

Предупреждение

Типичное математическое высказывание выглядит так: «Теорема. Если то-то, то то-то» или же в другой записи: $\mathcal{A} \longrightarrow \mathcal{B}$. Обучение математике обычно сосредоточено вокруг подробного обсуждения техники логического вывода \longrightarrow одного утверждения из другого, а за рамками обычно остается вопрос появления утверждения \mathcal{A} и того, зачем вообще в нашем мире появление утверждения \mathcal{B} . С другой стороны, все мы знаем из истории, что Математика во многом развивалась вовсе не из взявшихся из ниоткуда формальных утверждений-аксиом, но пыталась отвечать на некоторые изначально не математические вопросы (которые на то время представляли общественный интерес).

Позднее произошло разделение наук и многие математические утверждения стали рассматриваться как самоценные, хотя есть и иные точки зрения на образование, например в формулировке от создателя математических основ классической термодинамики и статистической механики Гиббса: «Математик может позволить себе допустить всё, что ему заблагорассудится, но физика-то здравый смысл покидать не должен».

В этом коротком курсе мы рассмотрим серию математических сюжетов, которые по идее должны прояснять появление и практическое развитие некоторых спектральных методов. При этом постараемся ограничиться несложными темами, которые уже вошли в повседневное употребление, а в общеуниверситетские математические курсы традиционно входят так, что многие приложения остаются скрытыми. В целом этот курс построен вокруг избранных примеров практического использования спектральных методов, вот некоторые из тех, которые возможно изложить, почти не выходя за рамки начальных сведений из математического анализа и линейной алгебры:

- Цифровая обработка сигналов, свертки и фильтрация.
- Цифровая обработка видеообразов. Эффект Гиббса.
- Быстрые алгоритмы умножения, их связь с преобразованием Фурье.
- Компьютерная томография
- Соотношение неопределенностей в его разных формах и связи с квантовой физикой.

Мы начнем с краткого «инженерно-физического» изложения идеи преобразования Фурье (спектрального преобразования), которое встречается во множестве областей Математики. При этом мы не будем подменять или дублировать стандартный подход к изложению, по мере возможности ссылаясь на результаты традиционных курсов математического и функционального анализа.

1 Идея преобразования Фурье

Дискретное преобразование Фурье

Основной объект изучения – последовательности $\mathbf{f} = f(t) = (f_0, f_1, \dots, f_{N-1})$ (на практике это обычно измерения сигнала в равноотстоящих друг от друга моментах времени $t_m = m\Delta t$) конечной длины N . Разумеется, такую запись следует понимать как координаты N -мерного вектора в базисе векторов $\mathbf{e}_k = (0, 0, 1, \dots, 0)$ (единичка стоит на k -том месте). Иными словами, $f(t_m) = \sum_0^{N-1} f_k e_k(t_m)$

- Вспомните, что такое *ортонормальный* базис. Если базис $\mathbf{e}_k = (0, 0, 1, \dots, 0)$ ортонормальный, то длина вектора $(f_0, f_1, \dots, f_{N-1})$ легко находится через его координаты: $\sqrt{\sum f_k^2}$. Квадрат длины вектора в теории сигналов называется *энергией* сигнала.

Для простоты будем пока считать $\Delta t = 1$ так что $t_m = m$, в случае изменения шкалы все формулы без труда переписываются явно (см ниже). Векторы $e_k(m)$ базиса $\{\mathbf{e}_k\}$ естественно называть единичными импульсами, а координаты в этом базисе – представлением сигнала во временной или импульсной области.

- Верно ли, что базис из импульсов ортонормальный?

Преобразование Фурье – это преобразование координат вектора, если вместо импульсного базиса выбрать некоторый другой, который называется *базисом Фурье*. Мотивация выбора нового базиса (Фурье) такова: – он должен быть удобен с той точки зрения, чтобы в нем проще отслеживать большинство физически значимых процессов. Очевидно, длина вектора не зависит от того в каком ортонормированном базисе выписывать его координаты, формула для энергии выглядит одинаково в каждом из них. *Равенством Парсеваля* называется выражение, приравнивающее энергии (то есть квадраты длин) одного и того же сигнала, выписанные в импульсном базисе и в базисе Фурье.

Сигналы записей физических процессов, в случае непрерывного времени как правило описываются на математическом языке как решения некоторых дифференциальных уравнений. Дискретный аналог дифференцирования – это разностный оператор $D = \frac{T-E}{\Delta t}$, здесь E обозначает тождественный оператор, а T – сдвиг на Δt : $T(f_0, f_1, \dots) = (f_1, f_2, \dots)$

Для конечных последовательностей определение оператора сдвига T нуждается в уточнении его действия на концах последовательности, в дискретном конечномерном случае можно «закольцевать» последовательность так чтобы все ее члены расположились в вершинах правильного N -угольника на единичной окружности и оператор T оказался бы поворотом «на одно деление» этого многоугольника:

$$T(f_0, f_1, \dots) = (f_1, f_2, \dots, f_0)$$

Оператор сдвига переводит базисные импульсы друг в друга, под воздействием оператора D координаты f_k произвольного вектора перепутываются весьма сложным образом. Ситуация станет проще, если бы удалось найти такие новые базисные векторы \mathbf{u}_k , что под воздействием оператора D каждый из них просто умножается на *собственное число* λ_k («растягивается» в λ_k раз) – тогда оператор D изменял бы координаты \hat{f}_k (выписанные в базисе $\{\mathbf{u}_k\}$) любого вектора несложно – каждая координата \hat{f}_k умножалась бы на раз и навсегда определенное число $1/\lambda_k$ (подумайте, почему это так). Такой базис для оператора D называется *собственным* и матрица оператора D в этом базисе диагональна.

