слайд 1

Название курса: Современные методы анализа и обработки сигналов и изображений

слайд 2

Рекомендуемая литература

1. Захарова Т.В., Шестаков О.В. Теория вейвлетов и ее применение в обработке сигналов. М.: МастерПринт. 2018.

2. Скворцова Н.Н., Шестаков О.В., Малахов Д.В. Методы численного анализа стохастических сигналов. Лекции по курсу «прикладная радиофизика». М.: МИРЭА. 2011.

3. Ивченко Г.И., Медведев Ю.И. Математическая статистика. М.: Высш. шк. 1984.

4. Шестаков О.В. Вероятностные модели в томографии. М.: МАКС Пресс. 2008.

5. Малла С. Вейвлеты в обработке сигналов. М: Мир. 2005.

6. Троицкий И.Н. Статистическая теория томографии. М.: Радио и связь. 1989.

7. Boggess A., Narkowich F. A First Course in Wavelets with Fourier Analysis. Upper Saddle River: Prentice Hall, 2001.

слайд 3

Основные темы курса

1. Вероятностный анализ больших массивов данных, содержащих выборки сигналов/изображений.

2. Фурье-анализ и его применение для анализа/обработки стационарных сигналов/изображений. Асимптотические свойства методов Фурье-анализа при возрастании объема анализируемых данных.

3. Вейвлет-анализ и его применение для анализа/обработки нестационарных сигналов/изображений. Асимптотические свойства методов вейвлет-анализа при возрастании объема анализируемых данных.

4. Применение методов вейвлет-анализа для решения обратных статистических задач. Методы нелинейной регуляризации и их асимптотические свойства при возрастании объема анализируемых данных.

5. Методы реконструкции томографических изображений. Задачи локальной реконструкции и улучшения качества изображений.

слайд 4

Задачи анализа и обработки сигналов возникают во многих областях, включая радиотехнику, геофизику, астрономию, физику плазмы, медицину, экономику и многие другие.

На схеме представлена примерная классификация типов сигналов.

слайд 5

Математическое описание сигналов:

Аналоговый сигнал (analog signal) является непрерывной функцией непрерывного аргумента, т.е. определен для любого значения аргумента.

Дискретный сигнал (discrete signal) по своим значениям также является непрерывной функцией, но определенной только для дискретных значений аргумента. По множеству своих значений он является конечным (счетным) и описывается дискретной последовательностью отсчетов (samples) $y_i = y(n\Delta t\}$, где Δt - интервал между отсчетами, $n = 1, 2, ..., N$. Величина, обратная шагу дискретизации: $\Delta f = 1/\Delta t$, называется частотой дискретизации (sampling frequency). Если дискретный сигнал получен дискретизацией (sampling) аналогового сигнала, то он представляет собой последовательность отсчетов, значения которых в точности равны значениям исходного сигнала по координатам $n\Delta t$. При $\Delta t = const$ (равномерная дискретизация данных) дискретный сигнал можно описывать сокращенным обозначением $y(n)$. При неравномерной дискретизации сигнала обозначения дискретных последовательностей обычно заключаются в фигурные скобки - $\{x(t_i)\}$, а значения отсчетов приводятся в виде таблиц с указанием значений координат t_i.

Цифровой сигнал (digital signal) квантован по своим значениям и дискретен по аргументу. Он описывается квантованной решетчатой функцией $y_n = Q_k[y(n\Delta t)]$, где Q_k- функция квантования с числом уровней квантования k, при этом интервалы квантования могут быть как с равномерным распределением, так и с неравномерным, например - логарифмическим. Задается цифровой сигнал, как правило, в виде временной выборки (temporal sample) числовых данных - числового массива по последовательным значениям аргумента при $\Delta t = const$, но в общем случае сигнал может задаваться и в виде таблицы для произвольных значений аргумента.

В принципе, квантованными по своим значениям могут быть и аналоговые сигналы, зарегистрированные соответствующей аппаратурой (дискретно-аналоговые сигналы), которые принято называть дискретно-аналоговыми. Но выделять эти сигналы в отдельный тип не имеет смысла - они остаются аналоговыми кусочно-непрерывными сигналами с шагом квантования, который определяется допустимой погрешностью измерений.

Операция дискретизации осуществляет преобразование аналоговых сигналов (функций), непрерывных по аргументу, в функции мгновенных значений сигналов по дискретному аргументу. Дискретизация обычно производится с постоянным шагом по аргументу (*равномерная дискретизация*), значения $x(n\Delta t)$ представляют собой отсчеты функции $x(t)$ в моменты времени $t = n\Delta t$, $n = 1, 2, ..., N$. Частота, с которой выполняются замеры аналогового сигнала, называется *частотой дискретизации*. В общем случае, сетка отсчетов по аргументу может быть произвольной, как, например, $x(t_n), t_1, t_2, ..., t_N$ или задаваться по определенному закону. В результате дискретизации непрерывный (*аналоговый*) сигнал переводится в последовательность чисел.

Операция аналого-цифрового преобразования (АЦП) заключается в преобразовании дискретного сигнала $x(t_n)$ в цифровой сигнал $x(n) = x_n \approx x(t_n), n = 0, 1, 2, ..., N$, как правило, кодированный в двоичной системе счисления. Процесс преобразования отсчетов сигнала в числа называется квантованием по уровню, а возникающие при этом потери информации за счет округления – ошибками или шумами квантования. При преобразовании аналогового сигнала непосредственно в цифровой сигнал операции дискретизации и квантования совмещаются.

Цифро-аналоговое преобразование (ЦАП) преобразует цифровой код в аналоговый сигнал (ток, напряжение или заряд). Цифро-аналоговые преобразователи являются интерфейсом между дискретным цифровым миром и аналоговыми сигналами.

слайд 6

Корреляционный анализ применяется для поиска «связи» между двумя физическими явлениями (процессами).

Рассмотрим ансамбли двух временных выборок $\{x_n(t)\}$и $\{y_n(t)\}$ при достаточно большом объеме данных n. Мы предполагаем, что между ними существует взаимосвязь. Тогда должен существовать такой параметр a, что с его помощью в фиксированный момент времени t_0 временные выборки совмещаются наилучшим образом. В простейшем случае предполагаем линейную зависимость между двумя параметрами.

Значение параметра a не должно зависеть от конкретной пары выборок и индекса n, поэтому его величину находим из минимума выражения
$\frac{1}{n}\sum\limits_{k=1}^{n} x_k(t_0) - a\frac{1}{n}\sum\limits_{k=1}^{n} y_k(t_0) = \varepsilon(a, t_0),$

Это выражение будет минимальным по модулю (равным нулю), если
$a = a(t_0) = \frac{1}{n}\sum\limits_{k=1}^{n} x_k(t_0)/\frac{1}{n}\sum\limits_{k=1}^{n} y_k(t_0),$

При неограниченном возрастании числа временных выборок n параметр $a(t_0)$ при выполнении определенных условий (независимости, стационарности) стремится к отношению математических ожиданий
$a(t_0) \to \frac{Ex(t_0)}{Ey(t_0)}.$

Введенный выше параметр a обладает одним существенным недостатком. Предположим, что средние значения временных выборок $x(t_0)$ и $y(t_0)$ равны 0, тогда $\varepsilon(a, t_0) = 0$ при всех значениях a.

слайд 7

Введем иной критерий сравнения процессов $x(t)$ и $y(t)$ по ансамблям их временных выборок в момент времени t_0. Чтобы критерий отражал суть явлений, он должен быть связан с энергией, содержащейся в разности сигналов $\mathrm{x}(\mathrm{t}_0)$ и $\mathrm{y}(\mathrm{t}_0)$, поскольку, в конечном счете, нас интересует в сигнале его энергия или мощность. Рассмотрим величину
$z(b, t_0) = \frac{1}{n}\sum\limits_{k=1}^{n} |x_k(t_0) - by_k(t_0)|^2$

Найдем значение коэффициента b, минимизирующее $z(b, t_0)$. Для этого решим уравнение
$\frac{\partial z(b,t_0)}{\partial b} = 0$, т.е. $b = \frac{1}{n}\sum\limits_{k=1}^{n} x_k(t_0)y_k(t_0) \Big/ \frac{1}{n}\sum\limits_{k=1}^{n} |y_k(t_o)|^2$

При этом
$$z_{\min} = \frac{1}{n}[\sum\limits_{k=1}^{n} \left|x_k(t_0)\right|^2] \left[1 - \frac{\{\frac{1}{n}\sum\limits_{k=1}^{n} x_k(t_0)y_k(t_0)\}^2}{\frac{1}{n}\sum\limits_{k=1}^{n}|x_k(t_0)|^2\frac{1}{n}\sum\limits_{k=1}^{n}|y_k(t_0)|^2}\right]$$

Обозначим

$\rho = \frac{\frac{1}{n}\sum\limits_{k=1}^{n} x_k(t_0)y_k(t_0)}{\sqrt{\frac{1}{n}\sum\limits_{k=1}^{n}|x_k(t_0)|^2 \frac{1}{n}\sum\limits_{k=1}^{n}|y_k(t_0)|^2}}$, тогда $z_{\min} = \frac{1}{n}[\sum\limits_{k=1}^{n} |x_k(t_0)|^2] [1-\rho^2]$.

Величина ρ, характеризующая связь между двумя процессами, называется коэффициентом корреляции. $z_{\min} = 0$ при $\rho = \pm 1$.

Величину $\hat{C}_{x,y}(t_0, t_0) = \frac{1}{n}\sum\limits_{k=1}^{n} x_k(t_0)y_k(t_0)$ принято называть ковариацией. Можно рассмотреть связь между двумя процессами в различные моменты времени $x(t_0)$ и $y(t_0 - \tau)$. В этом случае ковариацию, зависящую от смещения τ, называют корреляционной функцией

$\hat{C}_{x,y}(t_0, \tau) = \frac{1}{n}\sum\limits_{k=1}^{n} x_k(t_0)y_k(t_0 - \tau)$.

Коэффициент корреляции и корреляционная функция тесно связаны с энергией взаимодействия между процессами. Отметим, что некореллированность является более слабым свойством случайных величин, чем их независимость.

слайд 8

К сожалению, в физическом эксперименте трудно (а часто невозможно) организовать измерения ансамбля временных реализаций какого-либо параметра случайного процесса. Легче осуществить измерения одной реализации, но сколь угодно большой длины.

Некоторые стационарные случайные процессы обладают свойством, заключающимся в том, что каждый член ансамбля реализаций ведет себя в статистическом смысле так же, как и весь ансамбль (эргодичность). В этом случае все характеристики случайного процесса можно проанализировать путем исследования свойств только одной реализации. Часто оказывается полезным предположить эргодичность процесса, если отсутствуют веские доводы физического характера, препятствующие этому. В физике из-за сложности создания ансамбля реализаций методы корреляционного анализа применяют, как правило, для исследования эргодических процессов. Обычно ограничиваются предположением об эргодичности 2-го порядка, т.е. возможностью вычислять по одной реализации моменты первого и второго порядка.

слайд 9

Корреляционная функция для эргодических процессов $x(t)$ и $y(t)$ вычисляется по одной реализации следующим образом:

$$C_{x,y}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} x(t)y(t-\tau)dt$$

Измерение корреляции между временными выборками на разных интервалах одного и того же процесса $x(t)$ позволяет изучать влияние предыстории процесса на его дальнейшее развитие. Такая функция называется автокорреляционной (АКФ).

$$C_{x,x}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} x(t)x(t-\tau)dt.$$

Изучение любой временной выборки полезно начинать с построения автокорреляционной функции, потому что она обладает целым рядом свойств, которые позволяют провести первичный анализ процесса. Автокорреляционная функция обладает следующими свойствами:

Четность $C_{x,x}(\tau) = C_{x,x}(-\tau)$.

Автокорреляционная функция имеет максимум при $\tau = 0$. Если $\tau \neq 0$, то $|C_{x,x}(\tau)| \leq C_{x,x}(0)$. При $\tau = 0$ величина $C_{x,x}(0)$ представляет собой среднее значение квадрата (дисперсии, средней мощности) процесса $x(t)$.

При неограниченном росте времени задержки τ $C_{x,x}(\infty) = 0$.

Автокорреляционная функция периодического сигнала является периодической функцией.

Автокорреляционная функция сохраняет частотную информацию о сигнале и теряет фазовую информацию.

слайд 10

Невозможно измерить бесконечную временную выборку процесса, в любом эксперименте происходит анализ временных выборок конечной длины.

Рассмотрим гармонический процесса с частотой ω_0, для которого интеграл АКФ может быть точно вычислена аналитически.

$$C_{x,x}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} A\sin(\omega_0 t)A\sin(\omega_0(t-\tau))dt =$$

$$= \lim_{T\to\infty} \left(\frac{A^2}{2}(\cos(\omega_0\tau) + \frac{\cos(\omega_0 T)}{2\omega_0 T}\sin(\omega_0\tau))\right).$$

Второе слагаемое стремиться к 0 при $T \to \infty$. Оценка АКФ по конечному отрезку времени:

$$C_{x,x}(\tau) = \frac{1}{T-\tau} \int\limits_{-T/2+\tau}^{T/2} x(t)x(t-\tau)dt$$

Следовательно, оценка АКФ близка к точному значению АКФ только при $T \gg T_0$, где $T_0 = 2\pi/\omega_0$- период измеряемого гармонического сигнала. Таким образом, в оценке АКФ гармонической функции ошибка возрастает при уменьшении T или при уменьшении частоты гармонического сигнала.

слайд 11

Для временной выборки $\{x_k\}$ при интервале дискретизации данных $\Delta t = const$ АКФ вычисляется по интервалам $\Delta\tau = \Delta t$ и обычно записывается, как дискретная функция номеров n сдвига отсчетов - $n\Delta t$:

$$C_{x,x}(n) = \frac{1}{n} \sum_{k=-\infty}^{\infty} x_k x_{k+n}.$$

Дискретная АКФ имеет такие же свойства, как и непрерывная АКФ. Она также является четной, а ее значение при $n = 0$ равно мощности дискретного сигнала.

Дискретные сигналы обычно задаются в виде числовых массивов определенной длины с нумерацией отсчетов $k = 0, 1, ..., N$, а вычисление дискретной АКФ выполняется в одностороннем варианте с учетом длины массивов по формуле:

$$C_{x,x}(n) = \frac{N}{(N+1-n)n} \sum_{k=0}^{N-n} x_k x_{k+n}.$$

Множитель $N/(N + 1 - n)$ в данной функции является поправочным коэффициентом на постепенное уменьшение числа перемножаемых и суммируемых значений (от Nдо $N - n$) по мере увеличения сдвига n.

слайд 12

Как правило, следует придерживаться двух условий:

1. выбирать временной интервал более чем в 5 раз превышающий наибольший период колебаний

2. совмещать данные не более чем на 20% их длины

Полезно провести сравнение автокорреляционной функции временной выборки с δ-функцией — автокорреляционной функцией белого шума. Если автокорреляционная функция отлична от δ-функции, что характерно для многих стохастических сигналов (в радиотехнике, экономике, геофизике и др.), то необходимо использовать дополнительные методы анализа данных (например, кроме спектрального дополнительно вероятностный анализ).

Существование такой автокорреляционной функции указывает на некоторую функцию влияния, приводящую к неоднородности и не независимости исследуемой временной выборки, что не позволяет анализировать по ней природу случайного процесса. Для анализа необходимо перейти к выборке случайного процесса, которая была бы независимой и однородной. Для решения указанной задачи, можно применить подход, известный в математической статистике (прежде всего, в анализе временных рядов). А именно, предлагается осуществить переход от статистического анализа амплитуд временной выборки случайного процесса $\{x_n(t)\}, n = 0, 1, ..., N$ к анализу и моделированию его приращений (по сути первой производной по времени) $\{\Delta x_n(t)\}, n = 0, 1, ..., N - 1$.

$\Delta x_n = x_n - x_{n-1}$.

Естественно, если автокорреляционная функция для приращений тоже не окажется δ- функцией то необходимо рассчитать второе приращения и т.д. Как правило, уже третий (а иногда и второй) шаг становится невозможным из-за ухода сигнала (из-за малой разрядности АЦП) под шум. Однако если получена действительно однородная независимая выборка, то возможен анализ ее вероятностных характеристик.

слайд 13

Вероятностный анализ временных выборок

Как уже отмечалось, для анализа необходимо перейти к выборке, которая была бы независимой и однородной. При этом объем выборки n должен быть достаточно большим. Для наглядного представления статистических данных часто используются гистограммы. Гистограмма временной выборки $\{x_k\}$ объема n представляет собой столбиковый график, построенный по полученным за определенный период данным, которые разбиваются на m интервалов (количество интервалов может зависеть от n). Чаще всего длины интервалов выбирают одинаковыми, хотя это и не обязательно. Обозначим длину интервала разбиения через Δ. Число данных (элементов временной выборки), попадающих в j-й интервал, обозначим через ν_j. Высота столбика диаграммы на j-м интервале разбиения равна $\frac{\nu_j}{n\Delta}$. Гистограмма может быть построена для любой временной выборки. Однако, как правило, гистограммы оцениваются и анализируются для случайных стационарных сигналов.

слайд 14

Форма гистограммы дает первичное представление о процессе. Опишем наиболее распространенные формы:

а) симметричный или колоколообразный. Максимальная высота гистограммы оказывается в середине и постепенно снижается к концам. Форма симметрична. Такой тип гистограмм встречается чаще всего. В частности, колоколообразную форму имеет гауссово распределение, и по виду гистограммы можно в первом приближении сделать вывод об отклонении распределения наблюдаемой временной выборки от гауссова.

б) Положительно скошенная (отрицательно скошенная) гистограмма. Максимальная высота гистограммы локализуется слева (справа) от центра. Высоты довольно резко спадают при движении влево (вправо) и, наоборот, медленно вправо (влево). Форма асимметрична.

в) Двухпиковый тип (бимодальный тип). В окрестностях центра гистограммы высота низкая, зато есть по пику с каждой стороны. Такая форма встречается, когда смешиваются два распределения с далеко отстоящими средними значениями. Как правило, это говорит о разнородности данных.

Существуют и другие формы гистограммы, однако подробно останавливаться на них мы не будем.

Можно также строить и более гладкие оценки плотности (интерполяция – сглаженные гистограммы, ядерные оценки и т.п.).

слайд 15

Знание моментов распределения также многое может сказать о его виде и свойствах. Введем выборочные аналоги неизвестных истинных моментов распределения.

Пусть a, σ^2, b_l - теоретические среднее (математическое ожидание), дисперсия и l-й момент: $a = E\xi$, $\sigma^2 = E(\xi - E\xi)^2$, $b_l = E\xi^l$. Соответствующие выборочные характеристики определяются следующим образом:

$\overline{X} = \frac{1}{n}\sum\limits_{k=1}^{n} x_k$- выборочное среднее, оценка a,

$S^2 = \frac{1}{n}\sum\limits_{k=1}^{n}(x_k - \overline{X})^2$- выборочная дисперсия, оценка σ^2,

$\overline{X}^l = \frac{1}{n}\sum\limits_{k=1}^{n} x_k^l$ - выборочный l-й момент, оценка b_l.

Выборочное среднее $\overline{X}$ является несмещенной и состоятельной оценкой для теоретического среднего (математического ожидания), т.е. $E\overline{X} = a$ и $\overline{X}$ стремится к aпо вероятности при $n \to \infty$. Выборочный l-й момент также является несмещенной и состоятельной оценкой теоретического l-го момента. Выборочная дисперсия S^2 является состоятельной, но смещенной оценкой теоретической дисперсии: $ES^2 = \frac{n-1}{n}\sigma^2 \neq \sigma^2$.

Список числовых характеристик и их выборочных аналогов можно продолжить, рассмотрев, например, центральные и абсолютные моменты. Важную роль при изучении свойств распределения временной выборки наблюдаемого процесса играют коэффициенты асимметрии μ_3 и эксцесса μ_4 (для случайной величины ξ эти характеристики определяются следующим образом: $\mu_3 = \frac{E(\xi - E\xi)^3}{\sigma^3}$ и $\mu_4 = \frac{E(\xi - E\xi)^4}{\sigma^4} - 3$. Слагаемое -3 вводится для того, чтобы коэффициент эксцесса стандартного гауссова распределения был равен нулю), а также их выборочные аналоги:

$M_3 = \frac{A_3}{(S^2)^{3/2}}$, $A_3 = \frac{1}{n}\sum\limits_{k=1}^{n}(x_k - \overline{X})^3$ - выборочный центральный момент третьего порядка, и

$M_4 = \frac{A_4}{(S^2)^2} - 3$, где $A_4 = \frac{1}{n}\sum\limits_{k=1}^{n}(x_k - \overline{X})^4$ - выборочный центральный момент четвертого порядка.

Коэффициент асимметрии задает степень асимметричности плотности вероятности относительно оси, проходящей через ее центр тяжести. Коэффициент асимметрии определяется третьим центральным моментом распределения. В любом симметричном распределении с нулевым математическим ожиданием, например, гауссовом, все нечетные моменты, в том числе и третий, равны нулю, поэтому коэффициент асимметрии тоже равен нулю. Коэффициентом эксцесса показывает, насколько острую вершину имеет плотность вероятности по сравнению с гауссовым распределением. Если коэффициент эксцесса больше нуля, то распределение имеет более острую вершину, чем распределение Гаусса, если меньше нуля, то более плоскую.

Выборочный коэффициент эксцесса наряду с выборочным коэффициентом асимметрии часто используется для грубой предварительной проверки гипотезы о гауссовости распределения временной выборки.

слайд 16

Выборочные моменты также часто используются для построения оценок неизвестных параметров распределения временной выборки наблюдаемого процесса. При рассмотрении параметрических моделей любой теоретический момент b_l (если он существует) функционально зависит от некоторого параметра модели θ (θ может быть вектором), т.е. $b_l = f_l(\theta)$, $l = 1, ..., k$(k- размерность θ). Если функция f (векторная) обратима, то параметр θ является функцией от $b_1, ..., b_k$: $\theta = f^{-1}(b_1, ..., b_k)$. Подставив в эту функцию вместо неизвестного теоретического l-го момента b_l его выборочный аналог, мы получим оценку параметра θ: $\theta^* = f^{-1}(\overline{X}, ..., \overline{X}^k)$. Если функция f^{-1} непрерывна, то оценки параметров, получаемые методом моментов, являются состоятельными.

Замечание. При построении оценки параметра методом моментов может получиться так, что значение θ^* не принадлежит множеству допустимых значений параметра θ. В этом случае оценку корректируют. Например, в качестве θ^* берут ближайшую к $f^{-1}(\overline{X}, ..., \overline{X}^k)$ точку из области допустимых значений θ.

слайд 17

Статистические критерии позволяют проверять достоверность построенной модели поведения случайного процесса, в основе которой лежит гипотеза о виде распределения временной выборки этого процесса, базирующаяся на каких-то предварительных предположениях о его характере и предельных теоремах, следующих из этих предположений. Одним из самых простых и универсальных статистических критериев является критерий Хи-квадрат Пирсона.

Критерий Хи-квадрат Пирсона основан на группировке данных. Пусть истинное распределение временной выборки $\{x_k\}$ равно F. Проверяется основная гипотеза о том, что $F = F_0$, против альтернативы $F \neq F_0$. Для этого область значений предполагаемого теоретического распределения F_0 делится на некоторое число интервалов. Пусть, как и раньше, это число равно m, а число данных (элементов временной выборки), попадающих в j-й интервал, равно ν_j. Обозначим через p_j^0 ($p_j^0 > 0$, $p_1^0 + ... + p_m^0 = 1$) теоретическую вероятность попадания в j-й интервал случайной величины с распределением F_0. Вычислим величину

$$\mathrm{X}^2 = \sum_{j=1}^{m} \frac{(\nu_j - np_j^0)^2}{np_j^0}.$$

Величина X^2 называется статистикой Хи-квадрат. Фактически эта величина представляет собой квадрат некоего расстояния между двумя m-мерными векторами: вектором наблюдаемых относительных частот $\frac{\nu_j}{n}$ и вектором предполагаемых ненаблюдаемых вероятностей p_j^0. От евклидового расстояния это расстояние отличается тем, что разные координаты входят в него с разными весами. Критерий Хи-квадрат Пирсона основывается на сравнении величины X^2с некоторым порогом (о котором будет сказано чуть ниже). Если X^2 больше этого порога, то основная гипотеза о том, что $F = F_0$ отвергается, если же X^2 меньше порога, то считается, что основная гипотеза не противоречит данным наблюдений. Таким образом, на самом деле критерий Хи-квадрат Пирсона решает следующую задачу: для заданного набора вероятностей $p_1^0, \ldots, p_m^0$ ($p_j^0 > 0$, $p_1^0 + ... + p_m^0 = 1$) проверяется основная гипотеза том, что истинное распределение временной выборки обладает свойством: «вероятность попадания в j-й интервал равна p_j^0 для всех $j = 1, \ldots, m$» против сложной альтернативы: «хотя бы для одного из интервалов эта вероятность отличается от p_j^0».

слайд 18

Применимость критерия Хи-квадрат обосновывается теоремой Пирсона, которая утверждает, что если основная гипотеза верна, то при фиксированном m и при $n \to \infty$ распределение величины X^2 сближается с распределением χ^2 с $m - 1$ степенями свободы.

Чтобы понять, что происходит, когда основная гипотеза неверна, заметим, что по закону больших чисел относительные частоты $\frac{\nu_j}{n}$ стремятся к теоретическим вероятностям p_j при $n \to \infty$ для всех $j = 1, \ldots, m$. Поэтому при $n \to \infty$ величина $\mathrm{X}^2 = \sum_{j=1}^{m} \frac{(\nu_j - np_j^0)^2}{np_j^0}$ с большой вероятностью близка к величине $\sum_{j=1}^{m} \frac{(np_j - np_j^0)^2}{np_j^0} = n \sum_{j=1}^{m} \frac{(p_j - p_j^0)^2}{p_j^0}$. Если основная гипотеза неверна, то эта величина стремится к бесконечности при $n \to \infty$. Значит, основная гипотеза должна быть отвергнута, если полученное значение X^2 слишком велико. Термин "слишком велико"означает, что значение X^2 имеет малую вероятность, то есть превосходит критическое значение (порог), которое в силу теоремы Пирсона можно взять из таблиц распределения χ^2. Таким образом, окончательно критерий Хи-квадрат имеет следующий вид: пусть задана вероятность α (уровень значимости критерия), тогда если $X^2 \geq \chi_{1-\alpha,m-1}$, то основная гипотеза отвергается (Порог $\chi_{1-\alpha,m-1}$ называется $(1-\alpha)$-квантилью распределения χ^2 с $m-1$ степенями свободы, его значение берется из таблиц. Смысл этого значения заключается в том, что для случайной величины, имеющей распределение χ^2 с $m-1$ степенями свободы, вероятность превзойти порог $\chi_{1-\alpha,m-1}$ не превосходит α.), в противном случае гипотеза не противоречит данным наблюдений.

Асимптотический характер теоремы Пирсона, лежащий в основе этого правила, требует осторожности при его практическом использовании. На него можно полагаться только при больших n (большом объеме данных). Достаточно велико должно быть и n, и частоты ν_j, и все произведения np_j^0. Проблема применимости аппроксимации χ^2 (непрерывное распределение) к статистике X^2, распределение которой дискретно, довольно сложна. Согласно имеющемуся опыту, аппроксимация применима, если $n \geq 50$, $\nu_j \geq 5$ для всех $j = 1, \ldots, m$ и для ожидаемых теоретических вероятностей $np_j^0 > 10$ для всех $j = 1, \ldots, m$.

слайд 19

Критерий Хи-квадрат также часто применяют для проверки гипотезы о принадлежности распределения временной выборки некоторому параметрическому семейству, т.е. проверяется основная гипотеза о том, что $F \in \{F_\theta\}$, где $\theta \in \Theta \subset R^r$– неизвестный параметр (скалярный или векторный), а r – его размерность (количество параметров). Например, для гауссова распределения в качестве неизвестных параметров могут выступать математическое ожидание и дисперсия.

Пусть, как и раньше, область значений предполагаемого теоретического распределения делится на m интервалов, и число элементов временной выборки, попадающих в j-й интервал, равно ν_j. Однако теоретические вероятности попадания в j-й интервал теперь зависят от неизвестного параметра θ: $p_j^0 = p_j^0(\theta)$. Величина X^2 также зависит от θ

$X^2(\theta) = \sum\limits_{j=1}^{m} \frac{(\nu_j - np_j^0(\theta))^2}{np_j^0(\theta)}$,

и для того, чтобы воспользоваться критерием, необходимо вместо θ подставить его оценку. Пусть $\hat{\theta}$ – значение параметра θ, доставляющее минимум функции $X^2(\theta)$ при данном значении временной выборки $\{x_k\}$. Подставив вместо истинных вероятностей p_j^0 их оценки $p_j^0(\hat{\theta})$, получим величину

$$\mathrm{X}^2(\hat{\theta}) = \sum_{j=1}^{m} \frac{(\nu_j - np_j^0(\hat{\theta}))^2}{np_j^0(\hat{\theta})}.$$

Для величины $\mathrm{X}^2(\hat{\theta})$ справедлива теорема Фишера, аналогичная теореме Пирсона, однако теперь распределение $\mathrm{X}^2(\hat{\theta})$, если основная гипотеза верна, сближается с распределением χ^2 с $m-r-1$ степенями свободы. Далее процедура проверки гипотезы строится так же, как в предыдущем случае.

Оценку $\hat{\theta}$, минимизирующую функцию $\mathrm{X}^2(\theta)$ можно также получить, как оценку максимального правдоподобия для θ, построенную по вектору $\nu_1, \ldots, \nu_m$ из полиномиального распределения. В общем случае вычисление такой оценки, равно как и вычисление точки минимума функции $\mathrm{X}^2(\theta)$, возможно лишь численно.

Недостатком метода проверки достоверности модели на основе критерия Хи-квадрат является то, что группировка данных по интервалам приводит к некоторой потере информации. Кроме того, остается вопрос о выборе числа интервалов m и длине самих интервалов. Однако критерий Хи-квадрат имеет и некоторые достоинства: при его применении нет необходимости учитывать точные значения наблюдений, и несомненным преимуществом является универсальность этого критерия.

слайд 20

Оценка достоверности моделирования распределения временной выборки по критерию Колмогорова и Колмогорова-Смирнова

Еще одним известным критерием проверки гипотезы о виде распределения временной выборки случайного процесса является критерий Колмогорова. Этот критерий, так же как и критерий Хи-квадрат, служит для проверки основной гипотезы о том, что распределение F временной выборки $\{x_k\}$ равно F_0 против альтернативы $F \neq F_0$. Он применяется в тех случаях, когда функция F_0 непрерывна. Для критерия Колмогорова используется так называемая эмпирическая функция распределения:

$$F_n(x) = \frac{1}{n} \sum_{k=1}^{n} I(x_k \leq x),$$

Используя упорядоченную выборку, эмпирическую функцию распределения можно также записать в виде

$$F_n(x) = \frac{1}{n} \sum_{k=1}^{n} I(x_{(k)} \leq x).$$

Для построения критерия вычисляется величина

$D_n = \sup\limits_{-\infty<x<\infty} |F_n(x) - F_0(x)|,$

равная максимальному отклонению эмпирической функции распределения $F_n(x)$ от предполагаемой теоретической функции распределения $F_0(x)$. Фактически эта величина представляет собой равномерное расстояние между функциями $F_n(x)$ и $F_0(x)$. При каждом x функция $F_n(x)$ стремится п.в. к $F(x)$ при $n \to \infty$, и с ростом nпроисходит сближение $F_n(x)$ с теоретической функцией распределения $F(x)$, поэтому при больших n в том случае, когда гипотеза $F = F_0$ истинна, значение D_n не должно существенно отклоняться от нуля. Отсюда следует, что основная гипотеза должна отвергаться, если величинаD_n больше некоторого порога (критической границы), речь о котором пойдет чуть ниже. На практике вычисление величины D_n представляет собой более трудоемкую задачу, чем вычисление величины X^2, однако при этом не происходит той потери информации, которая наблюдается при использовании критерия Хи-квадрат.

