and selection via the lasso* // J Royal Statistical Soc:B. 1996; v. 58, 267–288.
- 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

References

- Hochbaum DS. *An efficient and effective tool for image segmentation, total variations and regularization* // In: Int Conf on Scale Space and Variational Methods in Computer Vision 2011 (pp. 338–349)
- Polisano K, Condat L, Clausel M, Perrier V. *A convex approach to superresolution and regularization of lines in images* // SIAM J Imaging Sciences. 2019, 211–58.
- Selesnick I.W., Arnold S., Dantham V.R. *Polynomial smoothing of time series with additive step discontinuities* // IEEE Trans Signal Processing. 2012. v. 60, 6305–6318.
- 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.