- Важное наблюдение: если базис собственный для оператора сдвига T , то он собственный и для оператора D . Проверить.

Собственный базис для оператора сдвига T найти можно если использовать комплексные числа: в векторном пространстве комплекснозначных последовательностей имеется, например, такой базис:

$$\mathbf{u}_k \quad \text{это} \quad \left(1, e^{i\frac{2\pi k}{N} \cdot 1}, e^{i\frac{2\pi k}{N} \cdot 2}, \dots, e^{i\frac{2\pi k}{N} \cdot m}, \dots\right)$$

- Найти *собственное число* ω_k , отвечающее *собственному* вектору \mathbf{u}_k оператора T . Как связаны между собой комплексные числа ω_{N-k} и ω_k ?

- Задача: ортогональный ли базис $\{\mathbf{u}_k\}$? Что нужно сделать чтоб он стал ортонормальным?

Координаты \hat{f}_* сигнала (т.е. вектора \mathbf{f}) в новом базисе $\{\mathbf{u}_k\}$ называются *спектром* сигнала, индексы координат можно связать с числами ω_k ; разложение сигнала по новым базисным векторам \mathbf{u}_k называют разложением по собственным гармоникам (частотам), все множество индексов ω_k называют спектральной областью:

$$\mathbf{f} = \sum_0^{N-1} \hat{f}_{\omega_k} \mathbf{u}_{\omega_k}$$

- В ортогональном базисе координаты ищут через перпендикулярные проекции вектора на базисный вектор. Написать формулу для проекции вектора \mathbf{f} на вектор единичной длины в направлении, указанном вектором $\mathbf{u}_k = \left(1, e^{i\frac{2\pi k}{N} \cdot 1}, e^{i\frac{2\pi k}{N} \cdot 2}, \dots, e^{i\frac{2\pi k}{N} \cdot m}, \dots\right)$ (не забывая при этом, что числа в формулах комплексные).

Явные выражения для комплексных чисел ω_k (см. упражнение выше) показывают, что они расположены на комплексной плоскости симметрично относительно вещественной оси, поэтому отвечающие им базисные векторы принято называть *положительными и отрицательными частотами* в соответствии со знаком мнимой части ω_k (можно также ориентироваться на сравнение k с $N/2$), поэтому говорят о положительной и отрицательной частях спектра. Важно напомнить, в случае конечных последовательностей сам сигнал и его спектр следует считать последовательностями длины N , спектральная и временная область в этом случае устроены одинаково. Это преобразование координат принято называть Дискретным Преобразованием Фурье (ДПФ).

Основное тождество Фурье состоит в том, что сигнал \mathbf{f} можно переписать из одного базиса в другой и обратно без потерь информации, вот оно:

$$f(r) = \frac{1}{N} \sum_s \left(\sum_t f(t) e^{-\frac{2\pi i}{N} st} \right) e^{\frac{2\pi i}{N} sr} \quad (1)$$

- Задача: Объяснить появление коэффициента $\frac{1}{N}$ в формуле тождества Фурье

Изменение шкалы времени и необоснованный предельный переход

Будем рассматривать четные значения $N = 2K$ и вместо индексов $0, 1, \dots, N-1$ симметричные индексы в тождестве Фурье (1), а именно, будем считать, что $-K < t, s \leq K$. Для произвольного фиксированного не обязательно целого числа $L > 0$ обозначим $\alpha_k = \frac{Lk}{K}$ тогда $\frac{\alpha_k}{L} = \frac{k}{K}$ и $-L < \alpha_k \leq L$. Введем еще обозначение $d\alpha = \alpha_{k+1} - \alpha_k = \frac{L}{K}$ и проследим за подстановкой новых обозначений в тождество (1):

$$f(r) = \frac{1}{N} \sum_s \left(\sum_t f(t) e^{-\frac{2\pi i}{N} st} \right) e^{\frac{2\pi i}{N} sr} \quad f(\alpha_r) = \frac{1}{2K} \sum_s \left(\sum_t f(\alpha_t) e^{-i\pi s \frac{t}{K}} \right) e^{i\pi s \frac{r}{K}} \quad (2)$$

$$f(\alpha_r) = \frac{1}{2L} \sum_s \left(\sum_t f(\alpha_t) e^{-i\pi s \frac{\alpha_t}{L}} d\alpha \right) e^{i\pi s \frac{\alpha_r}{L}} \quad (3)$$