слайд 21

Особенностью величины D_n является то обстоятельство, что ее распределение при истинности гипотезы $F = F_0$, не зависит от вида функции F_0. Действительно, полагая $x = F_0^{-1}(u)$, $0 \le u \le 1$, где $F_0^{-1}(u)$ - функция, обратная к $F_0(x)$, получаем

$D_n = \sup\limits_{-\infty<x<\infty} |F_n(F_0^{-1}(u)) - u|.$

Перейдем к новой выборке, используя формулу $u_k = F_0(x_k)$, $k = 1, \ldots, n$. Пусть $\{u_{(k)}\}$ - упорядоченная выборка, соответствующая выборке $\{u_k\}$. Функция $F_0(x)$ монотонна, поэтому $u_{(k)} = F_0(x_{(k)})$, и неравенства $x_{(k)} \le F_0^{-1}(u)$ эквивалентны неравенствам $u_{(k)} \le u$. Но тогда имеем

$F_n(F_0^{-1}(u)) = \frac{1}{n}\sum\limits_{k=1}^{n} I(x_{(k)} \le F_0^{-1}(u)) = \frac{1}{n}\sum\limits_{k=1}^{n} I(u_{(k)} \le u).$

Независимо от вида функции$F_0(x)$ (если только $F_0(x)$ непрерывна) величины u_k имеют равномерное распределение на отрезке [0,1]. Поэтому если $R_n(u)$ - эмпирическая функция распределения для выборки из равномерного распределения на отрезке [0,1], то $F_n(F_0^{-1}(u)) = R_n(u)$ и

$D_n = \sup\limits_{-\infty<x<\infty} |R_n(u) - u|.$

Следовательно, распределение D_n не зависит от $F_0(x)$ (если справедлива гипотеза о том, что $F = F_0$). Этот факт имеет принципиальное значение, так как достаточно вычислить и протабулировать распределение D_n только один раз, а именно для выборки из равномерного распределения на отрезке [0,1], и использовать ее для проверки гипотезы относительно произвольной непрерывной функции распределения $F_0(x)$.

слайд 22

Другой полезной особенностью величины D_n является ее асимптотическое поведение при $n \to \infty$. По теореме Колмогорова, если верна основная гипотеза, то

$$\lim_{n\to\infty} P(\sqrt{n}D_n \le t) = K(t) = \sum_{j=-\infty}^{\infty} (-1)^j e^{-2j^2t^2}.$$

При этом предельную функцию распределения $K(t)$ можно с хорошим приближением использовать для практических расчетов уже при $n \ge 20$. Отсюда следует, что при $n \ge 20$ критическую границу следует искать по следующему правилу, которое и составляет суть критерия Колмогорова: пусть задана вероятность α (уровень значимости критерия), тогда если $D_n \ge \frac{t_\alpha}{\sqrt{n}}$, то основная гипотеза отвергается, в противном случае делается вывод, что гипотеза не противоречит данным наблюдений. Порог t_α удовлетворяет соотношению $K(t_\alpha) = 1 - \alpha$. Смысл t_α заключается в том, что для случайной величины, имеющей функцию распределения $K(t)$, вероятность превзойти порог t_α не превосходит α. Значение t_αберется из таблиц.

Для небольших значений n точное распределение D_n также протабулировано и для расчета точных значений критической границы можно воспользоваться соответствующими таблицами.

слайд 23

Для проверки гипотезы о том, что две выборки $\{x_k\}$ и $\{y_k\}$ объема n и m имеют одинаковое распределение, используется критерий Колмогорова-Смирнова, основанный на величине

$$D_{nm} = \sup_{-\infty<x<\infty} |F_{1n}(x) - F_{2m}(x)|,$$

Где $F_{1n}(x)$ и $F_{2m}(x)$ - эмпирические функции распределения, построенные по выборкам $\{x_k\}$ и $\{y_k\}$ соответственно. Так как эмпирическая и теоретическая функции распределения при увеличении объема выборки сближаются, то в случаях, когда гипотеза верна, функции $F_{1n}(x)$ и $F_{2m}(x)$ при достаточно больших n и m должны быть близки друг к другу, и величина D_{nm} не должна существенно отклоняться от нуля. Следовательно, если D_{nm} больше некоторой критической границы, то гипотезу следует отвергнуть. Критическую границу при заданном уровне значимости α находят на основании теоремы Смирнова, которая утверждает, что если гипотеза о равенстве распределений верна, то

$\lim\limits_{n,m\to\infty} P(\sqrt{\frac{nm}{n+m}}D_{nm} \leq t) = K(t).$

Таким образом, суть критерия Колмогорова-Смирнова заключается в следующем правиле: пусть задан уровень значимости критерия α, если $D_{nm} \geq t_\alpha\sqrt{(n+m)/nm}$, где$K(t_\alpha) = 1 - \alpha$, то гипотеза о равенстве распределений отвергается, в противном случае гипотеза принимается. При этом вероятность ошибочно отвергнуть истинную гипотезу приблизительно равна α.

Указанное правило не зависит от конкретного вида теоретической функции распределения. Для приложений это имеет важное значение, так как истинное распределение временной выборки, как правило, неизвестно, и интерес представляет вопрос о том, не изменялось ли оно от выборки к выборке. Для применения критерия Колмогорова-Смирнова необходимо выполнение только условия непрерывности, которое обычно вытекает из физической природы изучаемого процесса и не требует специальной проверки.

слайд 24

Анализ параметра Херста для зависимых временных выборок

В практических ситуациях предположения о независимости и однородности временных выборок процесса часто нарушаются. Как уже говорилось, иногда ситуацию можно улучшить, перейдя к рассмотрению приращений процесса, однако это не всегда приводит к желаемому результату. В таких случаях можно рассмотреть модели, в основе которых лежат так называемые автомодельные процессы, открытые Колмогоровым и нашедшие свое применение в теории турбулентности и других областях физики. При рассмотрении таких процессов дисперсия выборочного среднего с ростом объема данных n убывает со скоростью $n^{-\alpha}$ (n - объем выборки), где α заключено между 0 и 2 (для независимых выборок дисперсия убывает со скоростью n^{-1}). Мандельброт обозначил параметр автомодельности или параметр сохранения дальнодействующей зависимости в процессе через H в честь Херста, который изучал статистические данные по разливам реки Нил и обнаружил эффект долгоживущих корреляций (см. [6]). Параметр $H = 1 - \frac{\alpha}{2}$ может изменяться от 0 до 1; случай $H = \frac{1}{2}$ соответствует независимости, $H > \frac{1}{2}$ - положительным корреляциям, а $H < \frac{1}{2}$ - отрицательным корреляциям.

Изучать автокорреляционные функции автомодельных процессов довольно сложно. Хотя все индивидуальные корреляции могут быть очень малыми и, следовательно, незначащими при проверке гипотезы, их сумма стремится к бесконечности, что влечет за собой бесконечную "память"и значимую корреляцию даже между далеким прошлым и отдаленным будущим. С другой стороны именно такие непривычные свойства необходимы для описания эмпирических выводов.

В настоящее время имеется несколько методов оценивания параметра H. Рассмотрим один из них: метод R/S-анализа, где R/S - это размах процесса накопления, деленный на стандартное отклонение процесса приращений для некоторого промежутка времени (t_0, t_0+t). Для временной выборки процесса отношениеR/S определяется следующим образом: $\frac{R}{S} = \frac{R(n)}{S(n)} = \frac{\max(0, Z_1, \ldots, Z_n) - \min(0, Z_1, \ldots, Z_n)}{\sqrt{S^2}}$,

где S^2 - выборочная дисперсия, а $Z_k = x_1 + \ldots + x_k - k\overline{X}$.

Точки графика $(\log n, \log \frac{R(n)}{S(n)})$ концентрируются вокруг прямой с наклоном H. Для определения H применяется метод наименьших квадратов.

Для почти нормальных данных этот метод менее эффективен, чем другие, но зато гораздо более робастен (т.е. невосприимчив к выбросам) и дает хорошие оценки параметра H даже для распределений с тяжелыми хвостами и ассиметричных распределений. Есть также и другие методы. В частности, основанные на Фурье-анализе и вейвлет-анализе.

слайд 25

Фурье-анализ является мощным аналитическим инструментом, применяемым в различных научных областях при исследовании стационарных колебательных процессов и явлений. При помощи Фурье-анализа пространственная или временна́я функция раскладывается на синусоидальные составляющие, каждая из которых имеет свою частоту, амплитуду и фазу. Это позволяет удобно исследовать стационарные процессы и обрабатывать стационарные сигналы.

Пусть функция $f(x)$ определена на отрезке $[-\pi, \pi]$. Тригонометрическим разложением функции называется сумма вида

$f(x) = a_0 + \sum_{k=1}^{\infty}(a_k \cos kx + b_k \sin kx)$.

Пример: Предположим, что сигнал задается функцией

$3 \sin 2t + \cos 4t - 40 \sin 7t + 5 \sin 200t$,

тогда он содержит волны с частотами 2, 4, 7 и 200. И, судя по размеру коэффициентов, можно сказать, что частота 7 доминирует над остальными частотами

Отметим, что для звукового сигнала волна с частотой 200 несет помехи типа шипения и свиста. Это – так называемый высокочастотный шум. Задача удаления высокочастотного шума является одной из самых важных при обработке сигнала. В данном примере нужно обнулить соответствующий коэффициент.

слайд 26

В приведенном примере, удалив из сигнала волновую компоненту с частотой 200, мы получим гладкий сигнал.

Еще одной важной задачей является задача сжатия данных для передачи сигнала с наименьшими затратами. Для этого сигнал f также раскладывают в тригонометрический ряд, а потом пересылаются только те коэффициенты a_k и b_k, которые больше, чем некоторое определенное значение. Коэффициенты, меньшие этого значения, не сильно влияют на величину сигнала и поэтому могут быть отброшены, причем число больших коэффициентов всегда конечно, так как в силу леммы Римана-Лебега a_k и b_k стремятся к нулю при $k \to \infty$.

слайд 27

Скалярным произведением $\langle f, g \rangle$ функций f и g из $L^2([a, b])$ называется $\langle f, g \rangle = \int\limits_a^b f(x)\overline{g(x)}dx$. Норма $\|f\| = \sqrt{\langle f, f \rangle}$.

Функции f и g ортогональны, если $\langle f, g \rangle = 0$.

Рассмотрим отрезок $[-\pi, \pi]$. Система

$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \ldots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}, \ldots$ ортонормирована в $L^2([-\pi, \pi])$.

слайд 28

Это вытекает из следующих соотношений. Рассмотрим тригонометрический ряд Фурье

$f(x) = a_0 + \sum_{k=1}^{\infty}(a_k \cos kx + b_k \sin kx)$.

В силу ортонормированности

$a_n = \frac{1}{\pi}\int\limits_{-\pi}^{\pi} f(x)\cos nx dx$,

$b_n = \frac{1}{\pi}\int\limits_{-\pi}^{\pi} f(x)\sin nx dx$,

$a_0 = \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} f(x) dx$.

слайд 29

Пример – прямоугольный импульс с периодическим продолжением.

$f(x) = \begin{cases} 0, & \text{при } -\pi \le x < 0 \\ 1, & \text{при } \quad 0 \le x < \pi. \end{cases}$

Разложение f в ряд Фурье выглядит следующим образом:

$f(x) = \frac{1}{2} + \sum_{k=1}^{\infty} \frac{1-(-1)^k}{\pi k} \sin kx$.

слайд 30

Аппроксимация функции f частичными суммами f_n ряда Фурье при $n = 3$ и $n = 11$

$f_n(x) = \frac{1}{2} + \sum_{k=1}^{n} \frac{1-(-1)^k}{\pi k} \sin kx$.

Наблюдается эффект Гиббса. В точках разрыва всплески (вверх и вниз). С ростом n ширина всплесков уменьшается, а высота нет.

слайд 31

Перейдем к рассмотрению других интервалов.

Лемма. Пусть $F(x)$ – некоторая 2π-периодическая функция и c – любое вещественное число, тогда $\int\limits_{-\pi+c}^{\pi+c} F(x)dx = \int\limits_{-\pi}^{\pi} F(x)dx$.

Следовательно, коэффициенты разложения на любом отрезке длиной 2π вычисляются по аналогичным формулам.

Рассмотрим отрезок $[-a, a]$. Функции $\cos(n\pi x/a)$ и $\sin(n\pi x/a)$ периодичны с периодом $2a$ и $f(x) = a_0 + \sum_{k=1}^{\infty}(a_k \cos(k\pi x/a) + b_k \sin(k\pi x/a))$, тогда $a_0 = \frac{1}{2a}\int\limits_{-a}^{a} f(t)dt$,

$a_n = \frac{1}{a}\int\limits_{-a}^{a} f(t)\cos(n\pi t/a)dt$,

$b_n = \frac{1}{a}\int\limits_{-a}^{a} f(t)\sin(n\pi t/a)dt$.

слайд 32

Если $f(x)$ четная, то в разложении только косинусы и a_0, если $f(x)$ нечетная – только синусы.

Если $f(x)$ определена на $[0, a]$, ее можно продолжить на $[-a, 0]$ четным или нечетным образом.

слайд 33

Комплексная форма ряда Фурье

Комплексная форма ряда Фурье короче и симметричнее своего вещественного аналога и поэтому чаще используется.

Вспомним основные свойства экспоненты от мнимого аргумента.

Последовательность функций

$\left\{\frac{e^{int}}{\sqrt{2\pi}}, n = 0, \pm 1, \pm 2, \dots\right\}$

является ортонормированной в пространстве $L^2([-\pi, \pi])$.

Для разложимой в ряд Фурье на отрезке $[-\pi, \pi]$ функции справедливо представление

$f(x) = \sum\limits_{n=-\infty}^{\infty} \alpha_n e^{inx}$,

где коэффициенты α_n, называемые коэффициентами Фурье, определяются формулами

$\alpha_n = \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} f(t)e^{-int}dt, \quad n \in Z.$

слайд 34

Связь между вещественной и комплексной формами рядов Фурье

Для простоты рассмотрим функцию, разложимую в ряд Фурье на отрезке $[-\pi, \pi]$.

$f(x) = \alpha_0 + \sum\limits_{n=1}^{\infty} \alpha_n e^{inx} + \sum\limits_{n=-\infty}^{-1} \alpha_n e^{inx}$

Для вещественной функции $\alpha_{-n} = \overline{\alpha_n}$. Следовательно,

$f(x) = \alpha_0 + \sum\limits_{n=1}^{\infty} \alpha_n e^{inx} + \overline{\sum\limits_{k=1}^{\infty} \alpha_k e^{ikx}}$.

Так как для комплексного числа $z + \overline{z} = 2\operatorname{Re} z$,

$f(x) = \alpha_0 + 2\operatorname{Re}\left(\sum\limits_{n=1}^{\infty} \alpha_n e^{inx}\right)$.

Далее

$\alpha_0 = \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} f(t)dt = a_0$,

$\alpha_n = \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} f(t)e^{-int}dt = \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} f(t)(\cos nt - i\sin nt)dt = \frac{1}{2}(a_n - ib_n)$.

слайд 35

$f(x) = \alpha_0 + 2\operatorname{Re}\left(\sum\limits_{n=1}^{\infty} \alpha_n e^{inx}\right) =$

$= a_0 + \operatorname{Re}\left(\sum\limits_{n=1}^{\infty} (a_n - ib_n)(\cos nt + i\sin nt)\right) =$

$= a_0 + \sum\limits_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$.

Эту цепочку равенств можно провести в обратном порядке, тем самым получив комплексную форму ряда Фурье из вещественной.

Комплексная форма более естественна. Разложение по волнам одной частоты. В каждой волне $|\alpha_n|$ – амплитуда, $arg(\alpha_n)$ – фаза (сдвиг относительно начала координат).

слайд 36

Рассмотрим теперь отрезок $[-a, a]$.

Последовательность функций

$$\left\{\frac{1}{\sqrt{2a}} e^{\frac{in\pi x}{a}}, n = 0, \pm 1, \pm 2, \dots\right\}$$

является ортонормированной в пространстве $L^2([-a, a])$.

Для разложимой в ряд Фурье на отрезке $[-a, a]$ функции

$$f(x) = \sum_{n=-\infty}^{\infty} \alpha_n e^{\frac{in\pi x}{a}}$$

$$\alpha_n = \frac{1}{2a} \int_{-a}^{a} f(t) e^{-\frac{in\pi x}{a}} dx, \quad n \in Z.$$

слайд 37

Существенную роль играет вопрос сходимости тригонометрического ряда Фурье функции f к самой этой функции в точке x. Сформулируем основные утверждения. Теоремы сходимости опираются на следующую лемму Римана-Лебега, которая интересна и сама по себе.

Лемма Римана-Лебега Если f кусочно-непрерывная функция на отрезке $[a, b]$, то

$$\lim_{k \to \infty} \int_a^b f(x) \cos kx dx = \lim_{k \to \infty} \int_a^b f(x) \sin kx dx = 0.$$

Отсюда следует, что только конечное число коэффициентов Фурье может быть больше по модулю любого заданного положительного числа. Этот факт лежит в основе метода сжатия данных. Сначала сигнал раскладывается в ряд Фурье, затем слагаемые с маленькими коэффициентами отбрасываются, и остается только конечное число слагаемых с коэффициентами большими некоторого порога чувствительности.

Для простоты далее рассмотрим только функции на отрезке $[-\pi, \pi]$.

Теорема 1. Пусть f – непрерывная 2π-периодическая функция, тогда в каждой точке x, в которой определена производная f, ряд Фурье функции f сходится к $f(x)$.

В данной теореме на функцию f накладывается довольно жесткое условие одновременной непрерывности и периодичности. Это требование снимается следующей теоремой сходимости для разрывных функций.

Теорема 2. Пусть f – кусочно-непрерывная функция 2π-периодическая. Предположим, что в точке x существует конечное левое и правое предельное значение производной функции f, тогда ряд Фурье функции f в точке x сходится к $\frac{f(x+0)+f(x-0)}{2}$.

В точках разрыва возникает эффект Гиббса.

слайд 38

Рассмотрим теперь равномерную сходимость рядов Фурье.

Напомним, что последовательность функций $F_n(x)$ сходится равномерно к $F(x)$ на отрезке $[a, b]$, если $\forall\varepsilon > 0\ \exists N$ (которое не зависит от x) такое, что при $n \geq N$ $|F_n(x) - F(x)| < \varepsilon$ для всех $x \in [a, b]$.

Ряд Фурье функции f сходится равномерно к $f(x)$ на отрезке $[-\pi, \pi]$, если последовательность частичных сумм

$S_N(x) = a_0 + \sum\limits_{k=1}^{N}(a_k \cos kx + b_k \sin kx)$,

где a_k и b_k – коэффициенты Фурье функции f, сходится равномерно к $f(x)$ на отрезке $[-\pi, \pi]$ при $N \to \infty$.

Определение. Функция называется кусочно-гладкой, если она непрерывна и её производная существует всюду, за исключением, быть может, конечного или счетного числа точек.

Сформулируем теорему о равномерной сходимости ряда Фурье на отрезке $[-\pi, \pi]$.

Теорема 3. Ряд Фурье кусочно-гладкой 2π-периодической функции f равномерно сходится к $f(x)$ на отрезке $[-\pi, \pi]$ $(f(-\pi) = f(\pi))$.

Заметим, что теоремы остаются верны в случае замены π на любое положительное число a для произвольной кусочно-гладкой $2a$-периодической функции.

слайд 39

Напомним, что пространство $L^2([-\pi, \pi])$ состоит из квадратично интегрируемых функций (т.е. таких функций f, что $\int\limits_{-\pi}^{\pi} |f(x)|^2 dx < \infty$) со скалярным произведением $\langle f, g\rangle$. Норма $\|f\|$ в этом пространстве определяется формулой

$\|f\|^2 = \int\limits_{-\pi}^{\pi} |f(x)|^2 dx.$

Справедливы:

$\langle f, g\rangle \leq \|f\|\|g\|$ – неравенство Коши-Буняковского

$\|f + \| \leq \|f\| + \|g\|$ – неравенство треугольника

Лемма Пусть $f \in L^2([-\pi, \pi])$, a_k и b_k – коэффициенты Фурье функции f и

$$f_N(x) = a_0 + \sum_{k=1}^{N}(a_k \cos kx + b_k \sin kx).$$

Пусть V_N – линейная оболочка, натянутая на систему

$\{1, \cos kx, \sin kx, \ \ 1 \le k \le N\}$. Тогда f_N – ближайший из V_N к f элемент в L^2-норме: $\|f - f_N\| = \min\limits_{g \in V_N} \|f - g\|$.

слайд 40

Сходимость рядов Фурье в среднем.

Теорема 4. Пусть $f \in L^2([-\pi, \pi])$ и

$f_N(x) = a_0 + \sum\limits_{k=1}^{N} (a_k \cos kx + b_k \sin kx)$,

где a_k и b_k – коэффициенты Фурье функции f, тогда f_N сходится к f в пространстве $L^2([-\pi, \pi])$, т.е. $\|f_N - f\| \to 0$ при $N \to \infty$.

Эта теорема для комплексной формы ряда Фурье выглядит следующим образом.

Теорема 4′. Пусть $f \in L^2([-\pi, \pi])$, α_n – коэффициенты комплексной формы ряда Фурье для f, тогда частичные суммы $f_N(x) = \sum\limits_{k=-N}^{N} \alpha_k e^{ikx}$ сходятся к f по норме пространства $L^2([-\pi, \pi])$ при $N \to \infty$.

Эти теоремы остаются верны в случае замены π на любое положительное число a.

слайд 41

Следствием важным следствием этих теорем являются следующие утверждения.

Теорема 5. (Равенство Парсеваля в вещественной форме)

Пусть $f \in L^2([-\pi, \pi])$ и $f(x) = a_0 + \sum\limits_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$, тогда

$\frac{1}{\pi} \int\limits_{-\pi}^{\pi} |f(x)|^2 dx = 2|a_0|^2 + \sum\limits_{k=1}^{\infty} \left(|a_k|^2 + |b_k|^2\right)$.

Теорема 5′. (Равенство Парсеваля в комплексной форме)

Пусть $f \in L^2([-\pi, \pi])$ и $f(x) = \sum\limits_{k=-\infty}^{\infty} \alpha_k e^{ikx}$, тогда

$\frac{1}{2\pi} \|f\|^2 = \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} |f(x)|^2 dx = \sum\limits_{k=-\infty}^{\infty} |\alpha_k|^2$.

Несложно показать, что для функций f и g из $L^2([-\pi, \pi])$ с соответствующими коэффициентами Фурье α_n и β_n справедливо равенство

$\frac{1}{2\pi}\langle f, g\rangle = \sum\limits_{n=-\infty}^{\infty} \alpha_n \overline{\beta_n}.$

Норма функции (сигнала) в пространстве L^2 в физике интерпретируется как энергия сигнала, квадрат коэффициента Фурье – часть энергии соответствующей частотной компоненты. В этом свете равенство Парсеваля показывает, что энергия сигнала складывается из энергий каждой частотной компоненты сигнала.

Также простым следствием теоремы $5'$ является неравенство Бесселя: Для любого натурального N: $\sum\limits_{k=-N}^{N} |\alpha_k|^2 \leq \frac{1}{2\pi}\|f\|^2.$

слайд 42

Преобразование Фурье (ПФ) является непрерывным аналогом ряда Фурье. Ряд Фурье представляет собой разложение периодического сигнала на волны с дискретным набором частот, т.о. сигнал представляется в виде суперпозиции гармонических колебаний. Аналогичное представление возможно и для непериодических сигналов, но уже в виде так называемого интеграла Фурье.

Рассмотрим ряд Фурье

$f(x) = \sum\limits_{k=-\infty}^{\infty} \alpha_k e^{\frac{ik\pi x}{a}},\ x \in [-a, a],\ \alpha_k = \frac{1}{2a}\int\limits_{-a}^{a} f(t)e^{-\frac{ik\pi t}{a}}dt.$

Пусть $a \to \infty$ и f удовлетворяет условиям для перехода к пределу. Тогда

$f(x) = \lim\limits_{a\to\infty} \sum\limits_{k=-\infty}^{\infty} \left(\frac{1}{2a}\int\limits_{-a}^{a} f(t)e^{-\frac{ik\pi t}{a}}dt\right) e^{\frac{ik\pi x}{a}} =$

$= \lim\limits_{a\to\infty} \sum\limits_{k=-\infty}^{\infty} \left(\frac{1}{2a}\int\limits_{-a}^{a} f(t)e^{\frac{ik\pi(x-t)}{a}}dt\right).$

Обозначим $\omega_k = \frac{k\pi}{a}$, $\Delta\omega = \omega_{k+1} - \omega_k = \frac{\pi}{a}$. Тогда $\Delta\omega \to 0$.

Получаем интегральную сумму

$f(x) = \lim\limits_{a\to\infty} \sum\limits_{k=-\infty}^{\infty} \left(\frac{1}{2\pi}\int\limits_{-a}^{a} f(t)e^{i\omega_k(x-t)}dt\right)\Delta\omega.$

$f(x) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(t)e^{i\omega(x-t)}dtd\omega = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}\left(\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} f(t)e^{-i\omega t}dt\right)e^{i\omega x}d\omega.$

$\widehat{f}(\omega) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} f(t)e^{-i\omega t}dt$ – прямое преобразование Фурье.

$f(x) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \widehat{f}(\omega) e^{i\omega x} d\omega$ – обратное преобразование Фурье.

слайд 43

Теорема 1. Если f непрерывно дифференцируемая функция и $\int\limits_{-\infty}^{\infty} |f(x)| dx < \infty$, тогда

$f(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} \widehat{f}(\omega) e^{i\omega x} d\omega$, где $\widehat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$.

ω играет роль частоты $\left(\omega_k = \frac{k\pi}{a}\right)$, $\widehat{f}(\omega)$ имеет смысл, аналогичный коэффициентам α_k в ряде Фурье.

$\left|\widehat{f}(\omega)\right|$ – амплитуда соответствующая ω, $arg\widehat{f}(\omega)$ – фаза.

У рядов Фурье и преобразований Фурье разные области определения.

Ряд Фурье для f на конечном отрезке – дискретный спектр сигнала.

Преобразование Фурье для f на R – непрерывный спектр сигнала.

слайд 44

Установим наиболее важные свойства преобразования Фурье. Введем сначала еще одно обозначение преобразования Фурье: $\mathcal{F}[f](\omega) = \widehat{f}(\omega)$. Оно удобно тем, что вводит понятие оператора преобразования Фурье $\mathcal{F}$. Таким образом, $\mathcal{F}[f] = \widehat{f}$. Аналогично вводится понятие обратного оператора преобразования Фурье $\mathcal{F}^{-1}$: $\mathcal{F}^{-1}[\mathcal{F}[f]] = f$.

Основные свойства преобразования Фурье и обратного преобразования Фурье сформулируем в следующей теореме.

Теорема 2. Пусть f и g – дифференцируемые функции вещественной переменной и $f(x) = 0$ на бесконечности, т.е. при $|x| > M$ (M – некоторое положительное число), тогда прямое и обратное преобразования Фурье обладают следующими свойствами:

1. Линейность: $\mathcal{F}[f+g] = \mathcal{F}[f] + \mathcal{F}[g], \quad \mathcal{F}[cf] = c\mathcal{F}[f]$.

$\mathcal{F}^{-1}[f+g] = \mathcal{F}^{-1}[f] + \mathcal{F}^{-1}[g], \quad \mathcal{F}^{-1}[cf] = c\mathcal{F}^{-1}[f]$.

2. Дифференцирование $\mathcal{F}[x^n f(x)](\omega) = i^n \frac{\partial^n}{\partial \omega^n} \{\mathcal{F}[f](\omega)\}$.

$\mathcal{F}^{-1}[\omega^n f(\omega)](x) = (-i)^n \frac{\partial^n}{\partial x^n} \{\mathcal{F}^{-1}[f](x)\}$.

3. Связь между степенью гладкости функции и скоростью убывания на бесконечности ПФ и обратного ПФ (чем больше производных в L^1

имеет f, тем быстрее убывает на бесконечности ее преобразование Фурье):

$\mathcal{F}[f^{(n)}](\omega) = (i\omega)^n \mathcal{F}[f](\omega)$,

$\mathcal{F}^{-1}[f^{(n)}](x) = (-ix)^n \mathcal{F}^{-1}[f](x)$,

где $f^{(n)}$ означает n-ую производную функции f.

4. ПФ сдвига функции (при сдвиге f ее ПФ сдвигается по фазе):

$\mathcal{F}[f(x-a)](\omega) = e^{-i\omega a}\mathcal{F}[f](\omega)$.

5. ПФ при масштабировании функции (при растяжении (сжатии) сигнала частота уменьшается (увеличивается)):

$\mathcal{F}[f(bx)](\omega) = \frac{1}{|b|}\mathcal{F}[f]\left(\frac{\omega}{b}\right)$.

слайд 45

Определение. Пусть f и g – интегрируемые функции. Тогда их сверткой называется функция $f * g(x)$, определяемая как

$$f * g(x) = \int\limits_{-\infty}^{\infty} f(x-t)g(t)dt = \int\limits_{-\infty}^{\infty} f(t)g(x-t)dt.$$

Теорема 3. Пусть f и g – две интегрируемые функции, тогда

$\mathcal{F}[f * g] = \sqrt{2\pi}\ \widehat{f} \cdot \widehat{g}, \quad \mathcal{F}^{-1}[\widehat{f} \cdot \widehat{g}] = \frac{1}{\sqrt{2\pi}}\ f * g$.

Еще одно свойство преобразования Фурье $\mathcal{F}$: оператор, сопряженный к $\mathcal{F}$, является обратным оператором к $\mathcal{F}$.

Теорема 4. Пусть f и g – квадратично интегрируемые функции, тогда $\langle \mathcal{F}[f], g\rangle = \langle f, \mathcal{F}^{-1}[g]\rangle$.

Следующее свойство – формула Планшереля, состоит в том, что преобразование Фурье сохраняет значение скалярного произведения (а, следовательно, и норму) в пространстве L^2.

Теорема 5. Пусть f и g – квадратично интегрируемые функции, тогда $\langle \mathcal{F}[f], \mathcal{F}[g]\rangle = \langle f, g\rangle$, $\langle \mathcal{F}^{-1}[f], \mathcal{F}^{-1}[g]\rangle = \langle f, g\rangle$.

В частности, $\|\mathcal{F}[f]\| = \|f\|$ (равенство Парсеваля).

Теорема 5 является аналогом равенства Парсеваля для рядов Фурье, т.к. также говорит о сохранении норм $\|\mathcal{F}[f]\| = \|f\|$. Это тождество имеет физическую интерпретацию. Как и в рядах Фурье формула Планшереля утверждает, что энергия сигнала во временнóй области $\|f\|^2$ равна энергии сигнала в частотной $\|\widehat{f}\|^2$.

слайд 46

До сих пор преобразование Фурье определялось для абсолютно интегрируемых функций (пространства $L^1(R)$). Распространим теперь определение преобразования Фурье на все функции из $L^2(R)$ с помощью теоремы 5.

Преобразование Фурье, определенное как интеграл – ограниченный линейный оператор $\mathcal{F} : L^1(R) \bigcap L^2(R) \to L^2(R)$. $L^1(R) \bigcap L^2(R)$ плотно в $L^2(R)$. Равенство Парсеваля, справедливое для всех функций $f \in L^2(R)$, распространяет преобразование Фурье на $L^2(R)$ с сохранением нормы: выберем последовательность $f_n \in L^1(R) \bigcap L^2(R)$, сходящуюся к f по норме ($f \in L^2(R)$): $\lim\limits_{n\to\infty} \|f_n - f\| = 0$. Так как имеет место сходимость, то $\|f_n - f_m\|$ мало при больших n, m. По равенству Парсеваля $\|f_n - f_m\| = \|\widehat{f}_n - \widehat{f}_m\|$. Поскольку пространство $L^2(R)$ полное, существует функция $\widehat{f} \in L^2(R)$: $\lim\limits_{n\to\infty} \|\widehat{f}_n - \widehat{f}\| = 0$. По определению полагаем эту функцию $\widehat{f}$ преобразованием Фурье функции $f \in \mathrm{L}^2(R)$ (f может не принадлежать $L^1(R)$).