При $K \rightarrow \infty$ сумма напоминает¹ интегральную и если удастся аккуратно доказать, что

$$\sum_t f(\alpha_t) e^{-i\pi s \frac{\alpha_t}{L}} d\alpha \sim \int_{-L}^L f(\alpha) e^{-\frac{i\pi s \alpha}{L}} d\alpha$$

то возникает известная из справочников формула преобразования Фурье F функций непрерывного аргумента $\alpha \in (-L, L]$

$$F : \mathbf{f} \mapsto \left\{ \frac{1}{2L} \int_{-L}^L f(\alpha) e^{-\frac{i\pi s \alpha}{L}} d\alpha \right\}_{s \in \mathbb{N}} \quad (4)$$

Таким образом, предельная версия дискретного тождества порождает *разложение в ряд Фурье*: грубо говоря, преобразование Фурье в этом случае задает *двойственность* между функциями на сегменте и бесконечными наборами коэффициентов, индексированных порядковыми номерами гармоник. Однако строгое обоснование этой двойственности требует дополнительных условий на функции² и достаточно развернутых доказательств. Традиционно ряды Фурье объясняют в математическом анализе несколько иначе, используя метрики в функциональных пространствах. Указанный в правой части набор чисел называется *коэффициентами ряда Фурье*, в начальных курсах математического анализа их принято выписывать не в комплексном виде, а в виде отдельно вещественной и мнимой части с использованием синусов и косинусов (что заметно усложняет запоминание формулы тождества Фурье).

- Задача: выписать тождество Фурье в случае вещественного сигнала без использования комплексных чисел.

Еще один необоснованный предельный переход

Продолжим в том же «неформально-инженерном» духе: еще одно переобозначение $\beta_s = \frac{\pi s}{L}$ дает в основном (дискретном) тождестве Фурье

$$f(\alpha_r) = \frac{1}{2\pi} \sum_s \left(\sum_t f(\alpha_t) e^{-i\beta_s \alpha_t} d\alpha \right) e^{i\beta_s \alpha_r} d\beta \quad (5)$$

здесь $d\beta = \frac{\pi}{L}$. Выше L было константой, а теперь рассмотрим предельное поведение выражения 5 при $K \rightarrow \infty$ и $L \rightarrow \infty$, опять-таки нетрудно увидеть³ в предельных выражениях соответствующие интегралы. В итоге получается так называемая *интегральная формула тождества Фурье*:

$$f(\gamma) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(\alpha) e^{-i\beta \alpha} d\alpha \right) e^{i\beta \gamma} d\beta \quad (6)$$

Большой вопрос здесь, о каких функциях здесь идет речь, как мы увидим в дальнейшем, желание распространить тождество Фурье приводит к тому, что привычного по математическому анализу запаса отображений $f : \mathbb{R} \rightarrow \mathbb{R}$ **уже не хватает**. Пока отметим, что импульсная

¹это как раз пример «физического рассуждения»

²чем мы никак не озаботились, поэтому аккуратное доказательство не будет коротким

³но доказать будет уже куда труднее

и частотная области здесь опять становятся одинаковыми – в этой версии это просто все вещественные числа. Строгое математическое изложение этого предельного перехода показывает важную разницу между понятием *базиса в линейном векторном пространстве* и бесконечным набором функций из континуального семейства в которое обращается базис Фурье в интегральном случае. Так что название «базис Фурье» – это в самом общем случае определенная вольность речи, которую физики, впрочем, используют не задумываясь. Для упражнения интуиции, попробуйте ответить на два вопроса

- Вопрос: выписать явно какую-нибудь функцию из базиса Фурье в этой версии.
- Трудный вопрос: что такое импульсная функция в импульсном базисе в интегральной версии?

Практический выбор какую область индексов называть импульсной, а какую — спектральной зависит от контекста: например, спектр бесконечного дискретного сигнала следует считать сосредоточенным на единичной окружности (то есть в индексации спектра используется α , в индексации импульсов целые числа – номера отсчетов), а спектр периодического сигнала (то есть опять-таки функции от аргумента α) следует индексировать целыми числами.

Происхождение трудностей при обработке сигналов

Замена переменных и последующий предельный переход по новому параметру заменяют в 1 знак конечной суммы на бесконечную или на интеграл. Важно помнить, что эти предельные переходы и преобразование (двойственность) Фурье даже если все переходы корректно определены все же *не коммутируют: если сделать предельный переход по параметру, а потом от результата взять преобразование Фурье, то получится не вполне то же самое, что при преобразовании Фурье и последующем предельном переходе* — это явление называется алиасингом (калька английского слова *aliasing*).

Практические цифровые алгоритмы обработки сигналов основан на представлении, что мы изучаем дискретными средствами сигналы непрерывного времени, поэтому прикладные формулы порождают множество неожиданных эффектов. Позже мы рассмотрим некоторые из тех, что легко поддаются визуализации.

Многомерные сигналы

Естественным обобщением ранее изложенного подхода служит рассмотрение сигналов, определенных для импульсного пространства, с импульсами закодированными многомерными наборами: в дискретном случае это будут наборы индексов $\mathbf{i} = (i_1, i_2, \dots, i_n)$, $0 \leq i_k < N_k$; в интегральном — $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$.

ДПФ и соответствующее интегральное преобразование в \mathbb{R}^2

Достаточно очевидные обобщения одномерных формул на общий многомерный случай мы обсуждать здесь не будем, а ограничимся лишь двумерным вариантом, поскольку двумерный сигнал в практическом понимании удобно связать с двумерной картинкой — image. Исключительно для простоты мы ограничимся примером формулы ДПФ для квадратных массивов (то есть $N_1 = N_2 = N$) — в конце-концов прямоугольный массив можно считать квадратным, дополняя его нулями.

Дискретный случай

Двумерное ДПФ — это преобразование $F : f(t_1, t_2) \rightarrow \hat{f}(\omega_1, \omega_2)$, заданное очевидным обобщением обычной формулы:

$$\begin{aligned}\hat{f}(\mathbf{s}) = \hat{f}(s_1, s_2) &= \sum_{\mathbf{t}, \mathbf{s}} f(\mathbf{t}) e^{-2\pi i \frac{(\mathbf{t}, \mathbf{s})}{N}} = \sum_{t_2=0}^{N-1} \sum_{t_1=0}^{N-1} f(t_1, t_2) \exp\left(-2\pi i \frac{t_1 s_1 + t_2 s_2}{N}\right) \\ &= \sum_{t_2=0}^{N-1} \left(\sum_{t_1=0}^{N-1} f(t_1, t_2) e^{-2\pi i \frac{t_1 s_1}{N}} \right) e^{-2\pi i \frac{t_2 s_2}{N}} \\ &= \sum_{t_1=0}^{N-1} \left(\sum_{t_2=0}^{N-1} f(t_1, t_2) e^{-2\pi i \frac{t_2 s_2}{N}} \right) e^{-2\pi i \frac{t_1 s_1}{N}}\end{aligned}$$

- Запишите обратное преобразование Фурье и тождество Фурье для этого случая

Практические навыки

Используя четные продолжения функций, мы приходим к специальному виду преобразования Фурье, в котором от комплексных экспонент в формуле останутся только косинусы. Соответствующее преобразование называется дискретным косинус-преобразованием, оно дало толчок развитию способов компактного архивирования (с потерями) визуальных образов. Идея заключается в том, что вклады высоких частот наш глаз воспринимает не так, как вклады низких — можно попробовать вообще не учитывать вклады высоких частот в визуальный образ. В современных программных пакетах достаточно много библиотек для обработки двумерных сигналов и в принципе начать можно с любого из них. Соответствующие функции преобразования Фурье в python можно найти в `np.fft.fft2`, `scipy.fftpack.dct`, обязательно внимательно ознакомьтесь с версией нормализующих коэффициентов при прямых и обратных преобразованиях.

Для нас чуть позже понадобится пакет `scikit-image` <https://scikit-image.org/docs> языка python следует ознакомиться с некоторыми функциями, реализующими обработку образов, так как наш интерес в этом курсе касается построения *синограмм*.

Преобразование Фурье и соотношение неопределенности

Существуют несколько разных формулировок принципа неопределенности, наиболее широко известна имеющая отношение к физике формулировка: принципиально невозможно одновременно точно определить координату и скорость микрочастицы. На самом деле, основные формулировки имеют отношение к свойствам сигнала и его преобразования Фурье и мы начнем с гораздо менее известного дискретного случая, сравнительно недавно сформулированного для обработки наблюдений за разреженными (sparse, то есть имеющими много нулевых значений) сигналами. Для удобства дальнейших формул выберем перенормированные формулы прямого и обратного дискретного преобразования Фурье:

$$\hat{f}(s) = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} f(t) e^{-2\pi i t s / N} \quad (1)$$

$$f(t) = \frac{1}{\sqrt{N}} \sum_{s=0}^{N-1} \hat{f}(s) e^{2\pi i s t / N} \quad (2)$$

Коэффициент в равенстве Парсеваля, тем самым, исчезает: $\sum_t |f(t)|^2 = \sum_s |\hat{f}(s)|^2$.

Интересный пример

Преобразование Фурье очевидным образом линейно.

1. Запишите матрицу преобразования Фурье для сигналов дискретного времени длины N . Какой у нее определитель?
2. Пусть $N = M^2$, рассмотрите в импульсной области сигнал дискретного времени, отличный от нуля только в моменты времени $0.M, 2M, \dots$ и равный в них соответственно M . Найдите его преобразование Фурье.