слайд 47

Можно обобщить рассмотренное одномерное преобразование Фурье на многомерный случай.

Пусть $x = (x_1, x_2, \ldots, x_n)$ — элемент n-мерного евклидова пространства R^n, а функция $f(x) = f(x_1, x_2, \ldots, x_n)$ определена на R^n и для нее сходится интеграл

$\int\limits_{R^n} |f(x)|\, dx = \int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} |f(x_1, x_2, \ldots, x_n)|\, dx_1 dx_2 \ldots dx_n < \infty.$

Тогда для функции $f(x)$ определено многомерное преобразование Фурье $\widehat{f}$ в точке $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$:

$\widehat{f}(\omega) = \frac{1}{(\sqrt{2\pi})^n} \int\limits_{R^n} f(x) e^{-i(\omega,x)} dx,$

где $(\omega, x) = \sum i = 1^n \omega_i x_i$ – скалярное произведение векторов ω и x.

При этом функцию $f(x)$ можно восстановить по ее образу Фурье:

$f(x) = \frac{1}{(\sqrt{2\pi})^n} \int\limits_{R^n} \widehat{f}(\omega) e^{i(\omega,x)} d\omega.$

Все свойства одномерного преобразования Фурье переносятся соответствующим образом на многомерный случай. Сформулируем некоторые из них. Для краткости изложения рассмотрим случай $n = 2$, который нам понадобится в дальнейшем.

1. Линейность:

$\mathcal{F}[af+bg](\omega_1,\omega_2)=a\mathcal{F}[f](\omega_1,\omega_2)+b\mathcal{F}[g](\omega_1,\omega_2)$,

$\mathcal{F}^{-1}[af+bg](x_1,x_2)=a\mathcal{F}^{-1}[f](x_1,x_2)+b\mathcal{F}^{-1}[g](x_1,x_2)$.

2. Дифференцирование ПФ:

$\mathcal{F}[x_1^{n_1}x_2^{n_2}f(x_1,x_2)](\omega_1,\omega_2)=i^{n_1+n_2}\frac{\partial^{n_1+n_2}}{\partial\omega_1^{n_1}\partial\omega_2^{n_2}}\{\mathcal{F}[f](\omega_1,\omega_2)\}$.

Дифференцирование обратного ПФ:

$\mathcal{F}^{-1}[\omega_1^{n_1}\omega_2^{n_2}g(\omega_1,\omega_2)](x_1,x_2)=(-i)^{n_1+n_2}\frac{\partial^{n_1+n_2}}{\partial x_1^{n_1}\partial x_2^{n_2}}\{\mathcal{F}^{-1}[g](x_1,x_2)\}$.

3. Связь между степенью гладкости функции и скоростью убывания на бесконечности ее ПФ:

$\mathcal{F}\left[\frac{\partial^{n_1+n_2}f(x_1,x_2)}{\partial x_1^{n_1}\partial x_2^{n_2}}\right](\omega_1,\omega_2)=i^{n_1+n_2}\omega_1^{n_1}\omega_2^{n_2}\mathcal{F}[f](\omega_1,\omega_2)$.

Связь между степенью гладкости функции и скоростью убывания на бесконечности обратного ПФ:

$\mathcal{F}^{-1}\left[\frac{\partial^{n_1+n_2}g(\omega_1,\omega_2)}{\partial\omega_1^{n_1}\partial\omega_2^{n_2}}\right](x_1,x_2)=(-i)^{n_1+n_2}x_1^{n_1}x_2^{n_2}\mathcal{F}^{-1}[g](x_1,x_2)$.

слайд 48

4. ПФ сдвига функции:

Обозначим $x=(x_1,x_2)$, $a=(a_1,a_2)$, тогда

$\mathcal{F}[f(x-a)](\omega_1,\omega_2)=e^{-i(\omega_1a_1+\omega_2a_2)}\mathcal{F}[f](\omega_1,\omega_2)$.

Тем самым ПФ переводит сдвиг по аргументу в сдвиг по фазе.

5. ПФ при масштабировании функции:

$\mathcal{F}[f(b_1x_1,b_2x_2)](\omega_1,\omega_2)=\frac{1}{|b_1b_2|}\mathcal{F}[f]\left(\frac{\omega_1}{b_1},\frac{\omega_2}{b_2}\right)$,

при $b_1\neq 0$, $b_2\neq 0$.

Пусть f и g – две интегрируемые функции на плоскости. Их сверткой назовем функцию $f*g$, определяемую как

$$f*g(x_1,x_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(u_1,u_2)g(x_1-u_1,x_2-u_2)du_1du_2.$$

6. Преобразование Фурье для сверток:

$\mathcal{F}[f*g](\omega_1,\omega_2)=2\pi\widehat{f}(\omega_1,\omega_2)\ \widehat{g}(\omega_1,\omega_2)$.

7. Формула Планшереля: $\langle f,g\rangle=\langle\widehat{f},\widehat{g}\rangle$. Положив $f=g$, получим равенство Парсеваля: $\|f\|=\left\|\widehat{f}\right\|$.

8. Пусть $\widehat{f}(\omega_1,\omega_2)$ – преобразование Фурье функции $f(x_1,x_2)$. Если функцию $f(x_1,x_2)$ повернуть на угол θ:

$f_\theta(x_1,x_2)=f(x_1\cos\theta-x_2\sin\theta,x_1\sin\theta+x_2\cos\theta)$,

то ее преобразование Фурье также будет повернуто на угол θ:

$\widehat{f}_\theta(\omega_1,\omega_2)=\widehat{f}(\omega_1\cos\theta-\omega_2\sin\theta,\omega_1\sin\theta+\omega_2\cos\theta)$.

слайд 49

Линейные фильтры.

На основе Фурье-анализа была развита теория фильтрации. Формально фильтр – это преобразование L, которое переводит сигнал f (обычно являющийся кусочно-гладкой функцией) в сигнал $\widetilde{f}$.

Фильтр L называется линейным, если выполняются свойства:

$L[f+g] = L[f] + L[g]$, $L[cf] = cL[f]$, где c – константа.

Еще одним свойством, присущим большинству фильтров является инвариантность по времени. Введем следующее обозначение: $f_a(t) = f(t-a)$. Тем самым функция f_a является задержкой сигнала f на a единиц по времени.

Определение. Фильтр L называется инвариантным по времени, если для любого сигнала f и для любого вещественного числа a, $L[f_a](t) = (Lf)(t-a)$ для всех t (или через введенное выше обозначение $L[f_a] = (Lf)_a$, т.е. L и задержка перестановочны).

Следующая лемма показывает, что комплексные синусоидальные волны $e^{i\omega t}$ являются собственными функциями линейных инвариантных по времени фильтров, т.е., если на вход такого фильтра поступает синусоидальный сигнал с частотой ω, то и на выходе будет также синусоидальный сигнал с той же частотой ω, но с другой амплитудой.

Лемма. Пусть L – линейный фильтр, инвариантный по времени, и ω – некоторое вещественное число. Тогда для L найдется такая функция $h(t)$, что $L(e^{i\omega t}) = \sqrt{2\pi}\,\widehat{h}(\omega)e^{i\omega t}$.

Следующая теорема позволяет сделать вывод о том, что любой линейный и инвариантный по времени фильтр представляет собой свертку сигнала с импульсным откликом h.

Теорема. Пусть L – инвариантный по времени линейный фильтр на пространстве сигналов f, которые являются кусочно-непрерывными функциями. Тогда существует интегрируемая функция h такая, что для любого сигнала f $L(f) = f * h$.

h называется импульсным откликом, так как свертка с δ-функцией (импульсом) на выходе дает h (фильтр преобразует импульс в h. $\widehat{h}$ называется системной функцией.

слайд 50

Дискретное преобразование Фурье

Аппарат преобразования Фурье и рядов Фурье используется для анализа непрерывных сигналов. Однако, на практике часто требуется исследовать сигнал заданный в виде числовой последовательности. Например, для цифровой обработки сигналов на компьютерах непрерывный сигнал аппроксимируется дискретным. В этом случае используется дискретное преобразование Фурье (ДПФ).

Обозначим через S_n множество n-периодических последовательностей $\{y\}$. Каждая последовательность $y = \{y_j\}_{j=-\infty}^{\infty}$ является периодическим дискретным сигналом, где y_j – это значение сигнала в момент времени $t = t_j$. Последовательность y_j называется n-периодической, если $y_{k+n} = y_k$ для любого целого числа k.

Определение. Пусть $y = \{y_j\}_{j=-\infty}^{\infty} \in S_n$. Дискретным преобразованием Фурье сигнала y называется последовательность $\{\widehat{y}_k\}_{k=-\infty}^{\infty}$, где

$$\widehat{y}_k = \sum_{j=0}^{n-1} y_j \exp(-2\pi i k j/n).$$

Непосредственное применение формулы ДПФ требует $O(n^2)$ операций умножения. При большом числе отсчетов n осуществление такого количества умножений может занять много времени. Например, для анализа непрерывной функции, заданной в 1000 точках, равномерно распределенных на интервале, необходимо выполнить примерно один миллион умножений.

слайд 51

Этот расчет может быть сделан за $O(n \log_2 n)$ операций с использованием алгоритма быстрого преобразования Фурье (БПФ).

Рассмотрим работу алгоритма БПФ в случае, когда n является степенью двойки.

Пусть $\{y\}$ – n-периодическая последовательность.

Введем множители $W_n^j = \exp(-2\pi i j/n)$.

Тогда формулу ДПФ можно переписать в виде

$$\widehat{y}_k = \sum_{j=0}^{n-1} y_j \, W_n^{kj}.$$

Разобьем $\{y\}$ на две $n/2$-периодические подпоследовательности $\{y^e\}$ и $\{y^o\}$, где $\{y^e\}$ состоит из членов последовательности $\{y\}$ с четными номерами, а $\{y^o\}$ состоит из членов последовательности $\{y\}$ с нечетными номерами, т.е.

$y^e_k = y_{2k}, \quad k = 0, 1, \ldots, n/2 - 1,$
$y^o_k = y_{2k+1}, \quad k = 0, 1, \ldots, n/2 - 1.$

ДПФ последовательностей $\{y^e\}$ и $\{y^o\}$ обозначим соответственно через $\{\widehat{y^e}\}$ и $\{\widehat{y^o}\}$. Несложно показать, что n-точечное ДПФ $\{\widehat{y}\}$ может быть выражено через два $n/2$-точечных ДПФ $\{\widehat{y^e}\}$ и $\{\widehat{y^o}\}$ следующим образом

$\widehat{y}_k = \widehat{y^e}_k + W_n^k \; \widehat{y^o}_k,$
$\widehat{y}_{n/2+k} = \widehat{y^e}_k - W_n^k \; \widehat{y^o}_k,$
при $k = 0, 1, \ldots, n/2 - 1.$

слайд 52

На этом шаге требуется $n/2$ умножений. Применим данную схему к $n/2$-периодическим последовательностям $\{y^e\}$ и $\{y^o\}$ для расчета ДПФ $\{\widehat{y^e}\}$ и $\{\widehat{y^o}\}$. Число умножений для них также сократится вдвое, т.е. для каждой последовательности будет произведено $n/4$ умножений, что в сумме даст $n/2$ умножений. Продолжая эти действия, мы каждый раз будем увеличивать число последовательностей вдвое и сокращать число умножений вдвое. Так как n является степенью двойки, то максимальное число делений последовательностей на 2 равно $\log_2 n$, при котором будут получены одноточечные последовательности.

Легко подсчитать, что общее число умножений будет равно $\frac{n}{2} \log_2 n$, что при больших n значительно меньше числа прямых вычислений, требующих n^2 умножений.

У этого алгоритма существуют обобщения на случай, когда n не является степенью 2.

слайд 53

Преобразование Фурье разлагает сигналы на ортогональные базисные функции (синусы и косинусы), определяя их частотные составляющие. Но этот метод не может локализовать частотные компоненты во времени, а только анализирует их наличие и величину. Преобразование Фурье не позволяет, например, определить, присутствовала ли данная частота в сигнале всегда, или она появилась в какой-то момент.

Рассмотрим два примера.

Пример 1. Пусть задана функция
$f(t) = \sin(2\pi \cdot 70t) + \sin(2\pi \cdot 250t).$

На рисунке изображено ее ДПФ. Шаг дискретизации был выбран равным 0.001.

слайд 54

Пример 2. Пусть задана функция

$$f(t) = \begin{cases} \sin(2\pi \cdot 70t), \text{ при } t \leqslant 0,3 \\ \sin(2\pi \cdot 250t), \text{ при } t > 0,3. \end{cases}$$

На рисунке изображено ее ДПФ.

Сравните с примером 1: в обоих случаях преобразование Фурье дает два пика в точках $2\pi \cdot 70$ и $2\pi \cdot 250$ и нельзя сказать в какой момент времени присутствовала та или иная частота.

слайд 55

Принцип неопределенности Гейзенберга

Идеально было бы использовать базисные функции с компактным носителем во временно́й области, чьи преобразования Фурье при этом имели бы компактный носитель в частотной области. Однако нельзя определить частотно-временны́е характеристики сигнала с какой угодно точностью. Принцип неопределенности Гейзенберга утверждает, что одновременной локализации по времени и частоте добиться невозможно.

Определим меры ширины носителей функции и ее преобразования Фурье, понимая под этим термином размер области, в которой функция (ее преобразование Фурье) заметно отлична от нуля. Пусть

$$x_0 = \frac{\int\limits_{-\infty}^{\infty} x|f(x)|^2 dx}{\int\limits_{-\infty}^{\infty} |f(x)|^2 dx} \quad \text{и} \quad \omega_0 = \frac{\int\limits_{-\infty}^{\infty} \omega\left|\widehat{f}(\omega)\right|^2 d\omega}{\int\limits_{-\infty}^{\infty} \left|\widehat{f}(\omega)\right|^2 d\omega}.$$

Величины x_0 и ω_0 (мы предполагаем, что они конечны) представляют собой взвешенные средние значения f и $\widehat{f}$ соответственно. Говорят, что сигнал f сконцентрирован вокруг точки (x_0, ω_0) в частотно-временно́й плоскости. Тогда величины Δ_x и Δ_ω, определяемые как

$$\Delta_x^2 = \frac{\int\limits_{-\infty}^{\infty} (x-x_0)^2|f(x)|^2 dx}{\int\limits_{-\infty}^{\infty} |f(x)|^2 dx} \text{ и } \Delta_\omega^2 = \frac{\int\limits_{-\infty}^{\infty} (\omega-\omega_0)^2\left|\widehat{f}(\omega)\right|^2 d\omega}{\int\limits_{-\infty}^{\infty} \left|\widehat{f}(\omega)\right|^2 d\omega},$$

являются мерами разброса функции сигнала f и ее преобразования Фурье $\widehat{f}$ относительно центра (x_0, ω_0). Эти величины и представляют собой меры ширины носителей f и $\widehat{f}$. Принцип неопределенности Гейзенберга утверждает, что $\Delta_x^2\Delta_\omega^2 \geq 1/4$ (см. [?]), причем функция Гаусса – единственная функция, для которой в этом неравенстве выполняется равенство.

слайд 56

Недостаток Фурье-анализа, обусловленный неспособностью осуществлять временну́ю локализацию сингулярностей сигналов, может быть частично устранен путем введения в преобразование так называемого окна $g(t)$ – движущейся функции, имеющей компактный носитель. Использование оконной функции позволяет представить результат анализа – образ Фурье – в виде функции двух переменных, а именно – частоты ω и времени u положения окна.

Пусть функция g является вещественной, четной и $\|g\| = 1$.

Определение. Пусть функция $f \in L^2(R)$, ее оконным преобразованием Фурье называется

$$Sf(u,\omega) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} f(t)g(t-u)e^{-i\omega t}dt.$$

Это преобразование также называют кратковременным преобразованием Фурье, поскольку интеграл Фурье берется по окрестности точки $t = u$, определяемой сомножителем $g(t-u)$.

Спектрограммой назовем функцию

$$|Sf(u,\omega)|^2 = \frac{1}{2\pi} \left| \int\limits_{-\infty}^{\infty} f(t)g(t-u)e^{-i\omega t}dt \right|^2.$$

Спектрограмма измеряет энергию сигнала f в частотно-временно́й окрестности точки (u,ω), называемой прямоугольником Гейзенберга.

Функция f может быть восстановлена по ее оконному преобразованию Фурье $Sf(u,\omega)$. В следующей теореме дается формула реконструкции сигнала f и доказывается, что оконное преобразование Фурье сохраняет энергию сигнала.

Теорема. Если $f \in L^2(R)$, то

$$f(t) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} Sf(u,\omega)g(t-u)e^{i\omega t}d\omega du$$

и

$$\int\limits_{-\infty}^{\infty} |f(t)|^2 dt = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} |Sf(u,\omega)|^2 d\omega du.$$ (формула сохранения энергии)

Одним из примеров оконного преобразования Фурье является преобразование Габора, в котором роль оконной функции играет гауссиан. Преобразование Габора является одним из лучших оконных преобразований Фурье, поскольку обеспечивает наилучшую частотно-временну́ю локализацию сигнала.

Оконные преобразования позволяют проанализировать либо высокие частоты в коротком окне времени, либо низкочастотную компоненту, но не оба колебания одновременно. Ширина окна остается постоянной для всех частот. Необходим подход, в котором для различных диапазонов частот использовались временны́е окна различной длительности. Это обеспечивается с помощью специальных базисных функций, которые Морле назвал вейвлетами (wavelets) – всплесками, маленькими волнами.

слайд 57

Вейвлет-анализ

Фурье-анализ дает информацию только о частотном спектре сигнала. Оконное преобразование Фурье представляет собой более удобный инструмент. Весь временной интервал разбивается на небольшие равные интервалы, которые анализируются отдельно с помощью преобразования Фурье. Однако использование равных интервалов затрудняет одновременное обнаружение коротких высокочастотных и более длительных низкочастотных участков сигнала. Вейвлеты позволяют получать временну́ю и частотную информацию, сужая окно для выделения коротких высокочастотных участков или расширяя его для анализа длительных низкочастотных колебаний. Впервые на практике вейвлеты были применены в геофизике для разведки нефтяных и минеральных месторождений. Однако математики разработали аппарат вейвлет-анализа для решения абстрактных проблем еще за 20 лет до этого.

Наибольшее развитие получила практика применения вейвлетов для решения задач сжатия и обработки сигналов и изображений, являющихся нестационарными по своей природе.

слайд 58

Вейвлеты Хаара

В вейвлет-анализе основную роль играют две функции: масштабирующая функция ϕ и вейлет-функция ψ. Эти две функции генерируют базисы, которые используются для разложения и реконструкции сигнала. Самый простой пример вейвлет-анализа основан на масштабирующей функции Хаара (см. рис):

$$\phi(x) = \begin{cases} 1, & \text{если } 0 \le x < 1 \\ 0, & \text{иначе.} \end{cases}$$

слайд 59

Функция $\phi(x-k)$ представляет собой копию функции $\phi(x)$, сдвинутую вправо на k единиц (если $k > 0$). Обозначим через V_0 пространство всех функций вида $\sum\limits_{k\in Z} a_k\phi(x-k), \quad a_k \in R$, где сумма содержит лишь конечное число слагаемых.

Пространство V_0 состоит из всех кусочно-постоянных функций, которые могут иметь разрывы только в целых точках. Рассмотрим теперь для любого целого неотрицательного j пространство V_j всех функций вида $\sum\limits_{k\in Z} a_k\phi(2^jx-k), \quad a_k \in R$.

Нетрудно видеть, что V_j состоит из кусочно постоянных функций, которые могут иметь разрывы только в точках $\{0, \pm 1/2^j, \pm 2/2^j, \pm 3/2^j, \ldots\}$.

Любая функция из V_0 содержится в V_1, т.е. $V_0 \subset V_1$. Также выполнено $V_1 \subset V_2$ и, продолжая эту цепочку:

$V_0 \subset V_1 \subset \cdots \subset V_j \subset \cdots$.

Эта вложенность строгая. V_j содержит информацию до деталей размера 2^{-j}. Назовем этот минимальный размер деталей разрешением. При увеличении j разрешение становится лучше. Тот факт, что $V_j \subset V_{j+1}$, означает, что при переходе к следующему уровню разрешения информация не теряется.

График функции $\phi(2^jx-k)$ представляет собой ступеньку шириной $1/2^j$. С помощью таких ступенек можно аппроксимировать анализируемый сигнал, поэтому хотелось бы иметь эффективный алгоритм разложения сигнала по его V_j-компонентам. Для этого требуется построить ортонормированный базис V_j (используя скалярное произведение в L^2).

Лемма 1. $f(x) \in V_0 \Leftrightarrow f(2^jx) \in V_j$, $f(x) \in V_j \Leftrightarrow f(2^{-j}x) \in V_0$.

Лемма 2. $\{2^{j/2}\phi(2^jx-k),\ k \in Z\}$ – ортонормированный базис в V_j $\left(\int\limits_{-\infty}^{\infty} |\phi(2^jx)|^2\,dx = 1/2^j\right)$.

слайд 60

Для анализа и обработки нужно уметь выделять компоненты сигнала, принадлежащие V_j, но не принадлежащие V_{j-1}. Здесь используется вейвлет-функция ψ. Идея заключается в разложении V_j на ортогональную сумму V_{j-1} и его дополнения. Начнем с $j = 1$ и построения ортогонального дополнения к V_0. Поскольку V_0 порождается сдвигами функции ϕ, разумно предположить, что ортогональное дополнение к V_0 порождается сдвигами некоторой функции ψ. Функция ψ должна удовлетворять двум требованиям:

1. ψ должна принадлежать V_1, т.е. $\psi(x) = \sum\limits_{l \in Z} a_l \phi(2x - l)$ для некоторых $a_l \in R$.

2. ψ должна быть ортогональна V_0, т.е. $\int\limits_{-\infty}^{\infty} \psi(x)\phi(x - k)dx = 0$ для всех целых k.

Функция, удовлетворяющая этим требованиям, – это вейвлет-функция Хаара (или просто вейвлет Хаара):

$\psi(x) = \phi(2x) - \phi(2x - 1)$.

График вейвлет-функции Хаара изображен на рисунке.

слайд 61

Функция из V_1 ортогональна V_0 тогда и только тогда, когда она представляет собой линейную комбинацию функций вида $\psi(x-k)$, где $k \in Z$. Обозначим множество всех таких функций через W_0. Тогда $V_1 = V_0 \oplus W_0$. Таким же образом можно установить, что справедлива следующая теорема.

Теорема 1. Пусть W_j – пространство всех конечных линейных комбинаций вида $\sum\limits_k a_k \psi(2^j x - k)$, $\quad a_k \in R$, тогда W_j представляет собой ортогональное дополнение к V_j в V_{j+1}, и $V_{j+1} = V_j \oplus W_j$.

Последовательно разлагая V_j, V_{j-1} и т.д., получаем $V_j = W_{j-1} \oplus V_{j-1} = W_{j-1} \oplus W_{j-2} \oplus V_{j-1} = \cdots = W_{j-1} \oplus W_{j-2} \oplus \cdots \oplus W_0 \oplus V_0$, поэтому любая функция $f \in V_j$ может быть представлена в виде

$f = w_{j-1} + w_{j-2} + \ldots + w_0 + f_0$,

где $w_l \in W_l$, $0 \le l < j$, а $f_0 \in V_0$. Компоненты w_l представляют собой "пики" функции f ширины $1/2^{l+1}$, которые не могут быть представлены в виде линейных комбинаций пиков другой ширины. Если устремить j к бесконечности, то можно показать, что справедлива следующая теорема.

Теорема 2. Пространство $L^2(R)$ может быть разложено в бесконечную ортогональную сумму

$L^2(R) = V_0 \oplus W_0 \oplus W_1 \oplus \cdots$,

т.е. любую функцию $f \in L^2(R)$ можно единственным образом представить в виде $f = f_0 + \sum\limits_{j=0}^{\infty} w_j$, где $f_0 \in V_0$, а $w_j \in W_j$. Ряд сходится в смысле L^2.

Идея доказательства этой теоремы заключается в том, что любую функцию из $L^2(R)$ можно приблизить непрерывными функциями, и что любую непрерывную функцию можно сколь угодно близко приблизить ступенчатыми функциями.

слайд 62

Для обработки или анализа сигнала нужно разложить его на компоненты по пространствам V_0 и $W_0, \ldots, W_{j-1}$.

Сначала для достаточно большого j необходимо аппроксимировать функцию f ступенчатой функцией $f_j \in V_j$

$f_j(x) = \sum\limits_{k \in Z} a_{j,k} \phi(2^j x - k)$.

В качестве коэффициентов $a_{j,k}$ берутся дискретные отсчеты функции f, т.е. $a_{j,k} = f(k/2^j)$.

Затем нужно разложить f_j по компонентам:

$f_j = f_0 + w_0 + \ldots + w_{j-2} + w_{j-1}$.

Компоненты w_l представляют собой пики шириной $1/2^{l+1}$. Для достаточно большого l эти пики становятся очень тонкими, и можно считать, что они представляют собой шум. Чтобы отфильтровать такой шум, нужно обнулить эти компоненты.

Нетрудно убедиться, что справедливы следующие соотношения (это видно из графиков функций).

Утверждение 3.

$\phi(2^j x) = (\phi(2^{j-1} x) + \psi(2^{j-1} x))/2$,

$\phi(2^j x - 1) = (\phi(2^{j-1} x) - \psi(2^{j-1} x))/2$.

слайд 63

Воспользуемся этим утверждением для построения алгоритма разложения.

Разделим сумму $f_j(x) = \sum\limits_{k \in Z} a_{j,k} \phi(2^j x - k)$ на два слагаемых:

$f_j(x) = \sum\limits_{k \in Z} a_{j,2k}\phi(2^j x - 2k) + \sum\limits_{k \in Z} a_{j,2k+1}\phi(2^j x - 2k - 1)$.

Из утверждения 3 следует, что

$\phi(2^j x - 2k) = (\phi(2^{j-1}x - k) + \psi(2^{j-1}x - k))/2$,

$\phi(2^j x - 2k - 1) = (\phi(2^{j-1}x - k) - \psi(2^{j-1}x - k))/2$.

Следовательно,

$f_j(x) = \sum\limits_{k \in Z} a_{j,2k}\frac{\phi(2^{j-1}x-k)+\psi(2^{j-1}x-k)}{2} + \sum\limits_{k \in Z} a_{j,2k+1}\frac{\phi(2^{j-1}x-k)-\psi(2^{j-1}x-k)}{2} =$

$= \sum\limits_{k \in Z}\left(\frac{a_{j,2k}+a_{j,2k+1}}{2}\right)\phi(2^{j-1}x - k) + \sum\limits_{k \in Z}\left(\frac{a_{j,2k}-a_{j,2k+1}}{2}\right)\psi(2^{j-1}x - k) =$

$= f_{j-1} + w_{j-1}$, где $f_{j-1} \in V_{j-1}$, а $w_{j-1} \in W_{j-1}$.

Таким образом, справедлива следующая теорема.

Теорема 3. Пусть $f_j(x) = \sum\limits_{k \in Z} a_{j,k}\phi(2^j x - k) \in V_j$, тогда f_j можно разложить в $f_j = f_{j-1} + w_{j-1}$:

$f_{j-1}(x) = \sum\limits_{k \in Z} a_{j-1,k}\phi(2^{j-1}x - k) \in V_{j-1}$,

$w_{j-1}(x) = \sum\limits_{k \in Z} b_{j-1,k}\psi(2^{j-1}x - k) \in W_{j-1}$,

где $a_{j-1,k} = \frac{a_{j,2k}+a_{j,2k+1}}{2}$, $b_{j-1,k} = \frac{a_{j,2k}-a_{j,2k+1}}{2}$.

Эту процедуру можно повторить, разложив f_{j-1} на f_{j-2} и w_{j-2}. Продолжая таким же образом, мы получим представление

$f_j = f_0 + w_0 + \ldots + w_{j-2} + w_{j-1}$.

слайд 64

Алгоритм реконструкции

После разложения f_j на компоненты можно приступать к обработке сигнала. Если цель обработки – подавление шума, то нужно обнулить компоненты w_l, соответствующие шумовым частотам. Если цель – компрессия данных, нужно обнулить компоненты w_l, коэффициенты которых близки к нулю. При этом сигнал изменится не сильно. В любом случае, после модификации коэффициентов необходимо реконструировать сигнал, т.е. из сигнала в форме

$f(x) = f_0(x) + w_0(x) + \ldots + w_{j-2}(x) + w_{j-1}(x)$, где

$f_0(x) = \sum\limits_{k \in Z} a_{0,k}\phi(x - k)$, $w_l(x) = \sum\limits_{k \in Z} b_{l,k}\psi(2^l x - k)$,

получить его представление в терминах базиса $\phi(2^j x - k)$ пространства V_j:

$f(x) = \sum\limits_{k \in Z} a_{j,k}\phi(2^j x - k)$.

Утверждение 4.
$\phi(2^{j-1}x) = \phi(2^j x) + \phi(2^j x - 1)$,
$\psi(2^{j-1}x) = \phi(2^j x) - \phi(2^j x - 1)$.

слайд 65

Имеем
$f_0(x) = \sum\limits_{k\in Z} a_{0,k}\phi(x-k) = \sum\limits_{k\in Z} a_{0,k}\phi(2x-2k) + a_{0,k}\phi(2x-2k-1)$,
т.е. $f_0(x) = \sum\limits_{k\in Z} a'_{1,k}\phi(2x-k)$, где $a'_{1,2k} = a_{0,k}$ и $a'_{1,2k+1} = a_{0,k}$.

Аналогично, $w_0(x) = \sum\limits_{k\in Z} b_{0,k}\psi(x-k)$ можно представить в виде
$w_0(x) = \sum\limits_{k\in Z} b'_{1,k}\phi(2x-k)$, где $b'_{1,2k} = b_{0,k}$ и $b'_{1,2k+1} = -b_{0,k}$.

Следовательно,
$f_0(x) + w_0(x) = \sum\limits_{k\in Z} a_{1,k}\phi(2x-k)$, где

$$a_{1,k} = a'_{1,k} + b'_{1,k} = \begin{cases} a_{0,k} + b_{0,k}, & \text{если } k \text{ четно} \\ a_{0,k} - b_{0,k}, & \text{если } k \text{ нечетно.} \end{cases}$$

Затем к сумме таким же способом прибавляется $w_1(x) = \sum\limits_{k\in Z} b_{1,k}\psi(x-k)$. Продолжая эту процедуру, получаем следующий алгоритм реконструкции.

Теорема 4. Пусть $f(x) = f_0(x) + w_0(x) + \ldots + w_{j-2}(x) + w_{j-1}(x)$, где
$f_0(x) = \sum\limits_{k\in Z} a_{0,k}\phi(x-k), \quad w_l(x) = \sum\limits_{k\in Z} b_{l,k}\psi(2^l x - k)$.

Тогда $f(x) = \sum\limits_{k\in Z} a_{j,k}\phi(2^j x - k)$, где $a_{l,k}$ определяются рекурсивно для $l = 1$, затем для $l = 2$ и т.д. до $l = j$ по следующим формулам:

$$a_{l,k} = \begin{cases} a_{l-1,k} + b_{l-1,k}, & \text{если } k \text{ четно} \\ a_{l-1,k} - b_{l-1,k}, & \text{если } k \text{ нечетно.} \end{cases}$$

слайд 66

Процедуры разложения и реконструкции в виде фильтров

Определение. Для последовательности $x = (\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots)$ сгущающей последовательностью Dx (downsample) называется последовательность

$Dx = (\ldots, x_{-2}, x_0, x_2, \ldots)$, т.е. $(Dx)_k = x_{2k}$ для всех $k \in Z$. Оператор D назовем downsample-оператором.