Соотношение неопределенности, дискретный сигнал на \mathbb{Z}/N

Для ненулевого сигнала f выполняется $\boxed{\text{supp } f \times \text{supp } \hat{f} \geq N}$. Действительно, в силу тождества Фурье

$$\max_t |f(t)| \leq \frac{1}{\sqrt{N}} \sum_{s=0}^{N-1} |\hat{f}(s)|$$

и, кроме того

$$\begin{aligned} \sum_{t=0}^{N-1} |f(t)|^2 &\leq |\text{supp } f| \cdot \max_t |f(t)|^2 \\ \left(\sum_{s=0}^{N-1} |\hat{f}(s)| \right)^2 &\leq |\text{supp } \hat{f}| \cdot \sum_{s=0}^{N-1} |\hat{f}(s)|^2 \end{aligned}$$

где второе неравенство — неравенство Коши для \hat{f} и $\mathbf{1}_{\text{supp } \hat{f}}$. Подставляя одно в другое:

$$\sum |f|^2 \leq |\text{supp } f| \cdot \max_t |f|^2 \leq |\text{supp } f| \cdot \frac{1}{N} \cdot \left(\sum |\hat{f}| \right)^2 \leq |\text{supp } f| \cdot \frac{1}{N} |\text{supp } \hat{f}| \cdot \sum |\hat{f}|^2$$

Осталось воспользоваться равенством Парсеваля. ■

Версия дискретного соотношения неопределенности для простого числа N

Для простого числа N возможна иная формулировка: $|\text{supp} f| + |\text{supp} \hat{f}| \geq N + 1$. Для объяснения потребуется теорема Чеботарева¹ о невырожденности любого квадратного минора в матрице с элементами $a_{ts} = e^{-2\pi i ts/N}$ при простом N . Понятно, что умножение сигнала на указанную матрицу как раз задает (с точностью до коэффициента) преобразование Фурье; рассмотрим теперь сигнал f , у которого $|\text{supp} f| = S$ и выберем соответственно столбцы с номерами, отвечающими S позициям этого носителя. Поскольку по теореме Чеботарева любые миноры на этих столбцах невырождены, то среди N значений \hat{f} нулей заведомо строго меньше, чем S , а значит $|\text{supp} \hat{f}| > N - S$. ■

Применение соотношения неопределенности: восстановление разреженного дискретного сигнала по части его спектра

Естественной характеристикой разреженности служит величина $|\text{supp} f|$: S -разреженные сигналы f — это те, у которых $|\text{supp} f| \leq S$. Сколько нужно знать позиций в \hat{f} , чтобы этот сигнал восстанавливался однозначно? Понятно, что при $2S < N$ пространство всех комплекснозначных сигналов на $0, 1, \dots, 2S - 1$ имеет размерность в точности $2S$, разделяя сигнал на два разных S -разреженных сигнала f и g , где f ненулевой только в позициях $0, 1, \dots, S - 1$, а g соответственно в позициях $S, S + 1, \dots, 2S - 1$, то можно указать такую пару, что значения спектров отличаются всего в одной частоте. А значит, однозначное восстановление S -разреженного сигнала — вообще говоря — невозможно, если рассматривать менее чем $2S$ значений его Фурье-образа. Выше было показано, что восстановление разреженного сигнала удобнее реализовать при простом N . На самом деле, в общем случае ситуация зависит еще и от того, какие именно позиции задействованы в S -разреженном сигнале.

Замечание

Для $N = M^2$ выше был построен такой сигнал f , чтобы $f = \hat{f}$. Заметим, что в этом примере при $M \gg 1$ даже весьма большое число $N - M - 1 > 2M$ не поможет отличить M -разреженный сигнал f от нуля! В случае простого числа N такой контрпример невозможен, конечно.

Таким образом неясно, как в общем случае искать разреженный сигнал не используя перебор всех вообще решений, а этот перебор нереально трудоемок. Действительно, можно просто решать систему $K \geq 2S$ уравнений с S неизвестными, благо для простого N в силу теоремы Чеботарева необходимые матрицы невырождены, но так как при этом вариантов решений много, то и неясно, как найти среди них S -разреженное (которое обеспечивает единственность)?

О доказательстве теоремы Чеботарева

См. на следующих страницах копии рассуждений из задачника В.Прасолова по линейной алгебре.

¹Доказательство теоремы Чеботарева никак не связано с преобразованием Фурье и опирается на достаточно известные свойства алгебры многочленов

жащих ее угловой минор порядка $m-1$. Очевидно, что $S_{m,n}^r = a_{11} \bar{S}_{m-1,n-1}^{r-1}$, причем к матрице $\bar{S}_{m-1,n-1}^{r-1}$ можно применить предположение индукции (случай $m-1=0$ был рассмотрен отдельно). Кроме того, если \bar{A}_{m-1} и \bar{A}_{n-1} — угловые миноры порядка $m-1$ и $n-1$ матрицы \bar{A} , то $A_m = a_{11} \bar{A}_{m-1}$ и $A_n = a_{11} \bar{A}_{n-1}$. Поэтому

$$|S_{m,n}^r| = a_{11}^t \bar{A}_{m-1}^{p_1} \bar{A}_{n-1}^{q_1} = a_{11}^{t-p_1-q_1} A_m^{p_1} A_n^{q_1},$$

где $t := C_{n-m}^{r-m}$, $p_1 = C_{n-m-1}^{r-m} = p$ и $q_1 = C_{n-m-1}^{r-m-1} = q$. Учтя, что $t = p + q$, получаем требуемое.