Определение. Для последовательности $x = (\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots)$ прорежающей последовательностью Ux (upsample) называется последовательность

$Ux = (\ldots, x_{-2}, 0, x_{-1}, 0, x_0, 0, x_1, 0, x_2, \ldots)$,

т.е. $(Ux)_k = \begin{cases} 0, & \text{если k нечетное} \\ x_{k/2}, & \text{если k четное.} \end{cases}$

Оператор U назовем upsample-оператором.

слайд 67

Разложение

Введем два дискретных фильтра H и L через соответствующие импульсные отклики. Пусть

$$h = (\ldots, 0, \ldots, 0, -\tfrac{1}{2}, \tfrac{1}{2}, 0, \ldots, 0, \ldots),$$

где $h_{-1} = -\frac{1}{2}, h_0 = \frac{1}{2}$,

$$l = (\ldots, 0, \ldots, 0, \tfrac{1}{2}, \tfrac{1}{2}, 0, \ldots, 0, \ldots),$$

где $l_{-1} = \frac{1}{2},\ l_0 = \frac{1}{2}$.

Определение. Высокочастотным фильтром H (high-pass filter) назовем оператор свертки, определенный по правилу

$$Hx = h * x, \text{ тем самым } (Hx)_k = \tfrac{1}{2}(x_k - x_{k+1}).$$

Определение. Низкочастотным фильтром L (low-pass filter) назовем оператор свертки, определенный по правилу

$$Lx = l * x, \text{ тем самым } (Lx)_k = \tfrac{1}{2}(x_k + x_{k+1}).$$

слайд 68

Коэффициенты $a_{j-1,k}$ и $b_{j-1,k}$ определяются через набор коэффициентов a^j уровня j по следующим правилам:

$$a_{j-1,k} = (DLa^j)_k,\ b_{j-1,k} = (DHa^j)_k.$$

Здесь DL и DH означает, что сначала последовательность коэффициентов a^j обрабатывается соответственно фильтром L (low-pass) или H (high-pass), а затем применяется downsample-оператор D. На схеме этот оператор обозначается $2\downarrow$.

После определения всех коэффициентов разложения сигнал представлен в виде суммы вкладов на разных масштабах. Теперь можно обрабатывать сигнал, изменяя коэффициенты $b_{j,k}$ соответственно желаемым целям обработки. Например, чтобы удалить (отфильтровать) из ряда высокие частоты, следует положить все $b_{j,k} = 0$ для j, превосходящих некоторое определенное число (порог). Если к тому же требуется отфильтровка частот на определенных временны́х интервалах, то соответствующим образом выбирается множество значений индекса k.

Если же мы хотим сжать данные, то для этого полагаются равными нулю те коэффициенты $b_{j,k}$, которые по абсолютному значению не превосходят некоторого выбранного порога.

слайд 69

введем два дискретных фильтра $\widetilde{H}$ и $\widetilde{L}$ через соответствующие импульсные отклики:

$\widetilde{h} = (\ldots, 0, \ldots, 0, 1, -1, 0, \ldots, 0, \ldots)$, где $\widetilde{h}_0 = 1, \widetilde{h}_1 = -1$,

и

$\widetilde{l} = (\ldots, 0, \ldots, 0, 1, 1, 0, \ldots, 0, \ldots)$, где $\widetilde{l}_0 = 1$, $\widetilde{l}_1 = 1$.

Таким образом,

$(\widetilde{H}x)_k = (\widetilde{h} * x)_k = x_k - x_{k-1}$,

$(\widetilde{L}x)_k = (\widetilde{l} * x)_k = x_k + x_{k-1}$.

слайд 70

Теперь, в соответствии с алгоритмом реконструкции можно вычислить коэффициенты a^j по рекуррентной формуле

$a^j = \widetilde{L}Ua^{j-1} + \widetilde{H}Ub^{j-1} \quad$ для $j = 1, \ldots, J$.

Здесь $\widetilde{L}U$ и $\widetilde{H}U$ означает, что сначала к последовательностям коэффициентов применяется upsample-оператор U (на схеме он обозначается $2\uparrow$), а затем Ua^{j-1} обрабатывается фильтром $\widetilde{L}$ (low-pass), а Ub^{j-1} фильтром $\widetilde{H}$ (high-pass).

Сначала для $j = 1$ рассчитывается вектор коэффициентов a^1 по известным векторам a^0 и b^0. Затем для $j = 2$ рассчитывается вектор a^2 по вычисленному a^1 и известному b^1 и так далее. В итоге при $j = J$ (на максимальном уровне) коэффициенты $a_{J,k}$ будут представлять собой значения сигнала в моменты времени $x = k/2^J$. Конечно же, вновь полученные коэффициенты будут отличаться от исходных $a_{J,k}$.

Итак, алгоритм обработки сигнала состоит из следующих шагов:

Дискретизация $\to$ разложение $\to$ анализ (обработка) $\to$ реконструкция.

слайд 71

Недостатком функций Хаара является их разрывность, вследствие чего они позволяют получить лишь грубую аппроксимацию непрерывных сигналов. В этом параграфе мы опишем схему разложения и реконструкции сигнала, опирающуюся на более общий вид функций ϕ и ψ. Эта схема, разработанная С. Малла, называется кратномасштабным анализом.

Начнем с общего определения кратномасштабного анализа. Пусть V_j, $j = \ldots, -2, -1, 0, 1, 2, \ldots$, – последовательность подпространств функций из $L^2(R)$. Эта последовательность называется кратномасштабным анализом с масштабирующей функцией ϕ, если выполнены следующие условия:

1. $V_j \subset V_{j+1}$.

2. Замыкание множества $\bigcup\limits_{j=-\infty}^{\infty} V_j$ совпадает с $L^2(R)$.

3. Замыкание множества $\bigcap\limits_{j=-\infty}^{\infty} V_j$ пусто.

4. $f(x) \in V_j \Longleftrightarrow f(2^{-j}x) \in V_0$.

5. $\phi(x) \in V_0$, и последовательность функций $\{\phi(x-k), k \in Z\}$ образует ортонормированный базис в V_0 (в смысле скалярного произведения в $L^2(R)$).

Пространства V_j называются аппроксимационными пространствами. Системе аппроксимационных пространств могут соответствовать различные масштабирующие функции ϕ. Возможно, самыми полезными масштабирющими функциями являются функции с компактным (или конечным) носителем. Примером такой функции является масштабирующая функция Хаара. Позднее мы приведем примеры масштабирующих функций, которые не только имеют компактный носитель, но и непрерывны.

слайд 72

Пример 1. Кратноамасштабный анализ, ассоциированный с масштабирующей функцией Хаара $\phi(x)$, является, пожалуй, одним из самых простых (в вычислительном смысле) примеров кратномасштабного анализа. Однако, как уже отмечалось выше, он плохо подходит для анализа гладких сигналов. Кроме того, преобразование Фурье $\widehat{\phi}(\omega)$ убывает довольно медленно при $|\omega| \to \infty$ (со скоростью $|\omega|^{-1}$), поэтому кратномасштабный анализ Хаара, предлагающий хорошую локализацию во временно́й области, не обеспечивает ее в частотной области.

Пример 2. Кратномасштабный анализ Шеннона. Для $j \in Z$ пространство V_j представляет собой пространство всех функций $f \in L^2(R)$, преобразование Фурье которых равно нулю вне интервала $[-2^j\pi, 2^j\pi]$.

Масштабирующая функция Шеннона равна $\phi(x) = \sin(\pi x)/(\pi x)$, при $x = 0$ полагают $\phi(0) = 1$. Используя формулу Планшереля и теорему Шеннона-Уиттакера, можно показать, что последовательность $\{\phi(x - k),\ k \in Z\}$ образует ортонормированный базис в V_0. Также легко проверяются другие свойства кратномасштабного анализа.

Поскольку преобразование Фурье любой функции $f_j(x)$ из пространства V_j обычно имеет разрывы в точках $\pm 2^j\pi$, сама функция $f_j(x)$ убывает со скоростью $|x|^{-1}$ при $|x| \to \infty$, даже если сигнал f, который она аппроксимирует, имеет конечный носитель. Следовательно, кратномасштабный анализ Шеннона, предлагая неплохую локализацию в частотной области, не обеспечивает ее во временно́й области.

слайд 73

Пример 3. Полиномиальная аппроксимация. Сплайны позволяют построить гладкую аппроксимацию сигнала с довольно быстрым асимптотическим убыванием. Аппроксимационное пространство V_j представляет собой множество функций из $L^2(R)$, непрерывно дифференцируемых $m-1$ раз ($m \geq 1$) и равных многочлену степени m на любом отрезке $[n2^{-j}, (n+1)2^{-j}]$, $n \in Z$. При $m = 0$ получаем кратномасштабный анализ Хаара. При $m = 1$ функции из V_j кусочно линейны и непрерывны. Аналогом масштабирующей функции является функция $\phi(x)$, чье преобразование Фурье равно

$\widehat{\phi}(\omega) = \left(\frac{\sin(\omega/2)}{\omega/2}\right)^{m+1} e^{-i\delta\omega/2}$,

где $\delta = 1$, если m четно (при этом ϕ имеет носитель, симметричный относительно $x = 1/2$), и $\delta = 0$, если m нечетно (при этом ϕ симметрична относительно $x = 0$). Последовательность $\{\phi(x - k),\ k \in Z\}$ не ортонормирована, однако образует устойчивый базис (базис Рисса).

Пример 4. Кратномасштабный анализ Добеши. Ингрид Добеши (Ingrid Daubechies) описала класс масштабирующих функций и ассоциированных с ними вейвлет-функций, имеющих компактный носитель и заданное количество непрерывных производных. В этот класс входят и функции Хаара (это функции Добеши первого порядка).

Вейвлет-функции Добеши классифицируются по числу N нулевых моментов, т.е.

$$\int\limits_{-\infty}^{\infty} x^k \psi_N(x) dx = 0, \quad k = 0, \cdots, N-1.$$

Нулевые моменты очень важны во многих приложениях вейвлет-анализа – компрессии сигнала, подавлении шума, локализации особенностей и т.д. Подробнее об этом будет сказано в дальнейшем.

Гладкость $\phi_N(x)$ и $\psi_N(x)$ растет с ростом количества нулевых моментов. При $N = 1$ получаются масштабирующая функция и вейвлет-функция Хаара, которые разрывны. При $N = 2$ функции $\phi_N(x)$ и $\psi_N(x)$ непрерывны, но не имеют непрерывных производных. Для больших N количество непрерывных производных у $\phi_N(x)$ и $\psi_N(x)$ примерно равно $N/5$.

На рисунках изображены функции Добеши 2-го, 4-го и 20-го порядков соответственно. Можно заметить, что с ростом порядка увеличивается гладкость этих функций.

слайд 74

Масштабирующая функция и вейвлет Добеши для $N = 2$.

слайд 75

Масштабирующая функция и вейвлет Добеши для $N = 4$.

слайд 76

Масштабирующая функция и вейвлет Добеши для $N = 20$.

слайд 77

Теорема 1. Пусть $\{V_j,\ j \in Z\}$ – кратномасштабный анализ с масштабирующей функцией ϕ, тогда для любого $j \in Z$ последовательность функций $\{\phi_{j,k}(x)=2^{j/2}\phi(2^jx-k),\ k\in Z\}$ является ортонормированным базисом в V_j.

Теорема 2. (масштабное соотношение) Пусть $\{V_j,\ j\in Z\}$ – кратномасштабный анализ с масштабирующей функцией ϕ, тогда справедливо следующее соотношение:

$\phi(x)=\sum\limits_{k\in Z}p_k\phi(2x-k)$, где $p_k=2\int\limits_{-\infty}^{\infty}\phi(x)\overline{\phi(2x-k)}dx$.

Более того,

$\phi(2^{j-1}x-l)=\sum\limits_{k\in Z}p_{k-2l}\phi(2^jx-k)$ или $\phi_{j-1,l}=2^{-1/2}\sum\limits_{k\in Z}p_{k-2l}\phi_{j,k}$.

Для КМА Хаара $p_0=p_1=1$, $p_k=0$ при $k\neq 0,1$.

Для КМА Добеши $(N=2)$

$p_0=\frac{1+\sqrt{3}}{4}$, $p_1=\frac{3+\sqrt{3}}{4}$, $p_2=\frac{3-\sqrt{3}}{4}$, $p_3=\frac{1-\sqrt{3}}{4}$, $p_k=0$ при $k\neq 0,1,2,3$.

слайд 78

Для того чтобы описать алгоритм разложения сигнала в общем случае, необходимо представить V_{j+1} в виде ортогональной суммы V_j и W_j (как это было сделано в случае функций Хаара). Следующая теорема позволяет получить вейвлет-функцию $\psi(x)$, сдвиги которой порождают пространство W_j.

Теорема 3. Пусть $\{V_j,\ j\in Z\}$ – кратномасштабный анализ с масштабирующей функцией ϕ, и пусть пространство W_j порождается семейством функций $\{\psi(2^jx-k),\ k\in Z\}$, где

$\psi(x)=\sum\limits_{k\in Z}(-1)^k\overline{p_{1-k}}\phi(2x-k)$,

а p_k те же, что в теореме 2, тогда $V_{j+1}=V_j\oplus W_j$. Кроме того, последовательность функций $\{\psi_{j,k}(x)=2^{j/2}\psi(2^jx-k),\ k\in Z\}$ образует ортонормированный базис в W_j. Функция ψ называется вейвлет-функцией или просто вейвлетом.

Замечание. Для масштабирующей функции Хаара коэффициенты p_0 и p_1 равны 1. По теореме 3 $\psi(x)=\phi(2x)-\phi(2x-1)$, что согласуется с определением вейвлет-функции Хаара.

Замечание. Простым следствием теоремы 3 является следующее разложение функций $\psi_{j,k}(x)=2^{j/2}\psi(2^jx-k)$:

$\psi_{j,k}=2^{-1/2}\sum\limits_{l\in Z}(-1)^l\overline{p_{1-l+2k}}\phi_{j+1,l}$.

слайд 79

Последовательно разлагая V_j, V_{j-1} и т.д., получаем

$$V_j = W_{j-1} \oplus W_{j-2} \oplus \ldots \oplus W_0 \oplus V_0.$$

Поскольку пространства V_j определены и для $j < 0$, эту процедуру можно продолжить:

$$V_j = W_{j-1} \oplus W_{j-2} \oplus \cdots \oplus W_0 \oplus W_{-1} \oplus W_{-1} \oplus \cdots.$$

При $j \to \infty$ получаем следующее утверждение.

Теорема 4. Пусть $\{V_j,\ j \in Z\}$ – кратномасштабный анализ с масштабирующей функцией ϕ. W_j – ортогональное дополнение к V_j до V_{j+1}. Пространство $L^2(R)$ может быть разложено в бесконечную ортогональную сумму

$$L^2(R) = \ldots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \ldots,$$

т.е. любую функцию $f \in L^2(R)$ можно единственным образом представить в виде

$$f = \sum_{j=-\infty}^{\infty} w_j,$$

где $w_j \in W_j$. Ряд сходится в смысле L^2.

слайд 80

Свойства масштабирующих функций. Достаточные условия.

Теорема 5. Пусть функция ϕ непрерывна, имеет компактный носитель и удовлетворяет следующим условиям:

1. $\int\limits_{-\infty}^{\infty} \phi(x)dx = 1$. (нормализация)

2. $\phi(x) = \sum\limits_{k \in Z} p_k \phi(2x - k)$ для конечного числа k. (масштабное соотношение)

3. $\int\limits_{-\infty}^{\infty} \phi(x-k)\phi(x-l)dx = \begin{cases} 1, & \text{если } k = l \\ 0, & \text{если } k \neq l \end{cases}$. (ортонормированность сдвигов)

Пусть V_j порождается семейством $\{\phi(2^j x - k),\ k \in Z\}$. Тогда $\{V_j,\ j \in Z\}$ – кратномасштабный анализ с масштабирующей функцией ϕ.

Ключевую роль при построении КМА играют коэффициенты p_k.

слайд 81

Масштабное соотношение можно переписать в терминах преобразования Фурье.

Утверждение 1. Масштабное соотношение $\phi(x) = \sum\limits_{k\in Z} p_k\phi(2x-k)$ эквивалентно соотношению $\widehat{\phi}(\omega) = \widehat{\phi}(\omega/2)P(e^{-i\omega/2})$, где многочлен P равен $P(z) = \frac{1}{2}\sum\limits_{k\in Z} p_k z^k$.

Утверждение 1 можно применить к $\widehat{\phi}(\omega/2)$ и получить равенство $\widehat{\phi}(\omega) = \widehat{\phi}(\omega/2^2)P(e^{-i\omega/2})P(e^{-i\omega/2^2})$.

Продолжая таким же образом, получаем для любого n
$\widehat{\phi}(\omega) = \widehat{\phi}(\omega/2^n)\prod\limits_{j=1}^{n} P(e^{-i\omega/2^j})$.

В пределе при $n \to \infty$ имеем
$\widehat{\phi}(\omega) = \widehat{\phi}(0)\prod\limits_{j=1}^{\infty} P(e^{-i\omega/2^j}) = \frac{1}{\sqrt{2\pi}}\prod\limits_{j=1}^{\infty} P(e^{-i\omega/2^j})$.

Для вейвлет-функции справедливо утверждение, аналогичное утверждению 1.

Пусть $Q(z) = -z\overline{P(-z)}$. можно показать, что $\widehat{\psi}(\omega) = \widehat{\phi}(\omega/2)Q(e^{-i\omega/2})$.

слайд 82

Необходимые условия для построения КМА

Теорема 6. Если функция ϕ удовлетворяет условию ортонормированности сдвигов и масштабному соотношению , тогда при всех $z \in C$, таких что $|z| = 1$, многочлен $P(z) = 2^{-1}\sum\limits_{k\in Z} p_k z^k$ удовлетворяет уравнению $|P(z)|^2 + |P(-z)|^2 = 1$.

Построение масштабирующей функции

Исходя из теоремы 6, для построения масштабирующей функции ϕ нужно построить многочлен $P(z)$, удовлетворяющий уравнению $|P(z)|^2 + |P(-z)|^2 = 1$ при $|z| = 1$, и определить ϕ через масштабное соотношение. Предположим, что мы построили такой многочлен $P(z)$, и предположим, что $P(1) = 1$, тогда масштабирующая функция ϕ строится с помощью следующего итерационного процесса. Возьмем масштабирующую функцию Хаара (она, очевидно, обладает свойством ортонормированности сдвигов из теоремы 6) и обозначим ее через $\phi_0(x)$. Далее итерационным образом определим для $n \geq 1$

$\phi_n(x) = \sum\limits_{k\in Z} p_k\phi_{n-1}(2x-k)$.

Следующая теорема дает условия, которым должен удовлетворять многочлен P, чтобы последовательность ϕ_n сходилась к масштабирующей функции ϕ, ассоциированной с многочленом P.

Теорема 7. Пусть многочлен $P(z) = 2^{-1} \sum\limits_{k \in Z} p_k z^k$ удовлетворяет следующим условиям:

$P(1) = 1$ и $|P(e^{-it})| > 0$ при $|t| \le \pi/2$,

$|P(z)|^2 + |P(-z)|^2 = 1$ при $|z| = 1$.

Обозначим через $\phi_0(x)$ масштабирующую функцию Хаара и определим ϕ_n с помощью описанной процедуры. Тогда при $n \to \infty$ последовательность ϕ_n сходится поточечно и в смысле L^2 к функции ϕ, удовлетворяющей условию ортонормированности сдвигов

$$\int\limits_{-\infty}^{\infty} \phi(x-k)\overline{\phi(x-l)}dx = \delta_{k,l}, \quad k, l \in Z,$$

и масштабному соотношению с коэффициентами p_k многочлена P.

То есть теорема 7 утверждает, что такая функция ϕ порождает некоторый КМА.

слайд 83

Опишем алгоритмы разложения и реконструкции аналогично тому, как это было сделано для КМА Хаара.

Разложение

Пусть $f_j \in V_j$, т.е. $f_j(x) = \sum\limits_{k \in Z} a_{j,k}\phi(2^j x - k)$. Поскольку функции $\phi_{j,k}$ образуют ортонормированный базис в V_j,

$$f_j = \sum_{k \in Z} a_{j,k}\phi_{j,k} = \sum_{k \in Z} \langle f_j, \phi_{j,k} \rangle \phi_{j,k},$$

а поскольку $V_j = V_{j-1} \oplus W_{j-1}$, мы также можем записать

$$f_j = f_{j-1} + w_{j-1} = \sum_{k \in Z} a_{j-1,k}\phi_{j-1,k} + \sum_{k \in Z} b_{j-1,k}\psi_{j-1,k} =$$
$$= \sum_{k \in Z} \langle f_j, \phi_{j-1,k} \rangle \phi_{j-1,k} + \sum_{k \in Z} \langle f_j, \psi_{j-1,k} \rangle \psi_{j-1,k},$$

Формулы разложения выражают коэффициенты во второй сумме через коэффициенты первой.

Используя масштабное соотношение, имеем

$$a_{j-1,k} = \langle f_j, \phi_{j-1,k} \rangle = 2^{-1/2} \sum_{l \in Z} \overline{p_{l-2k}} \langle f_j, \phi_{j,l} \rangle.$$

Далее, используя определение $\psi_{j,k}$ и теорему 3, получаем

$$b_{j-1,k} = \langle f_j, \psi_{j-1,k} \rangle = 2^{-1/2} \sum_{l \in Z} (-1)^l p_{1-l+2k} \langle f_j, \phi_{j,l} \rangle.$$

$a_{j-1,k}$ называются коэффициентами аппроксимации, а $b_{j-1,k}$ – коэффициентами детализации. Описанную процедуру можно повторить, разложив f_{j-1} на f_{j-2} и w_{j-2} и т.д. до $f_j = f_0 + w_1 + \cdots + w_{j-1}$.

слайд 84

После разложения f_j и последующей обработки коэффициентов, о которой уже говорилось, необходимо реконструировать сигнал. Для получения формулы реконструкции воспользуемся формулами разложения для функции $\phi_{j,k}$ (это можно сделать, поскольку $\phi_{j,k} \in V_j$).

Учитывая ортогональность функций $\phi_{j,k}$, имеем

$$\phi_{j,k} = \sum_{l \in Z} 2^{-1/2} \overline{p_{k-2l}} \phi_{j-1,l} + \sum_{l \in Z} 2^{-1/2} (-1)^k p_{1-k+2l} \psi_{j-1,l}.$$

Используя это разложение, получаем формулу реконструкции в следующем виде:

$$a_{j',k} = \langle f_j, \phi_{j',k} \rangle = \sum_{l \in Z} 2^{-1/2} \overline{p_{k-2l}} a_{j'-1,l} + \sum_{l \in Z} 2^{-1/2} (-1)^k p_{1-k+2l} b_{j'-1,l}.$$

Коэффициенты $a_{j',k}$ определяются рекурсивно для $j' = 1$, затем для $j' = 2$ и т.д. до $j' = j$. Таким образом восстанавливается сигнал в виде

$$f_j = \sum_{k \in Z} a_{j,k} \phi_{j,k}.$$

Перед разложением и последующими действиями нужно получить аппроксимацию f, т.е. f_j.

Наилучшая аппроксимация в смысле энергии – проекция f на V_j.

$$f_j(x) = P_j f(x) = \sum_{k \in Z} a_{j,k} \phi(2^j x - k), \text{ где } a_{j,k} = 2^j \int_{-\infty}^{\infty} f(x) \phi(2^j x - k) dx.$$

Утверждение 2. Пусть $\{V_j,\ j \in Z\}$ – кратномасштабный анализ с масштабирующей функцией ϕ, имеющей компактный носитель. Если $f \in L^2(R)$ непрерывна, то для достаточно больших j

$$a_{j,k} = 2^j \int_{-\infty}^{\infty} f(x) \phi(2^j x - k) dx \approx a f\left(\tfrac{k}{2^j}\right), \text{ где } a = \int_{-\infty}^{\infty} \phi(x) dx \ (= 1).$$

слайд 85

Опишем процедуры разложения и реконструкции в терминах дискретных фильтров.

Процедура разложения в виде фильтрации

Введем два дискретных фильтра H и L через соответствующие импульсные отклики. Пусть

$\{h\}:\ h_k = \frac{1}{2}(-1)^k p_{k+1},\ \{l\}:\ l_k = \frac{1}{2}\overline{p}_{-k}.$

Высокочастотный фильтр H (high-pass filter) определяется по правилу

$Hx = h * x.$

Низкочастотным фильтром L (low-pass filter) определяется по правилу

$Lx = l * x.$

слайд 86

Коэффициенты $a_{j-1,k}$ и $b_{j-1,k}$ определяются через набор коэффициентов a^j уровня j по следующим правилам:

$a_{j-1,k} = (DLa^j)_k,\ b_{j-1,k} = (DHa^j)_k.$

Здесь DL и DH означает, что сначала последовательность коэффициентов a^j обрабатывается соответственно фильтром L (low-pass) или H (high-pass), а затем применяется downsample-оператор D. На схеме этот оператор обозначается $2\downarrow$.

слайд 87

Процедура реконструкции в виде фильтрации

Введем два дискретных фильтра $\widetilde{H}$ и $\widetilde{L}$ через соответствующие импульсные отклики. Пусть

$\{\widetilde{h}\}:\ \widetilde{h}_k = (-1)^k \overline{p}_{1-k},\ \{\widetilde{l}\}:\ \widetilde{l}_k = p_k.$

Высокочастотный фильтр H (high-pass filter) определяется по правилу

$\widetilde{H}x = \widetilde{h} * x.$

Низкочастотным фильтром L (low-pass filter) определяется по правилу

$\widetilde{L}x = \widetilde{l} * x.$

слайд 88

Теперь, в соответствии с алгоритмом реконструкции можно вычислить коэффициенты a^j по рекуррентной формуле

$a^j = \widetilde{L}Ua^{j-1} + \widetilde{H}Ub^{j-1} \quad$ для $j = 1, \ldots, J.$

Здесь $\widetilde{L}U$ и $\widetilde{H}U$ означает, что сначала к последовательностям коэффициентов применяется upsample-оператор U (на схеме он обозначается

$2\uparrow$), а затем Ua^{j-1} обрабатывается фильтром $\widetilde{L}$ (low-pass), а Ub^{j-1} фильтром $\widetilde{H}$ (high-pass).

слайд 89

Краевые эффекты (особенности вычисления)

Рассмотрим следующий пример: пусть используются вейвлеты Добеши с 2 нулевыми моментами ($N = 2$) для разложения сигнала из n отсчетов.

a^j – вектор коэффициентов,

$l = \frac{1}{2}(\ldots, 0, p_3, p_2, p_1, p_0, 0, \ldots)$ (ненулевые коэффициенты при $k = -3, -2, -1, 0$).

$h = \frac{1}{2}(\ldots, 0, -p_0, p_1, -p_2, p_3, 0, \ldots)$ (ненулевые коэффициенты при $k = -1, 0, 1, 2$).

Тогда

$a_{j-1,k} = (DLa^j)_k = \frac{1}{2}(p_0 a_{j,2k} + p_1 a_{j,2k+1} + p_2 a_{j,2k+2} + p_3 a_{j,2k+3})$,

$b_{j-1,k} = (DHa^j)_k = \frac{1}{2}(p_3 a_{j,2k} - p_2 a_{j,2k+1} + p_1 a_{j,2k+2} - p_0 a_{j,2k+3})$.

Пусть, например, $n = 8$: $a^3 = \{s_0, \ldots, s_7\}$. Чтобы вычислить коэффициент $a_{2,3}$ необходимы коэффициенты $s_6, \ldots, s_9$. Но у нас нет s_9! Фильтрам необходимы отсчеты сигнала, которых нет.

слайд 90

Методы решения проблемы

1. Заполнение нулями (если сигнал длинный и его концы не имеют значения или сигнал внезапно начинается и заканчивается): $s_k =$ при $k < 0$ и $k > n - 1$.

2. Периодическое продолжение: $s_{k+n} = s_k$ (повторное использование отсчетов сигнала).

слайд 91

3. Линейная интерполяция (на обоих концах). Подходит, если сигнал не слишком зашумлен на краях.

4. Симметричное продолжение. Сигнал отражается четным образом на концах.

Нарушается гладкость функции сигнала, что, как правило, приводит к увеличению абсолютных значений соответствующих коэффициентов вейвлет-разложения (см. далее).

слайд 92

Вычислительная сложность

Обсудим вкратце, сколько операций требуется для вейвлет-разложения функции сигнала. Пусть f непрерывна и определена на отрезке $[0,1]$.

$a_{n,l} = f\left(\frac{l}{2^n}\right)$, $l = 0,\ldots,2^n-1$ – отсчеты сигнала ($N = 2^n$).

По формулам алгоритма разложения:

$a_{j-1,l} = 2^{-1/2}\sum\limits_{l\in Z}\overline{p_{k-2l}}a_{j,k}$.

$b_{j-1,l} = 2^{-1/2}\sum\limits_{l\in Z}(-1)^k p_{1-k+2l}a_{j,k}$.

Заметим, что коэффициентов a^{j-1} в два раза меньше, чем a^j, т.е. индекс l пробегает от 0 до $2^{j-1}-1$.

Пусть L – число ненулевых коэффициентов p_k (для Хаара $L = 2$, для Добеши с 2 нулевыми моментами $L = 4$). На масштабе $j-1$ нужно вычислить 2^{j-1} коэффициентов, для каждого из которых требуется $L+1$ умножений (включая умножение на $2^{-1/2}$). Такое же число умножений требуется для вычисления коэффициентов b^{j-1}.

При изменении j от n до 1 общее число умножений равно

$2(L+1)2^{n-1}+\ldots+2(L+1)2^0 = 2(L+1)(2^n-1)\approx 2(L+1)N$.

Таким образом, вычислительная сложность равна $O(N)$. Для сравнения вычислительная сложность быстрого преобразования Фурье (БПФ) равна $O(N\log N)$. Поэтому вейвлет-преобразование называют «сверхбыстрым».

Однако это сравнение нечестное, т.к. БПФ разлагает сигнал на все частотные компоненты от 0 до $\frac{N}{2}$. Вейвлет-преобразование разлагает сигнал на частотные компоненты из диапазонов от 2^{j-1} до 2^j. Следовательно, БПФ дает более детальное частотное представление.

слайд 93

Вейвлет-пакеты

Вейвлеты можно использовать для более детального частотного представления сигнала, используя вейвлет-пакеты. Опишем вкратце суть этого объекта.

Пусть $f_3 \in V_3$. После применения алгоритма разложения $f_3 = f_0 +$

$w_0 + w_1 + w_2$. На схеме изображено разложение в виде фильтрации.

Условно говоря, f_3 содержит частотные компоненты с номерами от 1 до 8. Тогда w_2 содержит частотные компоненты с номерами от 5 до 8, w_1 – с номерами 3 и 4, w_0 – с номером 2 и f_0 – с номером 1.

слайд 94

Идея вейвлет-пакетов заключается в применении high-pass и low-pass фильтров к w_2,, w_1 и w_0. Применение high-pass и low-pass фильтров к w_2 дает новые компоненты w_{21} и w_{20}, ортогональные друг к другу. Процесс итерационно повторяется до тех пор пока не получено разложение по всем частотным компонентам. Для f_3 полное разложение изображено на схеме.

Для каждой компоненты на k масштабе требуется $O(2^{n-k})$ умножений. Всего на k-м масштабе 2^k компонент. Следовательно, требуется $O(2^n)$ умножений. Общее число умножений на всех n масштабах $O(n2^n) = O(N \log N)$, т.е. такое же, как у быстрого преобразования Фурье.

слайд 95

Многомерный кратномасштабный анализ

Для описания плоских изображений, видео и трехмерных сцен требуется использовать функции нескольких переменных. Сепарабельный кратномасштабный анализ в пространствах с размерностью большей единицы получается с помощью тензорного произведения одномерных масштабирующих и вейвлет-функций. Мы рассмотрим эту процедуру для случая, когда размерность равна двум. Обобщение на любое конечное число измерений очевидно.