2.8. Теорема (Чоботарев). Пусть p — простое число и $\varepsilon = \exp(2\pi i/p)$. Тогда все миноры матрицы Вандермонда $\|a_{ij}\|_0^{p-1}$, где $a_{ij} = \varepsilon^{ij}$, отличны от нуля.

Доказательство [Решетняк, 1955]. Предположим, что

$$\begin{vmatrix} \varepsilon^{h_1 l_1} & \dots & \varepsilon^{h_1 l_j} \\ \varepsilon^{h_2 l_1} & \dots & \varepsilon^{h_2 l_j} \\ \dots & \dots & \dots \\ \varepsilon^{h_j l_1} & \dots & \varepsilon^{h_j l_j} \end{vmatrix} = 0.$$

Тогда существуют такие комплексные числа c_1, \dots, c_j , не все равные 0, что линейная комбинация столбцов с коэффициентами c_1, \dots, c_j равна нулю, т. е. числа $\varepsilon^{h_1}, \dots, \varepsilon^{h_j}$ являются корнями многочлена $c_1 x^{l_1} + \dots + c_j x^{l_j}$. Пусть

$$(x - \varepsilon^{h_1}) \dots (x - \varepsilon^{h_j}) = x^j - b_1 x^{j-1} + \dots \pm b_j. \quad (1)$$

Тогда

$$c_1 x^{l_1} + \dots + c_j x^{l_j} = (b_0 x^j - b_1 x^{j-1} + \dots \pm b_j)(a_s x^s + \dots + a_0), \quad (2)$$

где $b_0 = 1$ и $a_s \neq 0$. Для удобства будем считать, что $b_t = 0$ при $t > j$ и $t < 0$. Коэффициент при x^{j+s-t} в правой части уравнения (2) равен $\pm(a_s b_t - a_{s-1} b_{t-1} + \dots \pm a_0 b_{t-s})$. Степень многочлена (2) равна $s+j$, причем лишь коэффициенты при степенях l_1, \dots, l_j могут быть ненулевыми, а значит, найдется $s+1$ нулевых коэффициентов: $a_s b_t - a_{s-1} b_{t-1} + \dots \pm a_0 b_{t-s} = 0$ при $t = t_0, t_1, \dots, t_s$. Числа a_0, \dots, a_{s-1}, a_s не все нулевые, поэтому $|c_{kl}|_0^s = 0$ для $c_{kl} = b_{t-l}$. Пусть $t_k - l = \tau$. Формула (1) показывает, что b_τ можно представить в виде многочлена $f_\tau(\varepsilon)$ с целыми коэффициентами, причем этот многочлен является сум-

мой C_j^t степеней числа ε , а значит, $f_\tau(1) = C_j^t$. Так как $c_{ji} = b_i = f_i(\varepsilon)$, то $|c_{ij}|_0^s = g(\varepsilon)$, причем $g(1) = |c'_{ij}|_0^s$, где $c'_{ij} = C_j^{i-l}$. Многочлен $q(x) = x^{p-1} + \dots + x + 1$ неприводим над \mathbb{Q} (см. приложение, п. 2, теорема 2) и $q(\varepsilon) = 0$. Следовательно, $g(x) = q(x)\varphi(x)$, где φ — многочлен с целыми коэффициентами (см. приложение, п. 2, теорема 1). Поэтому $g(1) = q(1)\varphi(1) = p\varphi(1)$, т. е. $g(1)$ делится на p .

Чтобы прийти к противоречию, достаточно доказать, что число $g(1) = |c'_{ij}|_0^s$, где $c'_{ij} = C_j^{i_k-l}$, $0 \leq i_k \leq j+s$ и $0 < j+s \leq p-1$, не делится на p . Легко проверить, что $\Delta = |c'_{ij}|_0^s = |a_{ij}|_0^s$, где $a_{kl} = C_{j+l}^{i_k}$ (см. задачу 1.27), и

$$C_{j+l}^{i_k} = \left(1 - \frac{t}{j+l+1}\right) \cdots \left(1 - \frac{t}{j+s}\right) C_{j+s}^{i_k} = \varphi_{s-l}(t) C_{j+s}^{i_k}.$$

Следовательно,

$$\begin{aligned} \Delta &= \prod_{\lambda=0}^s C_{j+s}^{t_\lambda} \begin{vmatrix} \varphi_s(t_0) & \varphi_{s-1}(t_0) & \cdots & 1 \\ \varphi_s(t_1) & \varphi_{s-1}(t_1) & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_s(t_s) & \varphi_{s-1}(t_s) & \cdots & 1 \end{vmatrix} = \\ &= \pm \prod_{\lambda=0}^s (C_{j+s}^{t_\lambda} A_\lambda) \prod_{\mu > \nu} (t_\mu - t_\nu), \end{aligned}$$

где A_0, A_1, \dots, A_s — коэффициенты при наивысших степенях t в многочленах $\varphi_0(t) \equiv 1, \varphi_1(t), \dots, \varphi_s(t)$; степень многочлена $\varphi_i(t)$ равна i . Ясно, что в полученном произведении нет несократимых дробей с числителями, делящимися на p .