Пусть $\phi(x)$ и $\psi(x)$ – масштабирующая функция и вейвлет-функция, ассоциированные с одномерным кратномасштабным анализом. Определим $\Phi(x, y) = \phi(x)\phi(y)$. Последовательность функций

$\{\Phi_{j,m,n}(x, y) = 2^j \Phi(2^j x - m, 2^j y - n),\ m, n \in Z\}$

образует ортонормированный базис в декартовом произведении пространств $V_j = V_j' \bigotimes V_j'$, где V_j' – аппроксимационное пространство одномерного кратномасштабного анализа. Аппроксимацией функции $f \in L^2(R^2)$ является проекция на пространство V_j:

$f_j(x, y) = \sum\limits_{m,n \in Z} a_{j,m,n} \Phi_{j,m,n}(x, y),$

где $a_{j,m,n} = \langle f, \Phi_{j,m,n} \rangle = \langle f_j, \Phi_{j,m,n} \rangle$.

Опять же, при выполнении условий регулярности вместо точных значений $a_{j,m,n}$ можно использовать приближенные $a_{j,m,n} \approx f(m/2^j, n/2^j)$.

слайд 96

Обозначим через W_{j-1} ортогональное дополнение к V_{j-1} в V_j. Пусть $\Psi^{[1]}(x,y) = \phi(x)\psi(y)$, $\Psi^{[2]}(x,y) = \phi(y)\psi(x)$ и $\Psi^{[3]}(x,y) = \psi(x)\psi(y)$, тогда последовательность функций

$\{\Psi^{[1]}_{j-1,m,n}, \Psi^{[2]}_{j-1,m,n}, \Psi^{[3]}_{j-1,m,n},\ m,n \in Z\}$,

где $\Psi^{[l]}_{j,m,n}$ определяются так же, как в $\Phi_{j,m,n}(x,y)$, образует ортонормированный базис в пространстве W_{j-1}.

Проекция f на пространство W_{j-1} определяется выражением

$w_{j-1}(x,y) = d^{[1]}_{j-1}(x,y) + d^{[2]}_{j-1}(x,y) + d^{[3]}_{j-1}(x,y)$,

где $d^{[l]}_{j-1}(x,y) = \sum\limits_{m,n\in Z} b^{[l]}_{j-1,m,n}\Psi^{[l]}_{j-1,m,n}(x,y)$,

а $b^{[l]}_{j-1,m,n} = \langle f, \Psi^{[l]}_{j-1,m,n} \rangle = \langle f_j, \Psi^{[l]}_{j-1,m,n} \rangle$.

Компонента $d^{[1]}_{j-1}(x,y)$ описывает вертикальные детали,

$d^{[2]}_{j-1}(x,y)$ – горизонтальные детали,

а $d^{[3]}_{j-1}(x,y)$ – диагональные детали.

слайд 97

Разложение

Формулы разложения получаются аналогично формулам одномерного КМА с помощью масштабного соотношения и определения вейвлет-функции:

$$a_{j-1,m,n} = 2^{-1} \sum_{k,l\in Z} \overline{p_{k-2m}p_{l-2n}} a_{j,k,l},$$

$$b^{[1]}_{j-1,m,n} = 2^{-1} \sum_{k,l\in Z} \overline{p_{k-2m}}(-1)^l p_{1-l+2n} a_{j,k,l},$$

$$b^{[2]}_{j-1,m,n} = 2^{-1} \sum_{k,l\in Z} (-1)^k p_{1-k+2m}\overline{p_{l-2n}} a_{j,k,l},$$

$$b^{[3]}_{j-1,m,n} = 2^{-1} \sum_{k,l\in Z} (-1)^{k+l} p_{1-k+2m}p_{1-l+2n} a_{j,k,l}.$$

слайд 98

Реконструкция

Формула реконструкции получается из разложения

$\Phi_{j,m,n}(x,y) = 2^{-1} \sum\limits_{k,l\in Z} \overline{p_{m-2k}p_{n-2l}}\Phi_{j-1,k,l}(x,y)+$

$+2^{-1} \sum\limits_{k,l\in Z} \overline{p_{m-2k}}(-1)^n p_{1-n+2l}\Psi^{[1]}_{j-1,k,l}(x,y)+$

$+2^{-1} \sum\limits_{k,l\in Z} (-1)^m p_{1-m+2k}\overline{p_{n-2l}}\Psi^{[2]}_{j-1,k,l}(x,y)+$

$+2^{-1} \sum\limits_{k,l\in Z} (-1)^{m+n} p_{1-m+2k}p_{1-n+2l}\Psi^{[3]}_{j-1,k,l}(x,y).$

Умножая это выражение скалярно на f, получаем

$a_{j',m,n} = 2^{-1} \sum\limits_{k,l\in Z} \overline{p_{m-2k}p_{n-2l}}a_{j'-1,k,l}+$

$+2^{-1} \sum\limits_{k,l\in Z} \overline{p_{m-2k}}(-1)^n p_{1-n+2l}b^{[1]}_{j'-1,k,l}+$

$+2^{-1} \sum\limits_{k,l\in Z} (-1)^m p_{1-m+2k}\overline{p_{n-2l}}b^{[2]}_{j'-1,k,l}+$

$+2^{-1} \sum\limits_{k,l\in Z} (-1)^{m+n} p_{1-m+2k}p_{1-n+2l}b^{[3]}_{j'-1,k,l}.$

Как и в одномерном случае, коэффициенты $a_{j',m,n}$ определяются рекурсивно для $j'=1$, затем для $j'=2$ и т.д. до $j'=j$.

Для размерности $d \geq 2$ необходимо $2^d - 1$ базовых вейвлет-функций.

слайд 99

По аналогии с преобразованием Фурье можно ввести непрерывное вейвлет-преобразование. Вейвлет-преобразование, в отличие от преобразования Фурье, позволяет анализировать локальные свойства сигнала и отслеживать изменение частотного спектра. Чтобы определить вейвлет-преобразование, предположим, что дана функция $\psi(x) \in L^2(R)$, которая удовлетворяет условию

$$C_\psi = 2\pi \int\limits_0^\infty \left|\widehat{\psi}(a)\right|^2 \frac{da}{a} = 2\pi \int\limits_0^\infty \left|\widehat{\psi}(-a)\right|^2 \frac{da}{a} < \infty.$$

Это условие называется условием допустимости, а функция $\psi(x)$, удовлетворяющая ему, называется допустимой вейвлет-функцией (анализирующим вейвлетом). Для выполнения условия допустимости требуется, чтобы $\widehat{\psi}(0) = 0$, что в свою очередь влечет за собой равенство нулю интеграла $\int\limits_{-\infty}^{\infty} \psi(x)dx$. Таким образом, допустимая вейвлет-функция обладает хотя бы одним нулевым моментом. Определим вейвлет-преобразование от функции $f(x) \in L^2(R)$ как

$W_\psi f(a,b) = \int\limits_{-\infty}^{\infty} f(x)\frac{1}{\sqrt{a}}\overline{\psi\left(\frac{x-b}{a}\right)}dx, \quad b \in R, \ a > 0$. Если ввести обозначение

$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right)$,

то можно записать $W_\psi f(a,b) = \langle f, \psi_{a,b}\rangle$. Кроме того, легко видеть, что вейвлет-преобразование представляет собой свертку $f * \tilde{\psi}_{a,0}(b)$, где $\tilde{\psi}_{a,0}(x) = \overline{\psi_{a,0}(-x)}$. Для удобства введем еще одно обозначение вейвлет-преобразования, как оператора: $W_\psi[f](a,b) = W_\psi f(a,b)$.

слайд 100

Отметим некоторые свойства вейвлет-преобразования.

1. Линейность:

$W_\psi[f+g](a,b) = W_\psi[f](a,b) + W_\psi[g](a,b)$, $W_\psi[cf](a,b) = cW_\psi[f](a,b)$.

2. Инвариантность относительно сдвига:

$W_\psi[f(x-t)](a,b) = W_\psi[f(x)](a,b-t)$.

3. Инвариантность относительно масштабирования:

$W_\psi[f(cx)](a,b) = \frac{1}{\sqrt{c}}W_\psi[f(x)](ca,cb), \quad c > 0$.

4. Дифференцирование: если функции $f(x)$ и $\psi(x)$ n раз непрерывно дифференцируемы, и $\psi(x)$ быстро убывает на бесконечности вместе со своими производными, т.е. для всех $0 \le k \le n$ и любого $m \in N$ найдется константа C_m, что при всех $x \in R$

$\left|\frac{\partial^k}{\partial x^k}\{\psi(x)\}\right| \le \frac{C_m}{1+|x|^m}$,

то

$W_\psi\left[\frac{\partial^n}{\partial x^n}\{f(x)\}\right](a,b) = (-1)^n \int\limits_{-\infty}^{\infty} f(x)\frac{\partial^n}{\partial x^n}\{\overline{\psi_{a,b}(x)}\}dx$.

Из этого свойства следует, что безразлично, дифференцировать ли функцию или анализирующий вейвлет. Если анализирующий вейвлет задан формулой, то это может быть очень полезным для анализа сигналов. Проанализировать особенности высокого порядка или мелкомасштабные вариации сигнала с игнорированием крупномасштабных полиномиальных составляющих (тренда) можно дифференцированием нужного числа раз либо вейвлета, либо самого сигнала. Это свойство особенно полезно, когда сигнал задан дискретным рядом.

5. Связь между степенью гладкости функции и скоростью убывания ее вейвлет-преобразования при стремлении параметра масштаба к нулю:

если функция $f(x)$ n раз непрерывно дифференцируема в окрестности точки x_0, а $\psi(x)$ удовлетворяет требованиям, перечисленным в свойстве 4, и имеет n нулевых моментов, то $W_\psi f(a, x_0) \le Ca^{n+1/2}$ при $a \to 0$, где C – некоторая константа.

Это свойство позволяет анализировать локальную гладкость функции (в отличие от преобразования Фурье, которое позволяет анализировать лишь глобальную гладкость) и обнаруживать особенности (сингулярности).

Отметим, что дискретный аналог вейвлет-преобразования не является инвариантным относительно сдвига.

слайд 101

Для вейвлет-преобразования справедлив аналог формулы Планшереля.

Утверждение 1. Пусть $\psi(x) \in L^2(R)$ – допустимая вейвлет-функция и $C_\psi \ne 0$, тогда для $f(x) \in L^2(R)$ справедливо равенство

$$\int\limits_{-\infty}^{\infty} |f(x)|^2\, dx = \frac{1}{C_\psi} \int\limits_{0}^{\infty} \int\limits_{-\infty}^{\infty} |W_\psi f(a,b)|^2\, db \frac{da}{a^2}.$$

Приведем теперь формулу обращения для вейвлет-преобразования.

Утверждение 2. Пусть $\psi(x) \in L^2(R)$ – допустимая вейвлет-функция, $C_\psi \ne 0$ и $f(x) \in L^2(R)$, тогда

$$f(x) = \frac{1}{C_\psi} \int\limits_{0}^{\infty} \int\limits_{-\infty}^{\infty} W_\psi f(a,b) \psi_{a,b}(x) db \frac{da}{a^2}.$$

Как говорилось ранее, последовательность функций $\{\psi_{j,k},\ j,k \in Z\}$, порождающая кратномасштабный анализ, образует ортонормированный базис в $L^2(R)$, т.е. любую функцию $f \in L^2(R)$ можно представить в виде ряда

$$f = \sum_{j \in Z} \sum_{k \in Z} b_{j,k} \psi_{j,k}.$$

Легко видеть, что $b_{j,k} = \langle f, \psi_{j,k} \rangle = W_\psi f(2^{-j}, 2^{-j}k)$, т.е. коэффициенты ряда представляют собой непрерывное вейвлет-преобразование функции, вычисленное в точках $(2^{-j}, 2^{-j}k)$. Такой связи между рядами Фурье и преобразованием Фурье не существует, поскольку они определены на разных классах функций. Ряд Фурье определен для функций из $L^2([a,b])$ для некоторого отрезка $[a,b]$, а преобразование Фурье определено для функций из $L^2(R)$. В то же время, как вейвлет-ряды, так и непрерывные вейвлет-преобразования определены для функций из $L^2(R)$.

слайд 102

Так же как преобразование Фурье, непрерывное вейвлет-преобразование можно обобщить на многомерный случай. Мы рассмотрим только случай размерности 2. Одна из возможностей состоит в выборе вейвлет-функции $\psi(x, y)$, имеющей круговую симметрию, тогда ее преобразование Фурье также является функцией с круговой симметрией: $\widehat{\psi}(\omega_1, \omega_2) = \eta(\sqrt{\omega_1^2 + \omega_2^2})$, и условие допустимости превращается в

$C_\psi = (2\pi)^2 \int\limits_0^\infty |\eta(a)|^2 \frac{da}{a} < \infty.$

Вейвлет-преобразование от функции $f(x, y) \in L^2(R^2)$ по аналогии с одномерным случаем определяется как

$W_\psi f(a, b_1, b_2) = \langle f, \psi_{a,b_1,b_2} \rangle, \ \ b_1, b_2 \in R, \ a > 0,$

где $\psi_{a,b_1,b_2}(x, y) = \frac{1}{a}\psi\left(\frac{x-b_1}{a}, \frac{y-b_2}{a}\right),$

а формула обращения выглядит следующим образом:

$f(x, y) = \frac{1}{C_\psi} \int\limits_0^\infty \int\limits_{-\infty}^\infty \int\limits_{-\infty}^\infty W_\psi f(a, b_1, b_2)\psi_{a,b_1,b_2}(x, y) db_1 db_2 \frac{da}{a^3}.$

Другие возможные обобщения на многомерный случай можно найти в книге Добеши.

слайд 103

Приведем примеры часто используемых вейвлет-функций.

Пример 1. Вещественные вейвлет-функции часто получают с помощью дифференцирования функции Гаусса: $\psi_m(x) = (-1)^m \frac{d^m}{dx^m} e^{-x^2/2}, \widehat{\psi}_m(\omega) = (-i\omega)^m e^{-\omega^2/2}, \ \ m \geq 1$. Более высокие производные имеют больше нулевых моментов и позволяют анализировать особенности более высокого порядка.

Из-за формы графиков для $m = 1$ вейвлет-функцию называют WAVE-вейвлет.

слайд 104

а для $m = 2$ – MHAT-вейвлет или «мексиканская шляпа» (Mexican hat).

слайд 105

Пример 2. На основе функции Гаусса также строится часто используемый DOG-вейвлет (Difference of Gaussians):
$\psi(x) = e^{-x^2/2} - \frac{1}{2}e^{-x^2/8}$, $\widehat{\psi}(\omega) = e^{-\omega^2/2} - e^{-2\omega^2}$.

слайд 106

Пример 3. Хорошую локализацию в частотной области обеспечивает вейвлет-функция Мейера, которая определяется через свое преобразование Фурье следующим образом:

$$\widehat{\psi}(\omega) = \begin{cases} 2^{-1/2}\exp\left(-\frac{i\omega}{2}\right)\overline{\widehat{h}}\left(\frac{\omega}{2}+\pi\right) & \text{при } 2\pi/3 \le |\omega| < 4\pi/3, \\ 2^{-1/2}\exp\left(-\frac{i\omega}{2}\right)\widehat{h}\left(\frac{\omega}{4}\right) & \text{при } 4\pi/3 \le |\omega| \le 8\pi/3, \\ 0 & \text{иначе}, \end{cases}$$

где функция $\widehat{h}(\omega)$ удовлетворяет условиям

$$\left|\widehat{h}(\omega)\right|^2 + \left|\widehat{h}(\omega+\pi)\right|^2 = 2$$

и

$$\widehat{h}(\omega) = \begin{cases} \sqrt{2} & \text{при } \omega \in [-\pi/3, \pi/3], \\ 0 & \text{при } \omega \in [-\pi, -2\pi/3] \cup [2\pi/3, \pi]. \end{cases}$$

В области $[-2\pi/3, -\pi/3] \cup [\pi/3, 2\pi/3]$ поведение функции $\widehat{h}(\omega)$ может быть достаточно произвольным, лишь бы она удовлетворяла указанным условиям.

Вейвлеты Мейера имеют бесконечное число нулевых моментов, но не имеют компактного носителя. Функцию $\widehat{h}(\omega)$ можно выбрать так, чтобы $\widehat{\psi}(\omega)$ обладала заданным числом непрерывных производных и, таким образом, вейвлет-функция $\psi(x)$ достаточно быстро убывала на бесконечности.

слайд 107

Функцию $\widehat{h}(\omega)$ можно выбрать так, чтобы $\widehat{\psi}(\omega)$ обладала заданным числом непрерывных производных и, таким образом, вейвлет-функция $\psi(x)$ достаточно быстро убывала на бесконечности. На рис. **??** приведен график вейвлета Мейера для

$$\widehat{h}(\omega) = \sqrt{2}\cos\left[\frac{\pi}{2}\beta\left(\frac{3\,|\omega|}{\pi} - 1\right)\right] \quad \text{при} \quad |\omega| \in [\pi/3, 2\pi/3],$$

где

$$\beta(x) = x^4(35 - 84x + 70x^2 - 20x^3).$$

Так доопределенная функция $\widehat{h}(\omega)$ обеспечивает 3 непрерывные производные $\widehat{\psi}(\omega)$.

слайд 108

Пример 4. Наиболее часто используемая комплексная вейвлет-функция, хорошо локализованная и во временно́й и в частотной области, называется вейвлетом Морле:

$\psi(x) = e^{-x^2/2}(e^{i\omega_0 x} - e^{-\omega_0^2/2})$, $\widehat{\psi}(\omega) = e^{-(\omega-\omega_0)^2/2} - e^{-(\omega^2+\omega_0^2)/2}$.

Вейвлет Морле представляет собой плоскую волну, модулированную гауссианом. У вейвлета Морле равен нулю только нулевой момент. На рисунке изображен график вещественной части вейвлета Морле для $w_0 =$ 13.

слайд 109

Частотно-временная локализация

Как уже отмечалось, при анализе сигналов с помощью оконного преобразования Фурье частотно-временное окно остается фиксированным в том смысле, что его ширина не меняется при рассмотрении любого диапазона частот. Так как частота прямо пропорциональна числу периодов на единицу времени, то требуется более узкое временное окно для более точной локализации высокочастотных явлений и широкое временное окно для анализа низкочастотных составляющих сигнала. Следовательно, оконное преобразование Фурье плохо приспособлено для анализа сигналов, содержащих одновременно высокочастотные и низкочастотные составляющие. С другой стороны, вейвлет-преобразование при соответствующем выборе вейвлет-функции имеет гибкое частотно-временное окно, которое автоматически сужается при рассмотрении высокочастотных явлений и расширяется при изучении низкочастотных областей.

Действительно, если вейвлет-функция $\psi(x)$ заметно отлична от нуля лишь в некоторой окрестности начала координат, то $\psi_{a,b}(x)$ заметно отлична от нуля лишь в некоторой окрестности точки $x = b$. Причем эта окрестность сужается с уменьшением a. Кроме того, так как $\widehat{\psi}_{a,b}(\omega) = \sqrt{a}\widehat{\psi}(a\omega)e^{-ib\omega}$, то если предположить, что $\widehat{\psi}(\omega)$ заметно отлична от нуля лишь в некоторой окрестности точки $\omega = \xi$, $\widehat{\psi}_{a,b}(\omega)$ будет заметно отлична от нуля лишь в некоторой окрестности точки $\omega = \xi/a$. Следовательно, $W_\psi f(a,b)$ измеряет компоненту сигнала f в окрестности точки b с частотной характеристикой, пропорциональной $1/a$. Причем анализирующее окно сужается для измерения высокочастотных компонент и расширяется для измерения низкочастотных компонент. Таким образом, параметр b отвечает за положение частотно-временного окна во времени, т.е. за временну́ю локализацию, а параметр a (называемый параметром масштаба) отвечает за положение частотно-временного окна в частотной области, т.е. в некотором смысле отвечает за частотную локализацию. Аналогично при рассмотрении вейвлет-рядов параметр k отвечает за положение частотно-временного окна во времени, а параметр j отвечает за положение окна в частотной области.

слайд 110

Изложенная теория вейвлет-рядов и вейвлет-преобразований представляет собой аналог теорий рядов Фурье и преобразования Фурье. Для цифровой вейвлет-обработки сигналов на компьютерах необходим аналог дискретного преобразования Фурье. Таким аналогом является дискретное вейвлет-преобразование.

Пусть сигнал f задан в отсчетах $x_m = m/N$ $(m = 0, \dots, N-1)$, где $N = 2^J$ для некоторого J, т.е. представляет собой вектор длины N: $f[0], f[1], \dots, f[N-1]$. В дискретном случае скалярные произведения понимаются в смысле l_2:

$\langle f, g \rangle = \sum\limits_{m=0}^{N-1} f[m]g[m]$.

Дискретное вейвлет-преобразование (ДВП) представляет собой набор коэффициентов $d_{j,k}$ $(j = 0, \dots, J-1,\ k = 0, \dots, 2^j - 1)$, полученный по формуле $d = Wf$, где W – ортогональная матрица, удовлетворяющая соотношению $\sqrt{N} W_{(j,k),m} \approx \psi_{j,k}(m/N)$ (знак $\approx$ означает, что с отсчетами функций $\psi_{j,k}$ необходимо произвести незначительные модификации, чтобы сделать матрицу ортогональной. В матрице W также можно учесть разложение по базису, в котором используется не только вейвлет-функция ψ, но и масштабирующая функция ϕ. Подробнее о виде матрицы W см., например, в книге Малла.) Для прямого и обратного ДВП требуется $O(N)$ операций. Заметим, что для функции f коэффициенты ДВП последовательности ее отсчетов $f[0], f[1], \dots, f[N-1]$ приблизительно равны соответствующим коэффициентам ее вейвлет-ряда, умноженным на $\sqrt{N}$.

слайд 111

Прикладные задачи вейвлет-анализа

Основные практические задачи, при решении которых применяются методы вейвлет-анализа, связаны с компрессией данных, подавлением шумов и обнаружению особенностей (сингулярностей).

Важную роль для решения практических задач играет число нулевых моментов у вейвлет-функции, т.е. M, такое что

$\int\limits_{-\infty}^{\infty} x^k\psi(x)dx = 0 \;\; k = 0, \dots, M-1.$

Рассмотрим, например, вейвлет-функцию Добеши $\psi_2(x)$: $\int\limits_{-\infty}^{\infty} x^2\psi_2(x)dx \neq 0.$

Пусть функция сигнала f дважды непрерывно дифференцируема. Вычислим коэффициенты ее вейвлет-разложения. Учитывая размер носителя $\psi_2(x)$:

$b_{j,k} = \int\limits_{-\infty}^{\infty} f(x)2^{j/2}\psi_2(2^jx-k)dx = \int\limits_{0}^{3\cdot2^{-j}} f(x+2^{-j}k)2^{j/2}\psi_2(2^jx)dx$

Если j достаточно велико, то интервал интегрирования мал, и $f(x+2^{-j})k$ можно заменить квадратичной интерполяцией с помощью формулы Тейлора:

$f(x+2^{-j}k) \approx f(2{-}j) + xf'(2^{-j}k) + \frac{1}{2}x^2f''(2^{-j}k)$. Тогда

$b_{j,k} \approx \int\limits_{0}^{3\cdot2^{-j}} (f(2{-}j) + xf'(2^{-j}k) + \frac{1}{2}x^2f''(2^{-j}k))\psi_2(2^jx)dx.$

Этот интеграл вычисляется через первые три момента ψ_2. Первые два момента равны нулю. Вычислив третий момент, получим

$b_{j,k} \approx \frac{1}{16}\sqrt{\frac{3}{2\pi}}2^{-5j/2}f''(2^{-j}k).$

слайд 112

С помощью этих рассуждений можно, например, найти (окрестность) точки разрыва производной у функции. Эта операция называется обнаружением сингулярностей и позволяет, например, находить трещины в материалах.

Пусть например, дана функция

$$f(x) = \begin{cases} 0.37x + 1.37 & \text{при } x \in [-1, 2.8], \\ 1.58x - 2.03 & \text{при } x \in [2.8, 4]. \end{cases}$$

При использовании 256 отсчетов $j = 8$. Один применив разложение один раз, получаем вейвлет-коэффициенты при $j = 7$. Единственный значимый коэффииент $b_{7,97} = -0.01$. Остальные коэффициенты имеют порядок 10^{-14}. $k = 97$ соответствует $\psi_{7,97}$ с носителем на на отрезке $[2, 79, 2.83]$. Следовательно, сингулярность находится на этом отрезке (см. рисунок).

Чем больше у вейвлет-функции нулевых моментов, тем больше порядок сингулярности (номер соответствующей производной), который можно обнаружить.

Коэффициенты, не соответствующие области сингулярности, малы по абсолютному значению, и их можно обнулить, осуществив таким образом компрессию данных.

Методы подавления шума рассматриваются далее.

слайд 113

Методы подавления шума и сжатия данных

Пусть f – сигнал длины $N = 2^J$. Это значит, что данный сигнал f определяется набором чисел $f_0, f_1, \dots, f_{N-1}$. Обозначим через W возникающие помехи или шум, которые являются реализацией некоторого случайного процесса, тогда искаженный сигнал будет определяться как

$X = f + W$.

Обозначим координаты N-мерных векторов X, f и W в дискретном вейвлет-базисе $\{\psi_{j,k}\}$ (здесь $\psi_{j,k}$ – не ортогональный базис из функций, а базис, составляющий ортогональную матрицу дискретного вейвлет-преобразования).

$X_{j,k} = \langle X, \psi_{j,l} \rangle, \ \mu_{j,k} = \langle f, \psi_{j,k} \rangle, \ \varepsilon_{j,k} = \langle W, \psi_{j,k} \rangle.$

Тогда

$X_{j,k} = \mu_{j,k} + \varepsilon_{j,k}.$

Напомним, что $\mu_{j,k}$ приблизительно равны соответствующим коэффициентам вейвлет-ряда, умноженным на $\sqrt{N}$.

Предположим, что W – гауссовский белый шум с нулевым средним и дисперсией σ^2. Тогда в силу ортогональности вейвлет-преобразования (матрицы) случайные величины $\varepsilon_{j,k}$ будут также независимы и иметь гауссово распределение с нулевым средним и дисперсией σ^2.

Рассмотрим оценки $\widetilde{f}$ сигнала f, вычисленные с помощью диагонального оператора D в дискретном базисе $\{\psi_{j,k}\}$. Диагональный оператор оценивает каждую компоненту сигнала $\mu_{j,k}$ по соответствующей зашумленной компоненте $X_{j,k}$ с помощью функций $d_{j,k}(x)$:

$\widetilde{f} = D[X] = \sum\limits_{j,k} d_{j,k}(X_{j,k})\psi_{j,k}$.

Запишем функции d_m в виде $d_{j,k}(X_{j,k}) = a_{j,k}X_{j,k}$, где $a_{j,k}$ могут зависеть от $X_{j,k}$ ($a_{j,k} \in [0,1]$). Такой метод называется методом масштабирования коэффициентов.

слайд 114

Обозначим через $r_D(f)$ риск полученной оценки, вычисляемый по формуле

$r_D(f) = \mathsf{E}\|f - \widetilde{f}\|^2 = \sum\limits_{j,k} \mathsf{E}\left|\mu_{j,k} - d_{j,k}(X_{j,k})\right|^2$.

Так как $X_{j,k} = \mu_{j,k} + \varepsilon_{j,k}$, $\mathsf{E}\varepsilon_{j,k} = 0$ и $\mathsf{E}|\varepsilon_{j,k}|^2 = \sigma^2$, то

$\mathsf{E}\left|\mu_{j,k} - a_{j,k}X_{j,k}\right|^2 = |\mu_{j,k}|^2(1 - a_{j,k})^2 + \sigma^2 a_{j,k}^2$.

Для линейного оператора D (т.е. когда $a_{j,k}$ не зависит от $X_{j,k}$) это выражение принимает наименьшее значение при

$a_{j,k} = \frac{|\mu_{j,k}|^2}{|\mu_{j,k}|^2 + \sigma^2}$,

и минимальный риск равен

$r_{l,inf}(f) = \sum\limits_{j,k} \frac{|\mu_{j,k}|^2\sigma^2}{|\mu_{j,k}|^2 + \sigma^2}$.

В случае, когда D нелинейный, а коэффициенты $a_{j,k}$ принимают значения из множества $\{0,1\}$, минимум достигается при

$a_{j,k} = \begin{cases} 1, & \text{если } |\mu_{j,k}| \ge \sigma \\ 0, & \text{если } |\mu_{j,k}| < \sigma. \end{cases}$

При этом минимальный риск равен

$r_{inf}(f) = \sum\limits_{j,k} \min(|\mu_{j,k}|^2, \sigma^2)$.

Можно показать, что $r_{inf}(f) \ge r_{l,inf}(f) \ge \frac{1}{2} r_{inf}(f)$.

В обоих случаях оператор D не может использоваться практически, так как $a_{j,k}$ зависит от $|\mu_{j,k}|$, значение которого неизвестно.

слайд 115

Определение. Функции $d_{j,k}$ будем называть пороговыми, если обработка сигнала этими функциями зависит от сравнения амплитуды сигнала с некоторой константой (порогом).

Определение. Жесткой пороговой обработкой (с порогом T) называется процедура оценивания сигнала, при которой пороговые функции $d_{j,k}$ определяются как

$$d_{j,k}(x) = d_T(x) = \begin{cases} x, & \text{если } |x| > T \\ 0, & \text{если } |x| \le T. \end{cases}$$

Риск жесткой пороговой обработки равен

$$r_T(f) = \sum_{j,k} \mathsf{E} \left|\mu_{j,k} - d_T(X_{j,k})\right|^2,$$

где в силу того, что $X_{j,k} = \mu_{j,k} + \varepsilon_{j,k}$,

$$|\mu_{j,k} - d_T(X_{j,k})|^2 = \begin{cases} |\varepsilon_{j,k}|^2, & \text{если } |X_{j,k}| > T \\ |\mu_{j,k}|^2, & \text{если } |X_{j,k}| \le T. \end{cases}$$

Следовательно,

$$r_T(f) \ge r_{inf}(f) = \sum_{j,k} \min(|\mu_{j,k}|^2, \sigma^2).$$

слайд 116

Определение. Мягкой пороговой обработкой (с порогом T) называется процедура оценивания сигнала, при которой пороговые функции $d_{j,k}$ определяются как

$$d_{j,k}(x) = d_T(x) = \begin{cases} x - T, & \text{если } x > T \\ x + T, & \text{если } x < -T \\ 0, & \text{если } |x| \le T. \end{cases}$$

Отметим, что при мягкой пороговой обработке пороговые функции обладают свойством непрерывности, а при жесткой – пороговые функции разрывны. Отсюда вытекают и различия в полученных восстановленных сигналах.

Так увеличение порога приводит к большей гладкости аппроксимации при обоих видах обработки: мягкой и жесткой, но в ряде случаев все же остаются видимые локальные осцилляции. Причем такой эффект менее заметен при использовании процедуры мягкой пороговой обработки. Однако при использовании жесткой пороговой обработки не появляется дополнительного смещения, присущего мягкой пороговой обработке.

Метод пороговой обработки является близким к оптимальному (в смысле минимизации риска) для широкого класса сигналов, зашумленных аддитивным гауссовским белым шумом.

Помимо подавления шума описанные методы также позволяют решить задачу сжатия данных (с потерями).

слайд 117

На рисунке изображены графики функций жесткой и мягкой пороговых обработок.

слайд 118

Методы выбора порога

Универсальный порог

Донохо (D. Donoho) и Джонстон (I. Johnstone) предложили использовать так называемый «универсальный» порог, чей выбор обосновывается следующим утверждением.

Утверждение 1. Пусть $z_1, \ldots, z_n$ – независимые случайные величины, имеющие распределение $N(0, \sigma^2)$ и

$A_n = \{\max\limits_{1 \le i \le n} |z_i| > \sigma\sqrt{2\log n}\}$,

тогда

$P(A_n) \to 0 \quad$ при $n \to \infty$.

Более того, если

$B_n = \left\{\sigma\sqrt{2\log n} - \frac{\sigma \log n \log n}{\sqrt{\log n}} \le \max\limits_{1 \le i \le n} |z_i| \le \sigma\sqrt{2\log n}\right\}$,

$C_n(t) = \{\max\limits_{1 \le i \le n} |z_i| > \sigma t + \sigma\sqrt{2\log n}\}$,

то $P(B_n) \to 1$ при $n \to \infty$, и $P(C_n(t)) < e^{-\frac{t^2}{2}}$.

Локально энергия шума может быть сравнима и даже значительно превосходить энергию сигнала, поэтому, удалив шум при помощи высокого порога, мы потеряем полезную информацию. Напротив, низкий порог, оставляя входной сигнал практически без изменения, не удаляет почти всегда присутствующие слабые искажения. В этом случае и истинные нулевые коэффициенты сигнала будут вычислены с некоторой погрешностью. Тем самым в обработанный сигнал вносятся артефакты.

Порог $T = \sigma\sqrt{2\log n}$ во многом решает эту проблему. Согласно Утверждению 1, с большой вероятностью происходит удаление основного шума, так как вероятность A_n стремится к нулю, а оставшийся шум в обработанном сигнале будет незначительным. Стремление вероятности события B_n к единице означает, что максимальная амплитуда шума с ростом n все с большей вероятностью находится в окрестности универсального порога. А так как вероятность события C_n экспоненциально убывает, нет смысла выбирать порог, превосходящий универсальный. Универсальный порог зависит только от размера сигнала n и дисперсии шума σ^2.

слайд 119

При выборе универсального порога риск и мягкой, и жесткой пороговой обработки близок к риску

$r_{inf}(f) = \sum\limits_{j,k} \min(|\mu_{j,k}|^2, \sigma^2)$,

что обеспечивается следующей теоремой.

Теорема 1. Пусть $T = \sigma\sqrt{2\log N}$, $N = 2^J$ – объем сигнала. Риск $r_T(f)$ мягкой или жесткой пороговой оценки удовлетворяет при всех $N \geq 4$ неравенству

$r_T(f) \leq (2\log N + 1)(\sigma^2 + r_{inf}(f))$.

Из теоремы 1 следует, что при выборе универсального порога риск пороговой обработки близок к минимальному с точностью до логарифмического множителя для любого сигнала с конечной энергией (из $L^2(R)$). Именно поэтому порог назван «универсальным».

слайд 120

На практике дисперсия шума σ^2, как правило, неизвестна и ее необходимо оценивать. И поскольку оценку приходится строить по выборке сигнала, она должна быть устойчива к выбросам (робастна).

Определение. Пусть $X = (X_1, \ldots, X_n)$ – выборка, и $Z_i = |X_i - \mathrm{med}X|$, где med – выборочная медиана X. Тогда выборочное абсолютное медианное отклонение от медианы MAD (Median Absolute Deviation) определяется как выборочная медиана $Z = (Z_1, \ldots, Z_n)$: $MAD = \mathrm{med}Z$.

Оценка MAD является одной из самых устойчивых к выбросам оценок (робастной).

Утверждение 2. Пусть случайная величина ξ имеет нормальное распределение $N(a, \sigma^2)$, и m – медиана абсолютного отклонения ξ от ее медианы, тогда $\frac{m}{\sigma} = C_{3/4}$, где $C_{3/4}$ – 3/4-квантиль стандартного нормального распределения ($C_{3/4} \approx 0,6745$).

Исходя из утверждения 2 обычно в качестве оценки выбирается $\widehat{\sigma}$ выбирается величина

$\widehat{\sigma} = \frac{1}{0,6745} MAD$,

где MAD строится по половине всех коэффициентов на самом высоком масштабе вейвлет-разложения сигнала, так как если сигнал достаточно гладкий, то этот уровень содержит практически только шум.

Эти рассуждения приводят к определению порога $T_U = \widehat{\sigma}\sqrt{2\ln N}$. Эта модификация универсального порога уже не зависит ни от чего неизвестного.

слайд 121

SURE-порог

Риск обработки сигнала можно уменьшить, выбирая адаптивный порог, т.е. зависящий от зашумленных данных X.

Несмещенные оценки риска принято называть SURE-оценки, по аббревиатуре слов Stein Unbiased Risk Estimator. Порог, минимизирующий такую оценку при заданных зашумленных данных, также называют SURE-порогом.

Исследуем влияние порога на риск. Обозначим через $r_T(f)$ риск мягкой пороговой обработки с порогом T, который зависит от неизвестных значений $\mu_{j,k}$, и через $\widehat{r}_T(f)$ – оценку этого риска, зависящую от наблюдаемых зашумленных данных X. Напомним, что

$r_T(f) = \sum\limits_{j,k} \mathsf{E}|\mu_{j,k} - d_T(X_{j,k})|^2$.

Как разумно определить $\widehat{r}_T(f)$? Рассмотрим коэффициенты сигнала X. Заметим, что если $|X_{j,k}| \leq T$, то $d_T(X_{j,k}) = 0$, следовательно вклад в риск этого слагаемого равен $|\mu_{j,k}|^2$. Это число можно оценить разностью $|X_{j,k}|^2 - \sigma^2$, так как

$\mathsf{E}|X_{j,k}|^2 = |\mu_{j,k}|^2 + \mathsf{E}|\varepsilon_{j,k}|^2 = |\mu_{j,k}|^2 + \sigma^2$.

Если $|X_{j,k}| > T$, то $d_T(X_{j,k}) = X_{j,k} \pm T$ и вклад слагаемого равен

$\mathsf{E}|X_{j,k} - \varepsilon_{j,k} - X_{j,k} \pm T|^2 = \sigma^2 + T^2$.

Следовательно, риск мягкой пороговой обработки можно оценить следующим образом:

$\widehat{r}_T(f) = \sum\limits_{j,k} F(|X_{j,k}|^2)$,

где

$$F(u) = \begin{cases} u - \sigma^2, & \text{если } |u| \le T^2, \\ \sigma^2 + T^2, & \text{если } |u| > T^2. \end{cases}$$

В следующей теореме доказывается, что определенная выше оценка риска $\widehat{r}_T(f)$ является SURE-оценкой.

Теорема 2. При мягкой пороговой обработке оценка риска $\widehat{r}_T(f)$ не смещена, т.е. $\mathsf{E}\widehat{r}_T(f) = r_T(f)$.

При жесткой пороговой обработке несмещенную оценку риска построить нельзя.

слайд 122

Порог T_S, минимизирующий SURE-оценку $\widehat{r}_T(f)$, ищется эмпирически. Упорядочим координаты N-мерного сигнала $X_{j,k}$ по убыванию амплитуды, и обозначим через X_k^r k-ую координату ранжированного вектора. Так, что $|X_k^r| \ge |X_{k+1}^r|$ при $0 \le k < N-1$.

Пусть l – такой индекс, что

$|X_l^r| \le T < |X_{l-1}^r|$.

Тогда оценку риска можно переписать в виде

$$\widehat{r}_T(f) = \sum_{m=0}^{N-1} F(|X_m^r|^2) = \sum_{m=0}^{l-1} F(|X_m^r|^2) +$$
$$+ \sum_{m=l}^{N-1} F(|X_m^r|^2) = \sum_{m=0}^{l-1} (\sigma^2 + T^2) + \sum_{m=l}^{N-1} (|X_m^r|^2 - \sigma^2) =$$
$$= \sum_{m=l}^{N-1} |X_m^r|^2 - (N-l)\sigma^2 + l(\sigma^2 + T^2).$$

Отсюда видно, что оценка риска $\widehat{r}_T(f)$ при фиксированном l возрастает по T. Следовательно, взяв $|X_l^r|$ за значение порога, т.е. положив $T = |X_l^r|$, мы тем самым найдем минимум получившегося выражения. Порог T_S, минимизирующий $\widehat{r}_T(f)$, ищется сравнением N значений выражения для $\widehat{r}_T(f)$ в точках $T_m = |X_m^r|$, $0 \le m \le N-1$, соответственно. То значение $|X_m^r|$, на котором достигается минимум, принимается за T_S.

слайд 123

Однако, так вычисленное значение порога не всегда хорошо для обработки сигнала. Если энергия сигнала мала относительно энергии шума, т.е. $\|f\|^2 \ll \mathsf{E}\|W\|^2 = N\sigma^2$, то значение T_S может быть очень маленьким, вследствие чего сигнал остается сильно зашумленным. В этом случае, чтобы удалить шум, выбирают универсальный порог $T = \sigma\sqrt{2\log N}$.

Так как

$\mathsf{E}\|X\|^2 = \|f\|^2 + N\sigma^2$,

то $\|X\|^2 - N\sigma^2$ является несмещенной оценкой для энергии полезного сигнала $\|f\|^2$. Эта оценка энергии сигнала в дальнейшем сравнивается с величиной $\varepsilon_N = \sigma^2 N^{1/2}(\log N)^{3/2}$, называемой уровнем минимума энергии. Если она меньше этого уровня, значит шум перекрывает сигнал, и надо выбирать универсальный порог. Таким образом, порог определяется соотношениями

$$\widetilde{T}_S = \begin{cases} \sigma\sqrt{2\log N}, & \text{если } \|X\|^2 - N\sigma^2 \le \varepsilon_N \\ T_S, & \text{если } \|X\|^2 - N\sigma^2 > \varepsilon_N. \end{cases}$$

слайд 124

Пороговая обработка, инвариантная относительно сдвига

Напомним, что дискретное вейвлет-преобразование не является инвариантным относительно сдвига. У сдвинутого сигнала, вообще говоря, другие вейвлет-коэффициенты, что непосредственно влияет на результаты пороговой обработки.

Пороговые оценки сигнала могут быть значительно улучшены при помощи инвариантных сдвиговых алгоритмов, поскольку уточненная оценка вычисляется усреднением оценок для сдвинутых вариантов сигнала.

Рассмотрим сигналы f с периодом N в пространстве с дискретным вейвлет-базисом $\{\psi_{j,k}\}$.

Обозначим через f^p сигнал, полученный в результате сдвига сигнала f на p единиц вправо, т.е. $f_i^p = f_{i-p}$ для $i = 0, \ldots, N-1$. Пусть f_ψ – вектор коэффициентов разложения f по дискретному базису $\{\psi_{j,k}\}$.

У векторов f_ψ^p и f_ψ координаты не просто сдвинуты или переставлены, они совершенно различные.

Например, $f = (f_0, \ldots, f_{N-1}, f_0, \ldots)$, значит, $f^{-1} = (f_1, \ldots, f_0, f_1, \ldots)$. Но тогда

$f_{j,k} = \langle f, \psi_{j,k}\rangle \neq \langle f^{-1}, \psi_{j,k}\rangle = f_{j,k}^{-1}$.

Таким образом, сигнал, восстановленный в результате пороговой обработки по коэффициентам $f_{j,k}^p$, не является сдвигом сигнала, восстановленного по коэффициентам $f_{j,k}$.

слайд 125

Койфман и Донохо предложили алгоритм, при котором оцениваются все сдвиги сигнала f, затем производится обратный сдвиг самих оценок и их усреднение.

При всех $0 \le p < N$ оценка $\widetilde{f}^p$ сигнала f^p вычисляется с помощью пороговой обработки сдвинутых данных:

$\widetilde{f}^p = \sum\limits_{j,k} d_T(X^p_{j,k})\psi_{j,k}$,

где d_T – функция жесткой или мягкой пороговой обработки с порогом T.

Инвариантная относительно сдвига оценка получается сдвигом назад и усреднением этих оценок:

$\widetilde{f} = \frac{1}{N}\sum\limits_{m=0}^{N-1}(\widetilde{f}^p)^{-p}$.

Как видно, такая оценка требует в N раз больше вычислений, чем для стандартной пороговой оценки.

Свойство инвариантности относительно сдвига является необходимым при распознавании изображения. Когда изображение сдвигается, то его числовые характеристики также должны быть сдвинуты, но не изменены внутри изображения, т.е. представление изображения не должно зависеть от его расположения.

слайд 126

Обратные статистические задачи

В некоторых статистических задачах данные измеряются не напрямую, а после некоторого преобразования. Кроме того, в измерениях всегда присутствует шум, обусловленный несовершенством оборудования и различными случайными помехами.

Мы рассмотрим следующую модель данных: $X_i = (Kf)_i + \varepsilon_i, \quad i = 1, \dots, N (N = 2^J)$,

где X_i – наблюдаемые данные, K – некоторое линейное преобразование (линейный оператор), f – истинная (незашумленная) функция, которую необходимо оценить, а ε_i – случайные погрешности измерения. Будем предполагать, что все ε_i независимы и имеют одинаковое нормальное распределение с нулевым средним и дисперсией σ^2. Подобные модели используются в компьютерной томографии, физике плазмы, физической химии, астрономии и многих других областях.

Во многих случаях нельзя оценить функцию f, просто применив к данным X_i обратный оператор K^{-1}, поскольку такой оператор либо не существует, либо не является ограниченным. Статистические задачи такого рода называются некорректно поставленными, и для их решения применяются методы регуляризации часто в сочетании с сингулярным разложением.

слайд 127

Сингулярное разложение

Если K^*K является компактным оператором, то сингулярным разложением (Singular Value Decomposition) называется следующее представление функции f:

$f = \sum\limits_{m \in \mathbb{Z}} \kappa_m^{-1} \langle Kf, h_m \rangle e_m,$

где e_m – собственные функции оператора K^*K, κ_m^2 – его собственные значения, а $h_m = Ke_m / \|Ke_m\|_{L^2}$. Метод реконструкции, использующий сингулярное разложение, основан на идее, что «важные» коэффициенты расположены в начале этого ряда. Следовательно, для слагаемых выбирается последовательность весовых множителей t_m, которые близки к 1 для малых m и к 0 для больших. В результате получается так называемое оконное сингулярное разложение. Весовые множители выбираются таким образом, чтобы последовательность $\{t_m \kappa_m^{-1}\}$ была квадратично суммируема. Поскольку последовательность собственных значений компактного оператора K^*K стремится к нулю, использование таких весовых множителей необходимо для того, чтобы деление на близкие к нулю числа не приводило к расхождению ряда. Различный выбор весовых множителей t_m приводит к специальным типам регуляризации, например, квадратичной регуляризации Тихонова.

слайд 128

Сингулярное разложение представляет собой весьма популярный инструмент. Благодаря работе многих исследователей оно фактически стало парадигмой для анализа и решений линейных обратных задач. К тому же показано, что оконное сингулярное разложение при надлежащем выборе весовых множителей позволяет строить асимптотически наилучшие в минимаксном смысле линейные оценки.

Несмотря на свои привлекательные теоретические свойства, сингулярное разложение обладает некоторыми недостатками. Применение сингулярного разложения предполагает, что f имеет «экономное» представление в базисе собственных функций оператора K^*K, т.е. что лишь небольшое число коэффициентов в этом базисе являются «значимыми». Однако этот базис определяется только преобразованием K и не принимает в расчет свойства самой функции f. Эта особенность налагает некоторые ограничения на использование сингулярного разложения. Например, для операторов интегрирования и дифференцирования соответствующие собственные функции образуют базис Фурье из синусов и косинусов. Ряды Фурье подходят для представления пространственно однородных функций, но для функций, которые являются гладкими в одних областях и быстро меняющимися и имеющими локальные сингулярности в других, использование рядов Фурье неэффективно.

слайд 129

Как уже говорилось, в последние десятилетия значительно возросла популярность нелинейных методов обработки сигналов и изображений с помощью аппарата вейвлет-анализа. Объясняется это тем, что вейвлет-анализ позволяет гораздо более эффективно исследовать нестационарные сигналы, чем традиционный Фурье-анализ. В частности, чтобы обойти ограничения, присущие сингулярному разложению, Донохо предложил метод так называемого вейвлет-вейглет-разложения (Wavelet Vaguelette Decomposition). Функция $f \in L^2(R)$ представляется в виде ряда из сдвигов и растяжений некоторой вейвлет-функции ψ:

$$f = \sum_{j,k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k},$$

где $\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$. Как уже отмечалось ранее, семейство $\{\psi_{j,k}\}_{j,k \in Z}$ образует ортонормированный базис в $L^2(R)$.

Поскольку наблюдается не функция f, а ее линейное преобразование Kf, коэффициенты $\langle f, \psi_{j,k} \rangle$ вычислить напрямую нельзя. Идея метода вейвлет-вейглет-разложения заключается в том, чтобы выразить эти коэффициенты через Kf.

Основным свойством вейвлет-вейглет-разложения является то обстоятельство, что для некоторых линейных преобразований K существуют такие константы $\beta_{j,k}$, что последовательность функций $\xi_{j,k} = \beta_{j,k} K \psi_{j,k}$ образует устойчивый базис, т.е. существуют такие константы $0 < A \leq$

$B < \infty$, что

$$A \sum_{j,k} c_{j,k}^2 \le \left\| \sum_{j,k} c_{j,k} \xi_{j,k} \right\|_{L^2}^2 \le B \sum_{j,k} c_{j,k}^2$$

для всех квадратично суммируемых последовательностей $\{c_{j,k}\}$. Иногда это свойство называют «почти ортогональностью». Функции $\xi_{j,k}$ получили название «вейглеты».

слайд 130

Семейство функций $\{\xi_{j,k}\}$ существует не для всех линейных преобразований K. В дальнейшем мы будем рассматривать линейные преобразования K, однородные с показателем α, т.е.

$$K[f(a(x - x_0))] = a^{-\alpha}(Kf)[a(x - x_0)]$$

для любого x_0 и любого $a > 0$. Примерами однородных линейных преобразований служат оператор интегрирования, преобразование Абеля, лежащее в основе математической модели томографического анализа объектов, обладающих круговой симметрией, и некоторые виды операторов свертки. При положительном показателе α задача обращения однородного преобразования является некорректно поставленной. Если преобразование K однородно, то $\beta_{j,k} = \beta_{0,0} 2^{\alpha j}$, а функции $\xi_{j,k}$, так же как и вейвлеты, представляют собой сдвиги и растяжения некоторой функции ξ, т.е. «вейглеты» $\xi_{j,k}$ – это «почти» вейвлеты (за исключением свойства ортогональности).

слайд 131

Функция Kf разлагается в ряд по вейглет-базису:

$$Kf = \sum_{j,k} \langle Kf, \eta_{j,k} \rangle \xi_{j,k},$$

где $\{\eta_{j,k}\}$ – двойственный вейглет-базис, удовлетворяющий соотношению $K^* \eta_{j,k} = \beta_{j,k}^{-1} \psi_{j,k}$.

Базисы $\{\xi_{j,k}\}$ и $\{\eta_{j,k}\}$ биортогональны, и базис $\{\eta_{j,k}\}$ также является устойчивым и образуется с помощью сдвигов и растяжения функции η. С помощью базиса $\{\eta_{j,k}\}$ разложение f можно представить в виде

$$f = \sum_{j,k} \beta_{j,k} \langle Kf, \eta_{j,k} \rangle \psi_{j,k}.$$

Как видно, в этом представлении коэффициенты разложения выражаются не через функцию f, а через ее преобразование Kf. Эта формула является основой метода вейвлет-вейглет-разложения. Оценки пространственно неоднородных функций (т.е. функций, имеющих в разных областях разную степень регулярности), построенные по методу вейвлет-вейглет-разложения, в минимаксном смысле лучше, чем любые линейные оценки, и в частности, оконное сингулярное разложение.

слайд 132

Пример 1. Интегрирование. $Kf(x) = \int\limits_{-\infty}^{x} f(t)dt.$

Для того, чтобы можно было применить вейвлет-вейглет-разложение, функция ψ должна удовлетворять условиям:

$\int\limits_{-\infty}^{\infty} \left|\widehat{\psi}(\omega\right|^2 |\omega|^{-2} d\omega < \infty,\ \int\limits_{-\infty}^{\infty} \left|\widehat{\psi}(\omega\right|^2 |\omega|^2 d\omega < \infty,$

т.е. производная и интеграл от ψ должны лежать в $L^2(R)$. В этом случае

$\xi_{j,k}(x) = \beta_{j,k} \int\limits_{-\infty}^{x} \psi_{j,k}(t)dt,\ \eta_{j,k}(x) = -\beta_{j,k}^{-1}\psi'_{j,k}(x),$

где $\beta_{j,k} = \beta_{0,0}2^j$, $\beta_{0,0} = \|\eta\|$.

Из определения легко видеть биортогональность вейглет-базисов:

$\langle \xi_{j,k}, \eta_{i,m}\rangle = \langle \psi_{j,k}, \psi_{i,m}\rangle = \begin{cases} 1, & \text{если } (j,k) = (i,m), \\ 0, & \text{иначе.} \end{cases}$

Рост $\beta_{j,k}$ со скоростью 2^j указывает на степень некорректности задачи обращения.

слайд 133

Пример 2. Дробное интегрирование.

Пусть $H(t)$ – невырожденная функция, однородная с показателем 0: $H(at) = H(t)$ при любом $a > 0$.

Например, $H(t)$ может быть тождественной функцией, функцией Хевисайда или sgn.

$Kf(x) = \int\limits_{-\infty}^{\infty} f(t)\frac{H(t-x)}{|t-x|^{1-\alpha}}dt$ при некотором $\alpha \in (0,1)$.

Для функции Хевисайда $H(t) = \begin{cases} 1, & \text{при } x \geq 0, \\ 0, & \text{при } x < 0. \end{cases}$ и $\alpha = \frac{1}{2}$ получается преобразование Абеля.

В пространстве Фурье $\widehat{Kf}(\omega) = |\omega|^{-\alpha}\,\widehat{H}(\omega)\widehat{f}(\omega)$, $\widehat{H}$ также однородна с показателем 0.

Необходимые условия на функцию ψ для применимости вейвлет-вейглет-разложения получаются для так называемых дробных производных и дробных интегралов:
$\int\limits_{-\infty}^{\infty} \left|\widehat{\psi}(\omega\right|^{2\alpha} |\omega|^{-2\alpha}\, d\omega < \infty$, $\int\limits_{-\infty}^{\infty} \left|\widehat{\psi}(\omega\right|^{2\alpha} |\omega|^{2\alpha}\, d\omega < \infty$.

слайд 134

По формуле占Планшереля $\int\limits_{-\infty}^{\infty} f(t)g(t)dt = \int\limits_{-\infty}^{\infty} \widehat{f}(\omega)\widehat{g}(\omega)d\omega$ из соотношения $\beta_{j,k}\langle Kf, \eta_{j,k}\rangle = \langle f, \psi_{j,k}\rangle$ получаем

$\beta_{j,k}\widehat{\eta}_{j,k}(\omega)\,|\omega|^{-\alpha}\,\widehat{H}(\omega)\widehat{f}(\omega) = \widehat{\psi}_{j,k}(\omega)\widehat{f}(\omega)$.

Следовательно, формально

$\widehat{\eta}_{j,k}(\omega) = \widehat{\psi}_{j,k}(\omega)\,|\omega|^{\alpha}\,\beta_{j,k}^{-1}\widehat{H}^{-1}(\omega)$.

Чтобы $\eta_{j,k}$ и $\xi_{j,k}$ были корректно определены и принадлежали $L^2(R)$, достаточно чтобы ψ имела компактный носитель, два нулевых момента и две непрерывные производные.

В пространстве Фурье легко видеть биортогональность вейглет-базисов:
$\langle \xi_{j,k}, \eta_{i,m}\rangle = \langle \psi_{j,k}, \psi_{i,m}\rangle = \begin{cases} 1, & \text{если } (j,k) = (i,m), \\ 0, & \text{иначе.} \end{cases}$

Для дробного интегрирования $\beta_{j,k} = \beta_{0,0}2^{\alpha j}$, $\beta_{0,0} = \|\eta\|$.

При различных α получается разная степень некорректности задачи обращения. Например, для $\alpha = 1/2$, соответствующего преобразованию Абеля, получается меньшая степень некорректности, чем для интегрирования ($\alpha = 1$).

слайд 135

Помимо экономного представления функций нестационарных сигналов вейвлет-вейглет-разложение также позволяет решить проблему так называемого локального обращения однородного линейного оператора. Поясним идею этой проблемы «неформальным» образом. Некоторые виды таких операторов (например, преобразование Абеля) не допускают локального обращения в том смысле, что для получения значений функции f в некоторой окрестности необходимы все значения функции Kf.

Однако довольно часто полезная информация содержится в высокочастотных и короткоживущих компонентах сигнала, соответствующих «большим» значениям j в представлении

$f = \sum\limits_{j,k} \beta_{j,k} \langle Kf, \eta_{j,k} \rangle \psi_{j,k}$.

Оказывается, что если вейвлет функция ψ достаточно регулярна и имеет «большое» число нулевых моментов, то вейглет-функция η «достаточно быстро» убывает на бесконечности. Таким образом, при больших j (при которых $\eta_{j,k}$ уже «сильно» сжаты) для достаточно точного вычисления скалярных произведений $\langle Kf, \eta_{j,k} \rangle$ в вейвлет-вейглет-разложении f требуются значения Kf только из той окрестности, которая представляет интерес, и для вычисления высокочастотной компоненты f в этой окрестности следует производить суммирование при «больших» j и только по тем k, для которых $\psi_{j,k}$ «заметно» отличны от нуля в этой окрестности.

слайд 136

Абрамович и Сильверман предложили альтернативный метод разложения, получивший название вейглет-вейвлет-разложения. В этом методе в ряд по функциям $\psi_{j,k}$ раскладывается не функция сигнала f, а ее линейное преобразование Kf:

$Kf = \sum\limits_{j,k} \langle Kf, \psi_{j,k} \rangle \psi_{j,k}$.

Здесь для удобства мы не изменили обозначения для функций $\psi_{j,k}$, но в общем случае они отличаются от функций $\psi_{j,k}$ в вейвлет-вейглет-разложении.

Пусть $\beta_{j,k} = \|K^{-1}\psi_{j,k}\|$. Тогда для линейного однородного (с показателем α оператора $\beta_{j,k} = \beta_{0,0}2^{\alpha j}$).

Функция f представляется в виде ряда

$f = \sum\limits_{j,k} \beta_{j,k} \langle Kf, \psi_{j,k} \rangle u_{j,k}$,

где $u_{j,k} = K^{-1}\psi_{j,k}/\beta_{j,k}$ (по аналогии с предыдущим методом функции $u_{j,k}$ также называются «вейглетами»). Эта формула лежит в основе метода вейглет-вейвлет-разложения.

Последовательность $\{u_{j,k}\}$ не образует ортонормированную систему, однако если выполнены некоторые условия гладкости на $K^*\psi$ и $K^{-1}\psi$, то последовательность $\{u_{j,k}\}$ образует устойчивый базис.

Утверждение. Пусть существуют такие константы $A_i > 0$, $a_i > 0$ и

$b_i > 1$, $i = 1, 2$, что

$$\left|\widehat{K^{-1}\psi}(\omega)\right| \le A_1 \left|\omega\right|^{a_1} (1 + \left|\omega\right|^2)^{-(b_1+a_1)/2},$$
$$\left|\widehat{K^{*}\psi}(\omega)\right| \le A_2 \left|\omega\right|^{a_2} (1 + \left|\omega\right|^2)^{-(b_2+a_2)/2}$$

для всех $\omega \in R$, тогда последовательность $\{u_{j,k}\}$ образует устойчивый базис в $L^2(R)$.

По сути это условия на число нулевых моментов и непрерывных производных у ψ, похожие на условия применимости метода вейвлет-вейглет-разложения.

Анализ экспериментальных данных, приводимый в работе Абрамовича и Сильвермана, показывает, что методы вейвлет-вейглет-разложения и вейглет-вейвлет-разложения дают практически одинаковые результаты. Кроме того, в той же работе показано, что эти методы в минимаксном смысле эквиваленты.

Однако, например, в задаче локального обращения естественнее использовать метод вейглет-вейвлет-разложения, поскольку есть возможность сразу выбрать вейвлет-функцию ψ с компактным носителем и вычислять скалярные произведения $\langle Kf, \psi_{j,k}\rangle$ точно.

слайд 137

Особенности методов подавление шума

Напомним, что рассматривается модель

$X_i = (Kf)_i + \varepsilon_i, \qquad i = 1, \dots, N (N = 2^J),$

$\varepsilon_i \sim N(0, \sigma^2)$ и независимы.

При использовании вейвлет-вейглет-разложения дискретное вейглет-преобразование представляет собой умножение вектора отсчетов Kf на матрицу V (по аналогии с дискретным вейвлет-преобразованием). Однако теперь эта матрица не является ортогональной. После преобразования данные представляются в виде

$Y_{j,k} = \nu_{j,k} + z_{j,k},$

$\nu_{j,k} \approx \sqrt{N}\beta_{j,k}\langle Kf, \eta_{j,k}\rangle$, а $z_{j,k} \sim N(0, \sigma_j^2)$, где $\sigma_j^2 = \sigma^2\beta_{j,k}^2 = \sigma^2\beta_{0,0}2^{2\alpha j}$. Таким образом, дисперсия шума не является постоянной и растет с ростом j со скоростью $2^{2\alpha j}$, что является следствием некорректности задачи обращения K. Кроме того, дисперсия также зависит от выбранной вейвлет-функции ψ. Случайные величины $z_{j,k}$ уже не являются независимыми, однако в силу «почти ортогональности» вейглет-базиса этот факт игнорируется и обработка коэффициентов не отличается от случая независимых величин.

Поскольку дисперсия зависит от j пороговое значение в процедурах пороговой обработки также выбирается зависимым от j. Например, универсальный порог для каждого j вычисляется по формуле $T_j = \sigma_j \sqrt{2\log N}$.

слайд 138

При использовании вейглет-вейвлет-разложения вектор отсчетов Kf умножается на ортогональную матрицу W (дискретное вейвлет-преобразование) и после умножения на $\beta_{j,k}$ данные представляются в виде

$Y_{j,k} = \nu_{j,k} + z_{j,k}$,

$\nu_{j,k} \approx \sqrt{N}\beta_{j,k}\langle Kf, \psi_{j,k}\rangle$, $z_{j,k} \sim N(0, \sigma_j^2)$, где $\sigma_j^2 = \sigma^2\beta_{j,k}^2 = \sigma^2\beta_{0,0}2^{2\alpha j}$. Здесь случайные величины $z_{j,k}$ независимыми.

Порог также зависит от j (универсальный порог $T_j = \sigma_j\sqrt{2\log N}$).

Риск в обоих методах имеет одинаковый порядок. Можно предварительно нормировать коэффициенты разложения (поделить на $\beta_{j,k}$). Тогда при пороговой обработке используется один порог на всех масштабах j (например, универсальный порог $T = \sigma^2\beta_{0,0}\sqrt{2\log N}$). Результаты такой пороговой обработки получаются такими же.

слайд 139

Задачи вычислительной томографии

Томография — одно из важнейших направлений в области получения и обработки изображений. Образно говоря, томография позволяет заглянуть внутрь объекта, «просвечивая» его под различными углами обзора. Теоретические основы томографии были заложены в 1917 году Иоганом Радоном. В медицинских и биологических науках результаты Радона начали находить свое применение в 1960-х годах. Томографические методы революционизировали медицинскую диагностику, поскольку они позволили докторам «увидеть» внутренние органы, не подвергая пациента опасности. Создание компьютерных томографов (А. Кормак и Г. Н. Хаунсфилд) и их применение в биохимии (А. Клуг) отмечены Нобелевскими премиями (1979, 1982 гг.). Томографические методы применяются также в геологии, астрономии, сейсмологии, электронной микроскопии, диагностике плазмы, химии и во многих других областях.

Первые медицинские томографы для получения изображений внутренних тканей использовали рентгеновские лучи. Современная томография использует излучение самой различной физической природы: ультразвук, радио- и оптические сигналы, рентгеновские и γ-лучи, явление ядерно-магнитного резонанса и т. д. Для каждого вида излучения характерны свои особенности. Однако в большинстве случаев та информация, которая получается в процессе этих томографических экспериментов и с которой оперирует исследователь при восстановлении изображения, может быть описана очень похожими математическими зависимостями.