ЗАДАЧИ

2.1. Пусть A — матрица размера $n \times n$. Докажите, что $|A + \lambda E| = \lambda^n + \sum_{k=1}^n I_k \lambda^{n-k}$, где I_k — сумма всех C_n^k главных миноров k -го порядка матрицы A .

2.2. Докажите, что

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} & x_1 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & x_n \\ y_1 & \cdots & y_n & 0 \end{vmatrix} = - \sum_{i,j} x_i y_j A_{ij},$$

Приложение

НЕКОТОРЫЕ СВОЙСТВА МНОГОЧЛЕНОВ

1. Теорема. Пусть $f(x_1, \dots, x_n) = \sum a_{k_1, \dots, k_n} x_1^{k_1} \dots x_n^{k_n}$, причем $f(x_1, \dots, x_n) = 0$ для любого набора чисел x_1, \dots, x_n . Тогда все коэффициенты многочлена f равны нулю.

Доказательство. Ненулевой многочлен от одной переменной имеет конечное число корней, поэтому при $n = 1$ утверждение очевидно. Предположим, что утверждение верно для любого многочлена от $n - 1$ переменных. Многочлен f можно записать в виде

$$f(x_1, \dots, x_n) = x_n^k f_k(x_1, \dots, x_{n-1}) + \dots + f_0(x_1, \dots, x_{n-1}).$$

Многочлен $x_n^k f_k + \dots + f_0$, где $a_i = f_i(x_1, \dots, x_{n-1})$, принимает только нулевые значения, поэтому $f_i(x_1, \dots, x_{n-1}) = 0$, а значит, все коэффициенты многочлена $f_i(x_1, \dots, x_{n-1})$ равны нулю.

З а м е ч а н и е. Теорема верна лишь для полей, содержащих бесконечно много элементов. Ненулевой многочлен $x^p - x$ над полем \mathbb{F}_p принимает только нулевые значения.

2. Многочлен f с целочисленными коэффициентами называется *неприводимым* (над кольцом целых чисел), если его нельзя представить в виде произведения двух многочленов меньшей степени с целочисленными коэффициентами.

Теорема 1. Пусть многочлены f и g с целыми коэффициентами имеют общий корень, причем многочлен f неприводим. Тогда g/f — многочлен.

Доказательство. Будем последовательно производить деление с остатком: $g = a_1 f + b_1$, $f = a_2 b_1 + b_2$, $b_1 = a_3 b_2 + b_3$, ..., $b_{n-2} = a_{n-1} b_n$. Легко проверить, что b_n — наибольший общий делитель многочленов f и g . Все многочлены a_i и b_i имеют целые коэффициенты. Значит, наибольший общий делитель многочленов f и g над кольцом целых чисел совпадает с их наибольшим общим делителем над полем \mathbb{C} . Но над полем \mathbb{C} многочлены f и g имеют нетривиальный общий делитель, а значит, многочлены f и g имеют нетривиальный общий делитель r и над кольцом целых чисел. Так как многочлен f неприводим, то $r = \lambda f$, где λ — некоторое число.

Теорема 2 (критерий Эйзенштейна). Пусть $f(x) = a_0 + a_1 x + \dots + a_n x^n$ — многочлен с целыми коэффициентами, причем для некоторого простого числа p коэффициент a_n не делится на p , a_0, \dots, a_{n-1} делятся на p и a_0 не делится на p^2 . Тогда f — неприводимый многочлен.

Доказательство. Предположим, что

$$f = gh = \left(\sum b_k x^k \right) \left(\sum c_l x^l \right),$$

причем g и h не константы. Число $b_0 c_0 = a_0$ делится на p , поэтому одно из чисел b_0 и c_0 делится на p . Пусть, для определенности, b_0 делится на p . Тогда c_0 не делится на p , так как $a_0 = b_0 c_0$ не делится на p^2 . Если все числа b_i делятся на p , то a_n делится на p . Поэтому b_i не делится на p при некотором i , где $0 < i \leq \deg g < n$; можно считать, что i — наименьший номер числа b_i , не делящегося на p . С одной стороны, по условию число a_i делится на p . С другой стороны, $a_i = b_i c_0 + b_{i-1} c_1 + \dots + b_0 c_i$, причем все числа $b_{i-1} c_1, \dots, b_0 c_i$ делятся на p , а число $b_i c_0$ не делится на p . Получено противоречие.

Следствие. Если p — простое число, то многочлен $f(x) = x^{p-1} + \dots + x + 1$ неприводим.