В своей основе томографические методы основаны на реконструкции изображения по его проекциям. Математические соотношения, на которых основаны современные методы вычислительной томографии, заимствуются из интегральной геометрии и применяются к томографическим измерениям с учетом методов решения некорректных задач. В простейшей математической модели, которая применяется в трансмиссионной томографии с параллельной схемой сканирования, проекцией под заданным углом называется интеграл от описывающей изображение функции по направлению, определяемому этим углом.

Пусть на плоскости, где введена прямоугольная система координат $\{x, y\}$, задана функция $f(x, y)$ с компактным носителем, описывающее изображение изучаемого объекта. Проинтегрируем эту функцию по прямой $L_{\varphi,s}$, которая описывается уравнением

$x\cos\varphi + y\sin\varphi - s = 0$, где s – расстояние от начала координат до прямой, а φ – угол, образованный с осью x перпендикуляром, опущенным из начала координат на эту прямую. Результат интегрирования по всем возможным прямым обозначим через $Rf(\varphi, s)$. Таким образом, преобразование (или оператор) Радона определяется следующим образом:

$$Rf(\varphi, s) = \int\limits_{L_{\varphi,s}} f(x, y)dl = \int\limits_{-\infty}^{\infty} f(s\cos\varphi - t\sin\varphi, s\sin\varphi + t\cos\varphi)dt,$$

$s \in R,\ \varphi \in [0, 2\pi)$. При фиксированном значении φ функцию $Rf(\varphi, s)$ также называют проекцией функции $f(x, y)$ под углом φ. Этим отражается геометрический смысл преобразования Радона, а именно тот факт, что в этом преобразовании все значения функции $f(x, y)$, лежащие на прямой, как бы интегрально проецируются в соответствующую точку (φ, s).

Задача реконструкции томографического изображения состоит в обращении преобразования Rf и нахождении функции f.

слайд 140

Преобразование Радона является основной математической моделью эксперимента трансмиссионной томографии с так называемой параллельной схемой сканирования: при фиксированном ракурсе обзора прямые, вдоль которых распространяется сканирующее излучение, расположены параллельно друг другу.

слайд 141

На рисунке изображен так называемый фантом Шеппа-Логана и его преобразование Радона. Фантом Шеппа-Логана - это одно из самых известных тестовых изображений, на котором схематически показан мозг человека.

слайд 142

Отметим, что функция $Rf(\varphi, s)$, описывающая проекционные данные, обладает важным свойством симметричности:

$Rf(\varphi + \pi, -s) = Rf(\varphi, s)$.

В его справедливости можно убедиться непосредственно, сделав в формуле, определяющей $Rf(\varphi, s)$, соответствующую замену переменных. Смысл же этого свойства состоит в том, что любые пары (φ, s) и $(\varphi + \pi, -s)$ задают одну и ту же прямую.

Это свойство позволяет перейти от области $s \geq 0$, $\varphi \in [0, 2\pi)$ к области $s \in R$, $\varphi \in [0, \pi)$, которая более удобна для приложений и, например, дает возможность снизить дозу облучения в два раза при обследовании пациента.

Отправным пунктом для вывода формулы обращения преобразования Радона служит теорема о центральном сечении, которая устанавливает связь между преобразованием Фурье функции $f(x, y)$ и преобразованием Фурье ее радоновского образа $Rf(\varphi, \omega)$ по переменной s.

Теорема 1 (о центральном сечении). Если существуют нижеследующие интегралы, то при любых углах φ справедливо равенство

$\widehat{Rf}(\varphi, \omega) = \sqrt{2\pi}\widehat{f}(\omega \cos\varphi, \omega \sin\varphi)$,

где

$\widehat{Rf}(\varphi, \omega) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} Rf(\varphi, \omega) e^{-is\omega} ds$,

$$\widehat{f}(\omega_1, \omega_2) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f(x,y) e^{-i(\omega_1 x + \omega_2 y)} dx dy$$

слайд 143

Эта теорема называется теоремой о центральном сечении, поскольку в плоскости $\{\omega_1, \omega_2\}$ множество $\omega_1 = \omega \cos\varphi$, $\omega_2 = \omega \sin\varphi$ при фиксированном угле φ является прямой, проходящей через начало координат под углом φ к оси ω_1. Таким образом, теорема о центральном сечении утверждает, что одномерное преобразование Фурье проекции $Rf(\varphi, \omega)$ по переменной s равно функции, описывающей центральное сечение двумерного преобразования Фурье функции $f(x, y)$, соответствующего тому значению угла φ, при котором вычисляется преобразование Фурье функции $Rf(\varphi, \omega)$.

Несмотря на свою простоту, теорема о центральном сечении чрезвычайно важна, и на нее опираются при выводе многих алгоритмов реконструкции.

слайд 144

Приведем еще одну теорему, обобщающую теорему о центральном сечении.

Теорема 2 (обобщенная проекционная теорема). Пусть $\omega(s)$ — произвольная функция одной переменной, для которой существуют нижеследующие интегралы, тогда при любых углах φ справедливо равенство

$$\int\limits_{-\infty}^{\infty} Rf(\varphi, \omega) w(s) ds = \int\limits_{-\infty}^{\infty} dx \int\limits_{-\infty}^{\infty} f(x,y) w(x\cos\varphi + y\sin\varphi) dy,$$

Так как функция $w(s)$ произвольна, то, конкретизируя ее вид, можно вывести различные полезные соотношения. Во-первых, если положить $w(s) = \exp(-i2\pi\omega s)$, то получится теорема о центральном сечении.

Пусть $w(s) = 1$. Тогда из утверждения теоремы получаем

$$\int\limits_{-\infty}^{\infty} Rf(\varphi, s) ds = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f(x,y) dx dy.$$

Это равенство указывает на то, что интеграл от проекции по переменной s, взятый по всей области, где $Rf(\varphi, s)$ отлична от нуля, не зависит от угла φ. Более того, значение этого интеграла равно двойному интегралу от функции $f(x, y)$, взятому по всей области, где $f(x, y)$ отлична от нуля.

Пусть $w(s) = s^m$. Тогда

$J^{(m)}(\varphi) = \int\limits_{-\infty}^{\infty} Rf(\varphi, s)s^m ds = \int\limits_{-\infty}^{\infty} f(x, y)(x\cos\varphi + y\sin\varphi)^m dxdy =$

$= \sum\limits_{n=0}^{m} (a_{m,n}\cos n\varphi + b_{m,n}\sin n\varphi)$

Интеграл $J^{(m)}(\varphi)$ называют m-м моментом проекции $Rf(\varphi, s)$ для данного угла φ, и выведенное равенство означает, что m-й момент проекции представляется тригонометрическим многочленом от φ степени не выше чем m. Из последнего, в частности, следует, что при изменении величины φ, m-й момент изменяется не быстрее, чем $\cos m\varphi$ или $\sin m\varphi$.

Эти соотношения называют условиями согласованности Хельгассона-Людвига. Они вместе с условием симметричности характеризуют пространство образов преобразования Радона.

слайд 145

Формулы обращения и алгоритмы реконструкции

Приступим к выводу алгоритмов реконструкции, основанных на теореме о центральном сечении.

По двумерному преобразованию Фурье $\hat{f}(\omega_1, \omega_2)$ можно найти саму функцию $f(x, y)$:

$f(x, y) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \widehat{f}(\omega_1, \omega_2) e^{i(\omega_1 x + \omega_2 y)} d\omega_1 d\omega_2.$

Сделаем замену переменных $\omega_1 = \omega\cos\varphi$, $\omega_2 = \omega\sin\varphi$:

$f(x, y) = \frac{1}{2\pi} \int\limits_{0}^{2\pi} d\varphi \int\limits_{0}^{\infty} \omega\widehat{f}(\omega\cos\varphi, \omega\sin\varphi) e^{i\omega(x\cos\varphi + y\sin\varphi)} d\omega.$

Теперь воспользуемся теоремой о центральном сечении и подставим функцию $\widehat{Rf}(\varphi, \omega)$ вместо $\hat{f}(\omega\cos\varphi, \omega\sin\varphi)$, после чего получим

$f(x, y) = \frac{1}{(2\pi)^{3/2}} \int\limits_{0}^{2\pi} d\varphi \int\limits_{0}^{\infty} \omega\widehat{Rf}(\varphi, \omega) e^{i\omega(x\cos\varphi + y\sin\varphi)} d\omega.$

Это равенство и является искомой формулой обращения, позволяющей по $\widehat{Rf}(\varphi, \omega)$ найти функцию $f(x, y)$. Однако привлечение этого равенства для обработки данных томографических экспериментов оказывается неудобным из-за используемой в нем области интегрирования. Поэтому, воспользовавшись свойством симметричности преобразования, получаем формулу обращения в следующем виде:

$$f(x,y)=\frac{1}{(2\pi)^{3/2}}\int\limits_0^{\pi}d\varphi\int\limits_{-\infty}^{\infty}|\omega|\,\widehat{Rf}(\varphi,\omega)e^{i\omega(x\cos\varphi+y\sin\varphi)}d\omega.$$

слайд 146

Метод реконструкции, основанный на этой формуле, носит название алгоритм Фурье-синтеза. Суть алгоритма заключается в следующем:

1. Используя теорему о центральном сечении из $\widehat{Rf}(\varphi,\omega)$ получаем $\widehat{f}(\omega\cos\varphi,\omega\sin\varphi)$.

2. Из полярной системы координат переходим в декартову и из функции $\widehat{f}(\omega\cos\varphi,\omega\sin\varphi)$ получаем $\widehat{f}(\omega_1,\omega_2)$.

3. С помощью обратного преобразования Фурье из $\widehat{f}(\omega_1,\omega_2)$ получаем $f(x,y)$.

Алгоритм Фурье-синтеза хорош тем, что на последнем шаге возможно применение быстрого преобразования Фурье, что увеличивает скорость его работы. Однако на втором шаге при численной реализации необходимо производить интерполяцию при переходе от полярной сетки к декартовой, что приводит к потере точности. Таким образом, этот алгоритм быстр (его сложность при обработке изображения размером $N\times N$ равна $O(N^2\log N)$), но за счет этого теряет в качестве получаемого изображения.

слайд 147

Алгоритм фильтрации и обратного проецирования

Обозначим через $h(t)$ импульсный отклик фильтра с частотной характеристикой $|\omega|$ (из-за формы графика такой фильтра называется ramp - трамплин). Формально

$$h(t)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}|\omega|\,e^{i\omega t}d\omega\ \ (\widehat{h}(\omega)=|\omega|).$$

Алгоритм состоит из следующих шагов:

1. При каждом φ вычислим

$\widetilde{Rf}(\varphi,s)=\frac{1}{(2\pi)^{3/2}}\int\limits_{-\infty}^{\infty}Rf(\varphi,t)h(s-t)dt$ – фильтрация проекций.

2. Далее при каждом фиксированном φ из функции одной переменной $\widetilde{Rf}(\varphi,s)$ получаем функцию двух переменных $\widetilde{Rf}(\varphi,x\cos\varphi+y\sin\varphi)$, «растягивая» ее на плоскость вдоль прямой с углом наклона φ. Эта операция называется обратным проецированием.

3. Наконец, суммируя обратные проекции, получаем

$f(x,y) = \int\limits_0^\pi \widetilde{Rf}(\varphi, x\cos\varphi + y\sin\varphi)$.

В алгоритме фильтрации и обратного проецирования вместо преобразования Фурье проекций участвуют сами проекции. Этот метод имеет большую вычислительную сложность, чем алгоритм Фурье-синтеза (при обработке изображения размером $N \times N$ сложность равна $O(N^3)$), однако является более точным, так как не требуется интерполировать значения функции с полярной сетки на декартову.

слайд 148

Линейная регуляризация формул обращения

Предположим, что функция $Rf(\varphi, s)$ задана с некоторой погрешностью. Очевидно, что по таким проекционным данным осуществить точное восстановление функции $f(x,y)$ невозможно. Тогда естественно задачу реконструкции переформулировать следующим образом: требуется по приближенным проекционным данным найти приближенную функцию $f_\alpha(x,y)$, которая в каком-то смысле хорошо описывала бы искомую функцию $f(x,y)$.

Пусть проекционные данные имеют вид

$Rf_e(\varphi, s) = Rf(\varphi, s) + n(s, \varphi)$,

где функция $n(s,\varphi)$ описывает случайную погрешность.

Задача реконструкции изображения относится к некорректным задачам. Если подставить $Rf_e(\varphi, s)$ в формулу обращения, то ramp-фильтр подчеркнет высокие частоты шумовой составляющей $n(s,\varphi)$, и даже небольшие отклонения в $Rf_e(\varphi, s)$ могут привести к большим погрешностям в восстановленной функции изображения (интеграл в формуле обращения может вообще не сходиться).

Идея линейной регуляризации - подавить высокие частоты, которые считаются шумом. Для этого в фильтр вводится стабилизирующий множитель $W_\alpha(|\omega|)$, зависящий от параметра α, который указывает на то, насколько сильно подавляются высокие частоты. Таким образом, формулы реконструкции записываются в следующем виде:

$f_\alpha(x,y) = \frac{1}{(2\pi)^{3/2}} \int\limits_0^\pi d\varphi \int\limits_{-\infty}^{\infty} |\omega|\, W_\alpha(|\omega|) \widehat{Rf}(\varphi, \omega) e^{i\omega(x\cos\varphi + y\sin\varphi)} d\omega.$

(для алгоритма Фурье-синтеза)

или

$$f_\alpha(x,y) = \frac{1}{(2\pi)^{3/2}} \int\limits_0^\pi d\varphi \int\limits_{-\infty}^{\infty} Rf_e(\varphi,s) h_\alpha(x\cos\varphi + y\sin\varphi - s)ds,$$

где $h(t) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} W_\alpha(|\omega|)\,|\omega|\,e^{i\omega t}d\omega$

(для алгоритма фильтрации и обратного проецирования).

слайд 149

Множитель $W_\alpha(|\omega|)$ должен удовлетворять определенным требованиям.

При $\alpha = 0$ $W_0(|\omega|) = 1$. При $|\omega| \sim 0$ $W_\alpha(|\omega|) \sim 1$, а при $|\omega| \to \infty$ $W_\alpha(|\omega|) \sim 0$.

Здесь мы приведем примеры наиболее часто используемых функций $W_\alpha(|\omega|)$.

1. $W_\alpha(|\omega|) = e^{-\alpha|\omega|}$ – Гауссово окно.
2. $W_\alpha(|\omega|) = e^{-\alpha^2(|\omega|^2)/2}$ – экспоненциальное окно.
3. $W_\alpha(|\omega|) = \begin{cases} 1, & \text{если } |\omega| \le 1/\alpha \\ 0, & \text{если } |\omega| < 1/\alpha. \end{cases}$ – прямоугольное окно.
4. $W_\alpha(|\omega|) = \frac{\left||\omega|^{-2}\right|}{|\omega|^{-2}+\alpha|\omega|^{2p}}$ – полиномиальное окно порядка p.

Выбор того или иного вида стабилизирующего множителя $W_\alpha(|\omega|)$ достаточно произволен. Выбор же конкретного значения параметра α связан с ожидаемой интенсивностью искажений. Причем чем больше интенсивность ожидаемых искажений, тем больше должно быть значение параметра α.

Регуляризованные формулы обращения являются приближенными, так что если в них подставить точные проекционные данные, то будет восстановлена функция $f_\alpha(x,y)$, не равная $f(x,y)$. Различие будет тем больше, чем больше параметр α, и при $\alpha = 0$ различие нивелируется. Таким образом, используя регуляризованную формулу обращения, мы сознательно идем на подобную ошибку, но зато формула обращения оказывается устойчивой, так что при небольших искажениях в проекционных данных восстанавливается функция $f_\alpha(x,y)$, мало отличающаяся от $f(x,y)$.

слайд 150

Статистическое описание изображений

Процесс регистрации проекций сопровождается случайными помехами. Поэтому восстанавливаемое изображение является случайным, т. е. описывается случайной функцией (случайным полем) $f_\alpha(x,y)$. Основные статистические характеристики этой функции включают в себя математическое ожидание $\mathsf{E}f_\alpha(x,y) \neq f(x,y)$, дисперсию

$\sigma_\alpha^2(x,y) = \mathsf{E}(f_\alpha(x,y) - \mathsf{E}f_\alpha(x,y))^2$ и ковариационную функция

$K_\alpha(x_1,y_1,x_2,y_2) = \mathsf{E}(f_\alpha(x_1,y_1) - \mathsf{E}f_\alpha(x_1,y_1))(f_\alpha(x_2,y_2) - \mathsf{E}f_\alpha(x_2,y_2))$

или коэффициент корреляции

$k_\alpha(x_1,y_1,x_2,y_2) = K_\alpha(x_1,y_1,x_2,y_2)/(\sigma_\alpha(x_1,y_1)\sigma_\alpha(x_2,y_2))$.

слайд 151

Обсудим более подробно приведенные величины. Для однородного случайного поля ковариационная функция зависит от разности координат $x_1 - x_2$ и $y_1 - y_2$:

$K_\alpha(x_1,y_1,x_2,y_2) = K_h(x_1 - x_2, y_1 - y_2)$.

Преобразование Фурье

$$\widehat{K}_h(\omega_1,\omega_2) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} K_h(x,y)e^{-i(\omega_1 x+\omega_2 y)}dxdy$$

функции $K_h(x,y)$ называется пространственным спектром случайного однородного поля $f_\alpha(x,y)$.

Однородные поля, у которых ковариационная функция зависит не от направления вектора $\overrightarrow{r} = \{x_1 - x_2, y_1 - y_2\}$, а только от его модуля $r = |\overrightarrow{r}| = \sqrt{(x_1-x_2)^2 + (y_1-y_2)^2}$, называются изотропными. В этом случае, переходя к полярным координатам, получаем

$$K_\alpha(x_1,y_1,x_2,y_2) = K_i(r) = 2\pi\int\limits_0^\infty \omega\hat{K}_i(\omega)J_0(2\pi\omega r)d\omega.$$

Здесь $J_0(b) = \frac{1}{2\pi}\int\limits_0^{2\pi} e^{ib\cos(\varphi-\theta)}d\varphi$ – функция Бесселя 1-го рода 0-го порядка, $\omega = \sqrt{\omega_1^2 + \omega_2^2}$ и $\hat{K}_i(\omega) = \hat{K}_h(\omega_1,\omega_2)$. Функция $K_i(r)$ обладает важным свойством – с ростом r она убывает.

слайд 152

Для изотропного поля можно ввести эффективный радиус корреляции, т. е. такое r_k, что $K_i(r_k)$ становится значительно меньше дисперсии

σ_α^2 (дисперсия однородного изотропного случайного поля не зависит от x, y).

Часто эффективный радиус корреляции определяется как

$r_k = \int\limits_0^\infty K_i(r)dr/\sigma_\alpha^2 = \int\limits_0^\infty k_i(r)dr.$

Примерами ковариационной функции однородного изотропного случайного поля могут служить гауссовская и экспоненциальная функции:

$K_i(r) = \sigma_\alpha^2 \exp(-r^2/2l^2),$

$K_i(r) = \sigma_\alpha^2 \exp(-r/l).$

У гауссовской ковариационной функции $r_k = l\sqrt{\pi/2}$, а у экспоненциальной $r_k = l$.

слайд 153

Критерии качества изображений

Изучив статистическое описание изображения, можно предсказать качество конкретной реализации. Однако сначала нужно формализовать понятие “качественного изображения”, т. е. сформулировать количественные критерии качества.

К сожалению, не существует универсального критерия качества. Поэтому при оценке качества изображения пользуются частными критериями, каждый из которых отражает какую-нибудь определенную особенность изображения.

1. Отношение сигнал-шум

$q_n(x,y) = \frac{\mathsf{E} f_\alpha(x,y)}{\sigma_\alpha(x,y)}.$

В том случае, когда отклонение среднего изображения от истинного существенно меньше, чем дисперсия, величина q_n хорошо описывает качество наблюдаемого изображения. Отметим, что отношение сигнал-шум разное в разных точках изображения, а также зависит от самого вида истинного изображения.

2. Среднеквадратичное отклонение

$\Delta^2(x,y) = \frac{\mathsf{E}(f_\alpha(x,y) - f(x,y))^2}{f^2(x,y)}$

представляет собой нормированное среднеквадратичное отклонение восстановленной функции от истинной в данной точке (x, y).

3. Проинтегрированное среднеквадратичное отклонение

$\Delta^2 = \frac{\frac{1}{S}\iint\limits_\Omega \mathsf{E}(f_\alpha(x,y) - f(x,y))^2 dxdy}{\left[\frac{1}{S}\iint\limits_\Omega f(x,y)dxdy\right]^2},$

где S — площадь области Ω, в которой f отлична от нуля. В отличие от предыдущей величины Δ^2 не зависит от точки (x, y).

слайд 154

Так как $\mathsf{E}(f_\alpha(x,y) - f(x,y))^2 = \sigma_\alpha^2(x,y) + [f(x,y) - \mathsf{E}f_\alpha(x,y)]^2$, то вводя обозначения

$\Delta_n^2(x,y) = \sigma_\alpha^2(x,y)/f^2(x,y)$,

$\Delta_\alpha^2(x,y) = (f(x,y) - Ef_\alpha(x,y))^2/f^2(x,y)$,

$\Delta_n^2 = \frac{1}{S}\iint\limits_\Omega \sigma_\alpha^2(x,y)dxdy/\left[\frac{1}{S}\iint\limits_\Omega f(x,y)dxdy\right]^2$,

$\Delta_\alpha^2 = \dfrac{\frac{1}{S}\iint\limits_\Omega (f(x,y) - Ef_\alpha(x,y))^2 dxdy}{\left[\frac{1}{S}\iint\limits_\Omega f(x,y)dxdy\right]^2}$,

имеем

$\Delta^2(x,y) = \Delta_n^2(x,y) + \Delta_\alpha^2(x,y)$,

$\Delta^2 = \Delta_n^2 + \Delta_\alpha^2$.

Первое слагаемое в этих суммах дает количественную оценку случайных отклонений от среднего изображения, второе слагаемое определяет отклонение среднего изображения от истинного. Поэтому в критериях, использующих величину $\Delta^2(x,y)$ или Δ^2 можно отдельно проанализировать влияние случайных погрешностей и факта отличия среднего изображения от истинного на качество изображения и сопоставить их между собой.

Критерии качества, использующие значения q_n, $\Delta^2(x,y)$ и Δ^2, можно рассчитать, если известен вид изображения. Поэтому применение этих критериев на практике подразумевает предварительный выбор каких-то типичных изображений, для которых проводится томографический эксперимент и определяются величины q_n, $\Delta^2(x,y)$ или Δ^2. Затем по значениям этих величин определяются требования к условиям формирования изображений.

слайд 155

Модель с аддитивным шумом

Рассмотрим модель, в которой измеряемые проекционные данные представляют собой сумму истинных проекций и некоторого шума. Такая модель является наиболее простой и ее можно рассматривать в качестве первого приближения для более сложных ситуаций.

Итак, пусть проекционные данные имеют вид

$Rf_e(\varphi, s)f(s) = Rf(\varphi, s) + n(s, \varphi)$,

$Rf(\varphi, s)$ – радоновский образ искомой функции $f(x, y)$, а $n(s, \varphi)$ – однородная случайная функция, у которой

$\mathsf{E}n(s, \varphi) = 0$, $K_n(s_1 - s_2, \varphi_1 - \varphi_2) = K(s_1 - s_2)\delta(\varphi_1 - \varphi_2)$.

Наличие δ-корреляции по переменной φ отражает обычную на практике ситуацию, при которой проекции для различных углов φ регистрируются независимо друг от друга. В этом случае естественно считать, что шумы являются статистически независимыми, что и означает δ-коррелированность.

Спектральную плотность шума по координате s обозначим $\widehat{K}(\omega)$:

$K(s) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} \widehat{K}(\omega)e^{i\omega s}d\omega$.

Восстанавливаемое томографическое изображение описывается функцией

$f_\alpha(x, y) = \int\limits_0^\pi d\varphi \int\limits_{-\infty}^{\infty} Rf_e(\varphi, s)f(s)h_\alpha(x\cos\varphi + y\sin\varphi - s)ds$,

где

$h_\alpha(t) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} |\omega|\, W_\alpha(|\omega|)e^{i\omega t}d\omega$.

Среднее изображение равно

$\mathsf{E}f_\alpha(x, y) = \frac{1}{(2\pi)^{3/2}} \int\limits_0^\pi d\varphi \int\limits_{-\infty}^{\infty} Rf(\varphi, s)h_\alpha(x\cos\varphi + y\sin\varphi - s)ds$,

а шумовое изображение

$f_n(x, y) = \frac{1}{(2\pi)^{3/2}} \int\limits_0^\pi d\varphi \int\limits_{-\infty}^{\infty} n(s, \varphi)h_\alpha(x\cos\varphi + y\sin\varphi - s)ds$.

слайд 156

Так как $\mathsf{E}f_n(x, y) = 0$, дисперсия шумового изображения равна

$\sigma_\alpha^2(x, y) = \mathsf{E}f_n^2(x, y) = \sqrt{2\pi} \int\limits_0^\infty \omega^2 W_\alpha^2(|\omega|)\widehat{K}(\omega)d\omega$,

Исходя из вида дисперсии можно сделать три вывода: 1) дисперсия не зависит от вида восстанавливаемого изображения; 2) значение дисперсии не зависит от конкретной точки (x, y) изображения; 3) чем большую роль в алгоритме играет функция $W_\alpha(|\omega|)$ (чем больше она подавляет высокие частоты), тем меньше дисперсия.

Первый вывод следует из аддитивности шума. Второй является следствием рассмотрения функции, заданной и восстанавливаемой на всей плоскости $\{x, y\}$. Данный вывод можно приближенно отнести лишь к той части области, которая находится достаточно далеко от границы. Третий вывод является общим - именно для того, чтобы можно было регулировать дисперсию, и проводится регуляризация формул обращения.

слайд 157

Ковариационная функция шумового изображения равна

$K_\alpha(x_1, y_1, x_2, y_2) = \mathsf{E} f_n(x_1, y_1) f_n(x_2, y_2) = \sqrt{2\pi} \int\limits_0^\infty \omega^2 W_\alpha^2(|\omega|) \widehat{K}(\omega) J_0(\omega\rho) d\omega,$

где $\rho = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$, а $J_0(\rho) = \frac{1}{2\pi} \int\limits_0^{2\pi} e^{i\rho \cos(\varphi - \theta)} d\varphi$ – функция Бесселя 1-го рода 0-го порядка.

Ковариационная функция определяется только значением ρ. Такая зависимость, как говорилось выше, характерна для однородного изотропного случайного поля.

Как и для дисперсии, эту формулу можно использовать лишь приближенно в той части изображения, которая находится достаточно далеко от его границы.

слайд 158

Неполные данные. Проблемы реконструкции изображений.

На практике изображения описываются классом интегрируемых функций с компактным носителем. В этом классе преобразования Фурье функций принадлежат классу вещественных аналитических функций в R^2, а вещественные аналитические функции не могут обращаться в ноль на бесконечном наборе прямых, проходящих через начало координат, не будучи тождественно равными нулю, т.е. справедливо следующее утверждение.

Лемма. Пусть $f(z, w)$ и $g(z, w)$, $z, w \in C$, — две комплексные функции комплексных переменных, каждая из которых аналитична в C^2, и пусть $l_1, l_2, ...$ — последовательность различающихся прямых, принадлежащих действительной плоскости R^2 и проходящих через начало координат. Если $f(z, w)$ и $g(z, w)$ совпадают на множестве $L = \bigcup\limits_{i=1}^{\infty} l_i$, то они совпадают для всех $(z, w) \in C^2$.

слайд 159

Из этой леммы и теоремы о центральном сечении следует:

Теорема 1. Объект однозначно определяется любым бесконечным набором своих проекций.

Типичным примером задачи восстановления функции по неполному набору данных является задача восстановления по проекциям с ограниченной областью значений углов обзора. Такая ситуация возникает, например, в задачах реконструкции в микроскопии, где направления лучей должны быть ограничены относительно узким углом. Предположим, что функция f имеет носителем единичный круг с центром в начале координат. В задаче восстановления функции по проекциям с ограниченной областью значений углов обзора измерения $Rf(\varphi, s)$ доступны только для диапазона углов $\varphi \in [\varphi_1, \varphi_2]$, где $|\varphi_2 - \varphi_1| < \pi$. Заметим, что поскольку $Rf(\varphi, s) = Rf(\varphi + \pi, -s)$, угол обзора величиной π достаточен для получения полного набора данных. Из теоремы 1 следует, что функция f однозначно определяется проекциями с ограниченной областью значений углов обзора. Но проблема заключается в отсутствии устойчивости восстановления. Из теоремы о центральном сечении следует, что в данном случае проекционные данные определяют преобразование Фурье $\widehat{f}(\omega_1, \omega_2)$ в конусе $\{\omega_1 = \rho \cos\varphi,\ \omega_2 = \rho \sin\varphi :\ \ \varphi \in [\varphi_1, \varphi_2],\ \rho \in R\}$. Восстановление f таким образом эквивалентно построению аналитического продолжения $\widehat{f}$, а задача аналитического продолжения является сильно неустойчивой.

слайд 160

Конечный набор проекций. Парадокс вычислительной томографии.

Хотя теорема 1 выглядит обнадеживающе, на практике вопрос состоит в том, когда объект однозначно определяется конечным набором своих проекций. Многие считали, что ложные объекты с многими проекциями, совпадающими с проекциями истинного объекта, хотя и существуют, но являются настолько “ненормальными”, что не восстанавливаются методами реконструкции или, по крайней мере, отвергаются интерпретатором. Однако справедлива теорема, которая, как казалось, вообще исключает всякую возможность практической реконструкции.

Теорема 2. Пусть объект описывается функцией f_0 с компактным носителем, дифференцируемой бесконечное число раз. Пусть известны его проекции по конечному числу направлений. Тогда существует другой объект, описываемый дифференцируемой бесконечное число раз функцией f, имеющий ту же форму, те же проекции по данным направлениям, что и первый объект, и совершенно произвольную структуру на любом компактном множестве, принадлежащем носителю f_0.

Эта теорема может быть переформулирована даже в более резкой форме.

Теорема 4'. Конечный набор проекционных данных ничего не говорит об объекте.

слайд 161

Шепп также привел следующий пример, подтверждающий это парадоксальное утверждение.

Пусть $f(x, y)$ – плотность вероятностного распределения на плоскости R^2 с компактным носителем $G \subset R^2$.

Пусть $\theta_1, \ldots, \theta_N$ – N направлений (N единичных векторов) на плоскости R^2.

Доказано, что для любой плотности $f: \ 0 \leq f(x, y) \leq 1, \ (x, y) \in G$, и для любого конечного числа направлений $\theta_1, \ldots, \theta_N, \ (N < \infty)$, существует еще одна плотность $f_0(x, y), \ (x, y) \in G$, такая что f_0 имеет те же проекции по направлениям $\theta_1, \ldots, \theta_N$, что и f, и принимает всего два значения: 0 или 1.

Этот результат приводит к следующему парадоксу компьютерной томографии. Для любого человеческого органа и соответствующих проекционных данных существует множество различных реконструкций, в частности, реконструкция состоящая только из кости и воздуха (плотности 1 и 0), которые имеют те же проекции, что и сам орган.

Такие примеры обычно игнорируются, поскольку томографы реконструируют приемлемые для практических целей изображения. Возможно, что объяснение этого парадокса состоит в том, что на практике точечная реконструкция невозможна. По-видимому, томографы реконструируют приемлемые изображения, поскольку все функции $f(x, y)$ с совпадающими по данному набору направлений проекциями, имеют (по существу) одинаковые проекции и по «подходящему» множеству направлений.

Другими словами, проекции «почти» совпадают и по окрестным направлениям, а все функции $f(x, y)$ с нулевыми проекциями по данному конечному набору направлений (так называемые «фантомы») являются сильно осциллирующими. Следовательно, для того чтобы уменьшить уровень неопределенности в случае конечного числа проекций, необходимо наложить ограничение на вариацию функции $f(x, y)$.

слайд 162

На рисунке показано несколько реконструкций фантома Шеппа-Логана при использовании различного числа проекций. Можно видеть, как улучшается качество при росте числа проекций.

слайд 163

Количественное решение парадокса вычислительной томографии.

Решение этого парадокса вычислительной томографии основано на оценке расстояния между функционалами от плотностей вероятностных мер с конечным числом (N) совпадающих проекций.

Пусть F_U – множество всех вероятностных плотностей с компактным носителем в единичном круге $U = \{(x, y) \in R^2 : x^2 + y^2 \leq 1\}$. Не ограничивая общности, можно считать, что изображения внутренней структуры объектов описываются функциями из класса F_U. Пусть $\phi_\sigma(x, y)$ – плотность нормального распределения $\phi_\sigma(x, y) = (2\pi\sigma^2)^{-1} \exp(-(x^2 + y^2)/(2\sigma^2))$, и пусть $f * \phi_\sigma$ и $g * \phi_\sigma$ $(f, g \in F_U)$ — свертки ϕ_σ с f и g.

Теорема 3. Пусть $n \geq 2$ — натуральное число, и существует $N = 2n$ направлений $\theta_1, \ldots, \theta_N$ на плоскости R^2 таких, что проекции плотностей f и g $(f, g \in F_U)$ по прямым в направлениях θ_j $(j = 1, \ldots, N)$ совпадают, тогда

$$\sup_{(x,y)\in R^2} |f * \phi_\sigma(x, y) - g * \phi_\sigma(x, y)| \leq \frac{1}{\left(\sqrt{\pi}\sigma^{n+2}2^{(3n-2)/2}\Gamma((n+1)/2)\right)},$$

где $\Gamma(z)$ — гамма-функция. Векторы $\theta_1, \ldots, \theta_N$ выбираются следующим образом:

$\theta_j = (\nu_j, -1)/(\nu_j^2 + 1)^{1/2}, \qquad j = 1, \ldots, n,$

$\theta_j = (1, \nu_{j-n})/(\nu_{j-n}^2 + 1)^{1/2}, \qquad j = n + 1, \ldots, 2n,$

где $\nu_k = \cos(\pi(2k - 1)/(2n)), \qquad k = 1, \ldots, n.$

Применение этой теоремы к плотностям f и f_0, упомянутым в конце предыдущего параграфа, дает качественное и количественное решение парадокса компьютерной томографии. Хотя расстояние по равномерной норме между f и f_0 равно 1, расстояние между сглаженными плотностями $f * \phi_\sigma$ и $g * \phi_\sigma$ мало для достаточно больших N.

слайд 164

Если вдобавок к условиям Теоремы 3 предположить, что $f(x, y)$ непрерывно дифференцируема, то можно получить оценку расстояния между $f(x, y)$ и $g * \phi_\sigma(x, y)$.

Теорема 4 Пусть $f \in F_U$ – непрерывно дифференцируемая функция плотности. При выполнении условий предыдущей теоремы, имеем

$$\sup_{(x,y)\in R^2} |f(x,y) - g * \phi_\sigma(x,y)| \le C_f \sqrt{\tfrac{\pi}{2}}\sigma + \frac{1}{\left(\sqrt{\pi}\sigma^{n+2} 2^{(3n-2)/2}\Gamma((n+1)/2)\right)},$$

где

$$C_f = \sup_{(x,y)\in R^2} \|\mathrm{grad} f(x,y)\|.$$

Для $\sigma(n) = n^{-1/2}$ получается неравенство

$$\sup_{(x,y)\in R^2} |f(x,y) - g * \phi_\sigma(x,y)| \le \widetilde{C} n^{-1/2},$$

где $\widetilde{C}$ зависит только от C_f.

Эта теорема дает оценку равномерного расстояния между искомой функцией f и ее реконструкцией, полученной по методу линейной регуляризации с помощью Гауссова окна.

слайд 165

Конечный набор проекций с погрешностями

Также получены аналогичные оценки равномерного расстояния между функциями из класса F_U, имеющими проекции по конечному числу направлений, отличающиеся друг от друга не более чем на ε, т. е. при наличии погрешностей при регистрации проекционных данных.

Теорема 5 Пусть проекции от плотностей f и g $(f, g \in F_U)$ по прямым в направлениях θ_j $(j = 1, \ldots, N)$, где θ_j те же, что и в теореме 3, отличаются не более чем на $\varepsilon \in (0, 1)$:

$$\sup_{s\in R} \left| \int\limits_{L(\theta_j,s)} f(x,y)dl - \int\limits_{L(\theta_j,s)} g(x,y)dl \right| \le \varepsilon, \quad j = 1 \ldots N,$$

где интеграл берется вдоль прямой $L(\theta_j, s)$ с направляющим вектором θ_j, сдвинутой относительно начала координат на величину s, тогда

$$\sup_{(x,y)\in R^2} |f * \phi_\sigma(x,y) - g * \phi_\sigma(x,y)| \le \frac{\varepsilon}{\pi\sigma^2} + \frac{\sqrt{2\pi}}{\sigma^3 n}.$$

слайд 166

Если дополнительно потребовать непрерывную дифференцируемость функций $f(x,y)$ и $g(x,y)$, то можно получить оценку расстояния между самими функциями.

Теорема 6. Пусть f и g ($f, g \in F_U$) непрерывно дифференцируемы. Тогда при выполнении условий предыдущей теоремы

$$\sup_{(x,y)\in R^2} |f(x,y) - g(x,y)| \le C_f\sqrt{\tfrac{\pi}{2}}\sigma + C_g\sqrt{\tfrac{\pi}{2}}\sigma + \frac{\varepsilon}{\pi\sigma^2} + \frac{\sqrt{2\pi}}{\sigma^3 n},$$

где

$$C_f = \sup_{(x,y)\in R^2} \|\mathrm{grad} f(x,y)\|,$$

$$C_g = \sup_{(x,y)\in R^2} \|\mathrm{grad} g(x,y)\|,$$

а σ — произвольное положительное число.

Если, например, взять $\sigma = \sqrt[3]{\varepsilon}$, а n порядка $O(\varepsilon^{-\frac{4}{3}})$, то

$$\sup_{(x,y)\in R^2} |f(x,y) - g(x,y)| \le C\varepsilon^{-\frac{1}{3}},$$

где C зависит от C_f и C_g.

слайд 167

Неполные данные. Локальная реконструкция.

Преобразование Радона, введенное в предыдущем параграфе, не допускает локального обращения в том смысле, что для реконструкции изображения в какой-либо точке требуется использование всех проекций – даже по прямым, которые не проходят вблизи этой точки. Это, в частности, означает, что для того чтобы получить изображение какого-то органа, необходимо просветить всего пациента (весь слой в районе органа).

Прежде чем обратиться к методам решения этой проблемы, перепишем формулу обращения преобразования Радона в более удобном виде.

Воспользовавшись свойством симметричности $Rf(\varphi+\pi, -s) = Rf(\varphi, s)$, можно заменить верхний предел интегрирования по углу с π на 2π в формуле обращения и поделить результат на 2:

$$f(x,y) = \frac{1}{2}\int_0^{2\pi} \widetilde{Rf}(\varphi, x\cos\varphi + y\sin\varphi)d\varphi,$$

где

$\widetilde{Rf}(\varphi, s) = \frac{1}{(2\pi)^{3/2}} \int\limits_{-\infty}^{\infty} Rf(\varphi, t) h(s - t) dt,$

$h(t) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} |\omega| \, e^{i\omega t} d\omega.$

Определим оператор обратного проецирования, применяемый к функции $g(\varphi, s)$ (оператор R^* является сопряженным к оператору R):

$R^*[g](x, y) = \int\limits_{0}^{2\pi} g(\varphi, x\cos\varphi + y\sin\varphi) d\varphi.$

Получается, что в алгоритме фильтрации и обратного проецирования сначала осуществляется свертка проекционных данных с функцией h, а затем к получившейся функции $\widetilde{Rf}$ применяется оператор обратного проецирования R^*.

В пространстве Фурье можно записать:

$f(x, y) = \frac{1}{2(2\pi)^{3/2}} \int\limits_{0}^{2\pi} d\varphi \int\limits_{-\infty}^{\infty} |\omega| \, \widehat{Rf}(\varphi, \omega) e^{i\omega(x\cos\varphi + y\sin\varphi)} d\omega.$

Определим (формально) через преобразование Фурье оператор I^α (потенциал Рисса), действующий на функцию $q(x)$:

$\widehat{I^\alpha[q]}(\omega) = |\omega|^{-\alpha} \, \widehat{q}(\omega).$

слайд 168

Тогда формулу обращения можно записать в следующем виде:

$f(x, y) = \frac{1}{4\pi} R^* \left[I^{-1}[Rf] \right](x, y),$

где оператор I^{-1} действует на $Rf(\varphi, s)$ при каждом фиксированном φ как на функцию от s.

Эту формулу можно обобщить. Пусть функции $f(x, y)$ и $g(\varphi, s)$ ($(x, y) \in R^2$, $\varphi \in [0, 2\pi)$, $s \in R$) интегрируемы и имеют компактный носитель, тогда

$f * R^*[g](x, y) = R^*[Rf * g](x, y),$

где свертка в левой части осуществляется относительно двух переменных (x, y), а свертка в правой части – относительно одной переменной s при каждом фиксированном φ.

Последняя формула лежит в основе многих методов реконструкции томографических изображений. Если выбрать функцию $g(\varphi, s)$ такой, что $R^*[g](x, y)$ «достаточно хорошо» приближает δ-функцию, то можно получить хорошую аппроксимацию для функции $f(x, y)$.

слайд 169

Поясним причину невозможности локального обращения Радона. Обозначим $G(x,y) = R^*[g](x,y)$. По формуле обращения имеем

$G(x,y) = \frac{1}{4\pi} R^* \left[I^{-1}[RG]\right](x,y)$.

Следовательно,

$g(\varphi,s) = \frac{1}{4\pi} I^{-1}[RG](\varphi,s)$.

По определению оператора I^{-1} и теореме о центральном сечении $g(\varphi,s)$ удовлетворяет соотношению

$\widehat{g}(\varphi,\omega) = \frac{1}{2\sqrt{2\pi}} |\omega| \widehat{G}(\omega\cos\varphi, \omega\sin\varphi)$.

Для локального обращения преобразования Радона с использованием проекций только по тем прямым, которые проходят вблизи точек реконструкции было бы достаточно, чтобы обе функции – $g(\varphi,s)$ и $G(x,y)$ – имели компактный носитель.

слайд 170

В этом случае можно было бы выбрать положительные числа a' и a ($a' > a$) такие, что

$\mathrm{supp}(G) \subseteq \{(x,y) : \sqrt{x^2+y^2} \leq a' - a\}$

и

$\mathrm{supp}(g) \subseteq [-a'+a, a'-a]$ для всех $\varphi \in [0,2\pi)$.

Это позволило бы точно восстановить функцию $f * G(x,y)$ для $\{(x,y) : \sqrt{x^2+y^2} \leq a\}$, используя $Rf(\varphi,s)$ только при $s \in [-a', a']$.

Предположив, что

$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} G(x,y)dxdy = 1$,

и положив $G_\varepsilon(x,y) = \varepsilon^{-2} G(x/\varepsilon, y/\varepsilon)$, можно было бы аппроксимировать $f(x,y)$ при $\{(x,y) : \sqrt{x^2+y^2} \leq a\}$ сверткой $f * G_\varepsilon(x,y)$ сколь угодно точно, выбирая соответствующее $\varepsilon > 0$.

слайд 171

Однако $g(\varphi,s)$ и $G(x,y)$ не могут одновременно иметь компактный носитель. Предположим, например, что функция $G(x,y)$ имеет компактный носитель, тогда хорошо известно, что $\widehat{G}(\omega_1,\omega_2)$ можно аналитически продолжить до целой функции в C^2, поэтому для любого угла $\varphi \in [0,2\pi)$ функция $\widehat{G}(\omega\cos\varphi, \omega\sin\varphi)$ представляет собой сужение целой функции на R. Для того чтобы функция $g(\varphi, s)$ имела компактный носитель, необходимо, чтобы функция $|\omega|\,\widehat{G}(\omega\cos\varphi, \omega\sin\varphi)$ также представляла собой сужение целой функции на R. Но это невозможно, так как предполагается, что

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} G(x,y)dxdy = 2\pi\widehat{G}(0,0) = 1,$$

а производная функции $|\omega|$ разрывна в точке $\omega = 0$. Следовательно, функция $g(\varphi, s)$ не может иметь компактного носителя и локальная реконструкция невозможна.

Замечание Преобразование Радона можно ввести в пространствах размерности $d \geq 2$. При этом оказывается, что когда d четно, задача реконструкции не локальна, а когда d нечетно, локальная реконструкция возможна.

слайд 172

Хотя компактность носителя функции $g(\varphi, s)$ обеспечить невозможно, можно обеспечить «достаточно» быстрое убывание $g(\varphi, s)$ на бесконечности. Для этого нужно потребовать, чтобы $|\omega|\,\widehat{G}(\omega\cos\varphi, \omega\sin\varphi)$ была как можно более гладкой. Поскольку $|\omega|\,\widehat{G}(\omega\cos\varphi, \omega\sin\varphi)$ бесконечно дифференцируема везде за исключением точки $\omega = 0$, для обеспечения желаемой гладкости необходимо, чтобы $\widehat{G}(\omega_1, \omega_2)$ имела ноль достаточно высокого порядка в точке $(0,0)$. Это означает наличие у $G(x,y)$ достаточно большого числа нулевых моментов. При этом самое большее, на что можно надеяться, это «почти» локально восстановить $f * G(x,y)$. При соответствующем выборе $G(x,y)$ фактически получается вейвлет-преобразование (двумерное) функции $f(x,y)$.

слайд 173

Напомним определение вейвлет-преобразования в двумерном случае.

Пусть вейвлет-функция $\Psi(x,y)$ имеет круговую симметрию. Ее преобразование Фурье также является функцией с круговой симметрией: $\widehat{\Psi}(\omega_1, \omega_2) = \eta(\sqrt{\omega_1^2 + \omega_2^2})$. Условие допустимости выражается соотношением

$$C_\psi = (2\pi)^2 \int\limits_0^{\infty} |\eta(a)|^2 \frac{da}{a} < \infty.$$

Вейвлет-преобразование от функции $f(x,y) \in L^2(R^2)$ определяется как

$$W_\psi f(a,b_1,b_2) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(x,y)\psi_{a,b_1,b_2}(x,y)dxdy, \quad b_1,b_2\in R,\ a>0,$$

где $\psi_{a,b_1,b_2}(x,y) = \frac{1}{a}\psi\left(\frac{x-b_1}{a}, \frac{y-b_2}{a}\right)$,

а формула обращения выглядит следующим образом:

$$f(x,y) = \frac{1}{C_\psi}\int\limits_{0}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} W_\psi f(a,b_1,b_2)\psi_{a,b_1,b_2}(x,y)db_1db_2\frac{da}{a^3}.$$

слайд 174

Покажем, как можно вычислить непрерывное вейвлет-преобразование функции с помощью непрерывных вейвлет-преобразований ее проекций.

Теорема 1. Пусть $\Psi(x,y)$ – допустимая вещественная вейвлет-функция, обладающая круговой симметрией и удовлетворяющая при некотором $C>0$ условию

$$\left|\widehat{\Psi}(\omega_1,\omega_2)\right| \leqslant C(1+\sqrt{\omega_1^2+\omega_2^2})^{-2} \text{ при всех } (\omega_1,\omega_2)\in R^2.$$

Определим четную функцию $\rho(s) = \rho(\theta,s)$ с помощью следующего соотношения:

$$\widehat{\rho}(\omega) = \frac{1}{2\sqrt{2\pi}}|\omega|\,\widehat{\Psi}(\omega\cos\theta,\omega\sin\theta)$$

(в силу круговой симметрии эта функция не зависит от θ), тогда функция $\rho(s)$ представляет собой допустимую вейвлет-функцию, и для любой функции $f(x,y)\in L^1(R^2)\bigcap L^2(R^2)$ с конечным носителем

$$W_\Psi f(a,b_1,b_2) = a^{-1/2}\int\limits_0^{2\pi} W_\rho[R_\theta f](a, b_1\cos\theta+b_2\sin\theta)d\theta,\ a>0,\ b_1,b_2\in R.$$

Смысл теоремы 1 заключается в том, что вейвлет-преобразование функции $f(x,y)$ можно вычислить, применив оператор обратного проецирования к вейвлет-преобразованиям проекций $R_\theta f(s)$, используя допустимую вейвлет-функцию для каждого $\theta\in[0,2\pi)$.

слайд 175

Справедливо также следующее утверждение.

Теорема 2. Пусть $\rho(s)$ – допустимая четная вещественная вейвлет-функция, удовлетворяющая условию

$$\int\limits_0^{\infty}\frac{|\widehat{\rho}(\omega)|^2}{\omega^3}d\omega<\infty.$$

Определим функцию $\Psi(x,y)$ (обладающую круговой симметрией) соотношением

$\widehat{\Psi}(\omega_1, \omega_2) = 2\sqrt{2\pi}\frac{\widehat{\rho}(\sqrt{\omega_1^2+\omega_2^2})}{\sqrt{\omega_1^2+\omega_2^2}}$,

тогда $\Psi(x, y)$ является допустимой вейвлет-функцией, и

$W_\Psi f(a, b_1, b_2) = a^{-1/2} \int\limits_0^{2\pi} W_\rho[R_\theta f](a, b_1 \cos\theta + b_2 \sin\theta) d\theta,\ a > 0,\ b_1, b_2 \in R.$

Теорема 2 похожа на теорему 1, но в ней вейвлет-функция $\rho(s)$ фиксирована, и с помощью этой функции определяется вейвлет-функция $\Psi(x, y)$. Эти теоремы помогают понять, как вейвлет-преобразования можно использовать для обработки изображений и локализации задачи реконструкции. Для задач обработки изображений обычно заранее выбирается функция $\Psi(x, y)$, обладающая подходящими свойствами, а затем рассматриваются свойства $\rho(s)$. Для локализации нужна вейвлет-функция $\rho(s)$ с маленьким носителем и достаточно большим числом нулевых моментов. В этом случае область, в которой функция $\Psi(x, y)$ заметно отлична от нуля, будет иметь почти такой же радиус, что и носитель $\rho(s)$. Следовательно, вейвлет-преобразование функции $f(x, y)$ можно вычислить локально, используя локальную информацию о преобразовании Радона.

После вычисления вейвлет-преобразования $f(x, y)$ можно воспользоваться формулой обращения для непрерывного вейвлет-преобразования и получить саму функцию $f(x, y)$.

слайд 176

Пример локально восстановленного изображения фантома Шеппа-Логана.

слайд 177

Степень локализации зависит от скорости убывания оператора I^{-1} от вейвлет-функции. В следующей теореме приводится оценка этой скорости в зависимости от количества нулевых моментов у вейвлет-функции.

Теорема 3. Пусть функция $h(s)$ удовлетворяет следующим условиям:

1. $h(s)$ имеет компактный носитель,
2. $\int\limits_{-\infty}^{\infty} s^j h(s) ds = 0$ при $j = 0, \ldots, m$ для некоторого $m \geqslant 0$,

3. $\int\limits_{-\infty}^{\infty} \left|\omega^j \widehat{h}^{(k)}(\omega)\right| d\omega < \infty$ при $j = 0, 1$, и $k = 0, \ldots, m+1$, где $\widehat{h}^{(k)}(\omega)$ – k-я производная $\widehat{h}(\omega)$, и

4. $\int\limits_{-\infty}^{\infty} \left|\omega^j \widehat{h}^{(k)}(\omega)\right|^2 d\omega < \infty$ при $j = 0, 1$, и $k = 0, \ldots, m+1$,

тогда

$I^{-1}[h](s) = o(|s|^{-(m+1)})$ при $|s| \to \infty$,

и

$s^{m+1} I^{-1}[h](s) \in L^2(R)$.

Первая формула показывает скорость убывания функции $I^{-1}[h](s)$ для больших значений s, а вторая дает усредненную оценку убывания и показывает, что «энергия» $I^{-1}[h](s)$ мала вдали от начала координат.

слайд 178

Кратномасштабная реконструкция

Опишем теперь метод локальной реконструкции, основанный на вейвлет-вейглет-разложении функции f.

Пусть $\varphi(x)$ – одномерная масштабирующая функция, а $\psi(x)$ – одномерная вейвлет-функция. Определим

$\Phi_{j,k_1,k_2}(x,y) = 2^j \varphi\left(2^j x - k_1\right) \varphi\left(2^j y - k_2\right)$,

$\Psi^{[1]}_{j,k_1,k_2}(x,y) = 2^j \varphi\left(2^j x - k_1\right) \psi\left(2^j y - k_2\right)$,

$\Psi^{[2]}_{j,k_1,k_2}(x,y) = 2^j \psi\left(2^j x - k_1\right) \varphi\left(2^j y - k_2\right)$,

$\Psi^{[3]}_{j,k_1,k_2}(x,y) = 2^j \psi\left(2^j x - k_1\right) \psi\left(2^j y - k_2\right)$.

Это так называемое тензорное произведение двух одномерных кратномасштабных анализов. При любом $j_0 \in Z$ последовательность

$\left\{\Phi_{j_0,k_1,k_2},\, \Psi^{[\lambda]}_{j,k_1,k_2}\right\}$, $j, k_1, k_2 \in Z, j \geqslant j_0, \lambda = 1, 2, 3,$

является ортонормированным базисом в $L^2(R^2)$. Также ортонормированным базисом в $L^2(R^2)$ является последовательность

$\left\{\Psi^{[\lambda]}_{j,k_1,k_2}\right\}$, $j, k_1, k_2 \in Z, \lambda = 1, 2, 3.$

слайд 179

Рассмотрим разложение функции $f \in L^2(R^2)$ по второму «однородному» базису (разложение по первому базису рассматривается аналогично):

$$f = \sum_{\substack{j,k_1,k_2 \in Z,\\ \lambda=1,2,3}} \langle f, \Psi^{[\lambda]}_{j,k_1,k_2} \rangle \Psi^{[\lambda]}_{j,k_1,k_2}.$$

Используя формулу обращения, можно записать вейвлет-вейглет-разложение f в виде

$$f = \sum_{\substack{j,k_1,k_2 \in Z,\\ \lambda=1,2,3}} \beta^{[\lambda]}_{j,k_1,k_2} \langle Rf, \xi^{[\lambda]}_{j,k_1,k_2} \rangle \Psi^{[\lambda]}_{j,k_1,k_2},$$

где $\beta^{[\lambda]}_{j,k_1,k_2} = 2^{j/2}$, а

$$\xi^{[\lambda]}_{j,k_1,k_2}(\varphi, s) = \frac{2^{-j/2}}{4\pi} I^{-1}[R\Psi^{[\lambda]}_{j,k_1,k_2}](\varphi, s),$$

т.е. $\langle f, \Psi^{[\lambda]}_{j,k_1,k_2} \rangle = 2^{j/2} \langle Rf, \xi^{[\lambda]}_{j,k_1,k_2} \rangle$. Таким образом, эта формула может служить основой для методов реконструкции.

При определенных условиях последовательность функций

$$\left\{ \xi^{[\lambda]}_{j,k_1,k_2} \right\}, \quad j, k_1, k_2 \in \mathbb{Z}, \lambda = 1, 2, 3,$$

образует устойчивый базис в $L^2(T)$, где $T = [0, 2\pi] \times R$. Справедлива следующая лемма.

Лемма. Пусть $A > 0$, $a > 1/2$ и $b > 3/2$ такие константы, что $|\widehat{\varphi}(\omega)| \leqslant A(1 + |\omega|^2)^{-b/2}$ и $\left|\widehat{\psi}(\omega)\right| \leqslant A\,|\omega|^a\,(1 + |\omega|^2)^{-(b+a)/2}$ для всех $\omega \in R$, тогда последовательность $\left\{ \xi^{[\lambda]}_{j,k_1,k_2} \right\}$ образует устойчивый базис в $L^2(T)$.

слайд 180

Используя теорему о центральном сечении, можно записать

$$\xi^{[\lambda]}_{j,k_1,k_2}(\varphi, s) = \frac{2^{-j/2}}{4\pi} \int\limits_{-\infty}^{\infty} |\omega|\, \widehat{\Psi}^{[\lambda]}_{j,k_1,k_2}(\omega\cos\varphi, \omega\sin\varphi) e^{is\omega} d\omega.$$

Далее учитывая, что

$$\widehat{\Psi}^{[\lambda]}_{j,k_1,k_2}(\omega_1, \omega_2) = 2^{-j} e^{-i2^{-j}(k_1\omega_1 + k_2\omega_2)} \widehat{\Psi}^{[\lambda]}_{0,0,0}\left(2^{-j}\omega_1, 2^{-j}\omega_2\right),$$

получаем

$$\xi^{[\lambda]}_{j,k_1,k_2}(\varphi, s) = 2^{j/2} \xi^{[\lambda]}_{0,0,0}\left(\varphi, 2^j s - k_1\cos\varphi - k_2\sin\varphi\right).$$

Таким образом, $\xi^{[\lambda]}_{j,k_1,k_2}$ представляют собой «искривленные» сжатия базовых функций $\xi^{[\lambda]}_{0,0,0}$, и при увеличении j они «концентрируются» около некоторых синусоид в плоскости $\{\varphi, s\}$.

Эти свойства преобразования Радона позволяют провести аналогию между ним и однородными линейными преобразованиями. В частности, из вида $\beta^{[\lambda]}_{j,k_1,k_2} = 2^{j/2}$ следует, что степень некорректности задачи обращения преобразования Радона равна $1/2$.

слайд 181

Важным обстоятельством является тот факт, что если вейвлет-функция $\psi(x)$ имеет компактный носитель и достаточно большое число нулевых моментов, то функции $\xi^{[\lambda]}_{0,0,0}(\varphi, s)$, $\lambda = 1, 2, 3$, при каждом фиксированном $\varphi \in [0, 2\pi]$ существенно отличны от нуля только в области примерно такого же радиуса, что и радиус носителя вейвлет-функций $\Psi^{[\lambda]}_{0,0,0}(x, y)$. Следовательно, при «достаточно» больших j коэффициенты в разложении можно вычислить достаточно точно, используя значения $Rf(\varphi, s)$ при $\varphi \in [0, 2\pi)$ и s из интервала радиуса $2^{-j}T$ с центром в точке $2^{-j}(k_1 \cos\varphi + k_2 \sin\varphi)$, где T – радиус области, в которой функции $\Psi^{[\lambda]}_{0,0,0}(x, y)$ отличны от нуля. Эти значения $Rf(\varphi, s)$ представляют собой проекции функции $f(x, y)$ по прямым, проходящим в круге радиуса $2^{-j}T$ с центром в точке $(2^{-j}k_1, 2^{-j}k_2)$. Радиус круга становится меньше с ростом j, т.е. для восстановления более мелких деталей изображения требуется более локальная информация, чем для восстановления крупных деталей. Степень локализации, как и в случае непрерывного вейвлет-преобразования, зависит от скорости убывания функций $I^{-1}[\Psi^{[\lambda]}_{0,0,0}](\varphi, s)$ по переменной s. Зачастую в томографическом эксперименте представляют интерес именно мелкие детали. Это обстоятельство позволяет значительно снизить дозу облучения в процессе получения проекционных данных.

слайд 182

Статистические характеристики коэффициентов разложения

Рассмотрим проекционные данные в виде

$Rf_e(\varphi_m, s_n) = Rf(\varphi_m, s_n) + z_{m,n}$,

$m = 0, \ldots, N-1$, $n = 1, \ldots, N$ $(N = 2^J)$,

где $z_{m,n}$ – независимые случайные величины, имеющие нормальное распределение с нулевым средним и дисперсией σ^2. Для регуляризации метода реконструкции воспользуемся методом пороговой обработки.

В условиях указанной модели эмпирические аналоги коэффициентов в вейвлет-вейглет-разложении с учетом нормировки, аналогичной той, которая связывает дискретные и непрерывные вейвлет-коэффициенты, определяются следующими выражениями:

$X^{[\lambda]}_{j,k_1,k_2} = \mu^{[\lambda]}_{j,k_1,k_2} + \varepsilon^{[\lambda]}_{j,k_1,k_2}$,

где

$\mu^{[\lambda]}_{j,k_1,k_2} = 2^J \langle f, \Psi^{[\lambda]}_{j,k_1,k_2}\rangle = 2^{J+j/2}\langle Rf, \xi^{[\lambda]}_{j,k_1,k_2}\rangle,$
а
$\varepsilon^{[\lambda]}_{j,k_1,k_2}$ – нормально распределенные случайные величины с нулевым средним, которые уже не являются независимыми. Дисперсии случайных величин $\varepsilon^{[\lambda]}_{j,k_1,k_2}$ равны

$$\mathsf{D}\varepsilon^{[\lambda]}_{j,k_1,k_2} = \sigma^2_{\lambda,j} = 4\pi 2^j \sigma^2 \left\|\xi^{[\lambda]}_{0,0,0}\right\|^2.$$

Таким образом, дисперсия растет со скоростью 2^j, что является прямым следствием некорректности задачи реконструкции. Заметим, что в силу однородности преобразования Радона дисперсия не зависит от k_1 и k_2.

слайд 183

Регуляризация с помощью пороговой обработки

Линейная регуляризация с помощью функции окна может чрезмерно сглаживать восстановленное томографическое изображение. Чтобы повысить резкость и лучше восстанавливать мелкие детали изображения, информация о которых обычно содержится в верхней части спектра, можно использовать метод пороговой обработки коэффициентов вейвлет-вейглет-разложения функции изображения.

Напомним, что при пороговой обработке к каждому коэффициенту применяется функция $d_{T_{\lambda,j}}$, которая для жесткой пороговой обработки равна

$$d_{T_{\lambda,j}}(x) = \begin{cases} x, & \text{если } |x| > T_{\lambda,j} \\ 0, & \text{если } |x| \le T_{\lambda,j}, \end{cases}$$

а для мягкой пороговой обработки равна

$$d_{T_{\lambda,j}}(x) = \begin{cases} x - T_{\lambda,j}, & \text{если } x > T_{\lambda,j} \\ x + T_{\lambda,j}, & \text{если } x < -T_{\lambda,j} \\ 0, & \text{если } |x| \le T_{\lambda,j}. \end{cases}$$

Вид дисперсии коэффициентов разложения указывает на то, что порог $T_{\lambda,j}$ должен зависеть от масштаба j и от параметра λ. Например, универсальный порог вычисляется для каждого масштаба j по формуле $T_{\lambda,j} = \sigma_{\lambda,j}\sqrt{2\ln 2^{2j}}$.

слайд 184

На рисунке изображены примеры реконструкции изображения фантома Шеппа-Логана по зашумленным проекциям.

слайд 185

1. Методы вероятностного анализа выборок сигналов/изображений большого объема позволяют строить вероятностные модели, анализировать и прогнозировать течение реальных процессов и явлений, таких как Интернет-трафик, телекоммуникационный трафик, изменение финансовых индексов, турбулентная плазма и прочие.

2. Методы Фурье-анализа используются при анализе/обработке стационарных сигналов/изображений и позволяют осуществлять спектральный анализ, линейную фильтрацию, выделение в сигналах/изображениях диапазонов частот, подавление шума и сжатие больших массивов данных.

слайд 186

3. Вейвлет-анализ дает возможность анализировать эволюцию спектральных характеристик нестационарных сигналов/изображений и предоставляет эффективные инструменты для подавления шумов, сжатия больших объемов данных и локализации особенностей.

4. При решении обратных статистических задач, таких как обращение линейных операторов, применяемых в оптике, радиотехнике, компьютерной графике и других областях, методы вейвлет-анализа дают возможность учитывать не только вид оператора, но и свойства анализируемых сигналов/изображений, такие как разная степень гладкости на разных временных/пространственных участках.

5. Рассмотренные в курсе методы реконструкции томографических изображений позволяют решать такие задачи, как локальная реконструкция участков изображений, линейная и нелинейная регуляризация и оценка качества восстановленных изображений.