В самом деле, к многочлену

$$f(x+1) = \frac{(x+1)^p - 1}{(x+1) - 1} = x^{p-1} + C_p^1 x^{p-2} + \dots + C_p^{p-1}$$

можно применить критерий Эйзенштейна.

3. Теорема (Эрмит). Пусть в точках x_1, \dots, x_n заданы числа $y_1^{(0)}, y_1^{(1)}, \dots, y_1^{(\alpha_1-1)}, \dots, y_n^{(0)}, y_n^{(1)}, \dots, y_n^{(\alpha_n-1)}$; $m = \alpha_1 + \dots + \alpha_n - 1$. Тогда существует многочлен $H_m(x)$ степени не более m , для которого $H_m(x_j) = y_j^{(0)}$ и $H_m^{(i)}(x_j) = y_j^{(i)}$.

Доказательство. Пусть $k = \max(\alpha_1, \dots, \alpha_n)$. При $k = 1$ можно воспользоваться интерполяционным многочленом Лагранжа

$$L_n(x) = \sum_{j=1}^n \frac{(x-x_1) \dots (x-x_{j-1})(x-x_{j+1}) \dots (x-x_n)}{(x_j-x_1) \dots (x_j-x_{j-1})(x_j-x_{j+1}) \dots (x_j-x_n)} y_j^{(0)}.$$

Пусть $\omega_n(x) = (x-x_1) \dots (x-x_n)$. Возьмем произвольный многочлен H_{m-n} степени не более $m-n$ и сопоставим ему многочлен $H_m(x) = L_n(x) + \omega_n(x) H_{m-n}(x)$. Ясно, что $H_m(x_j) = y_j$ при любом многочлене H_{m-n} . Кроме того,

$$H_m'(x) = L_n'(x) + \omega_n'(x) H_{m-n}(x) + \omega_n(x) H_{m-n}'(x),$$

т. е. $H_m'(x_j) = L_n'(x_j) + \omega_n'(x_j) H_{m-n}(x_j)$. А так как $\omega_n'(x_j) \neq 0$, то в точках, в которых заданы значения $H_m'(x_j)$, можно определить соответствующие значения $H_{m-n}(x_j)$. Далее,

$$H_m''(x_j) = L_n''(x_j) + \omega_n''(x_j) H_{m-n}(x_j) + 2\omega_n'(x_j) H_{m-n}'(x_j).$$

Поэтому в точках, в которых заданы значения $H_m''(x_j)$, можно определить соответствующие значения $H_{m-n}'(x_j)$ и т. д. Таким образом, наша задача сводится к построению многочлена $H_{m-n}(x)$.

Соотношения между определителями

1.23. Пусть $b_{ij} = (-1)^{i+j} a_{ij}$. Докажите, что $|a_{ij}|_1^n = |b_{ij}|_1^n$.

1.24. Докажите, что

$$\begin{vmatrix} a_1 c_1 & a_2 d_1 & a_1 c_2 & a_2 d_2 \\ a_3 c_1 & a_4 d_1 & a_3 c_2 & a_4 d_2 \\ b_1 c_3 & b_2 d_3 & b_1 c_4 & b_2 d_4 \\ b_3 c_3 & b_4 d_3 & b_3 c_4 & b_4 d_4 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ a_3 & a_4 \end{vmatrix} \cdot \begin{vmatrix} b_1 & b_2 \\ b_3 & b_4 \end{vmatrix} \cdot \begin{vmatrix} c_1 & c_2 \\ c_3 & c_4 \end{vmatrix} \cdot \begin{vmatrix} d_1 & d_2 \\ d_3 & d_4 \end{vmatrix}.$$

1.25. Докажите, что

$$\begin{vmatrix} a_1 & 0 & 0 & b_1 & 0 & 0 \\ 0 & a_2 & 0 & 0 & b_2 & 0 \\ 0 & 0 & a_3 & 0 & 0 & b_3 \\ b_{11} & b_{12} & b_{13} & a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} & a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} & a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_1 a_{11} - b_1 b_{11} & a_2 a_{12} - b_2 b_{12} & a_3 a_{13} - b_3 b_{13} \\ a_1 a_{21} - b_1 b_{21} & a_2 a_{22} - b_2 b_{22} & a_3 a_{23} - b_3 b_{23} \\ a_1 a_{31} - b_1 b_{31} & a_2 a_{32} - b_2 b_{32} & a_3 a_{33} - b_3 b_{33} \end{vmatrix}.$$

1.26. Пусть $s_k = \sum_{i=1}^n a_{ki}$. Докажите, что

$$\begin{vmatrix} s_1 - a_{11} & \dots & s_1 - a_{1n} \\ \dots & \dots & \dots \\ s_n - a_{n1} & \dots & s_n - a_{nn} \end{vmatrix} = (-1)^{n-1} (n-1) \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

1.27. Докажите, что

$$\begin{vmatrix} C_n^{m_1} & C_n^{m_1-1} & \dots & C_n^{m_1-k} \\ \dots & \dots & \dots & \dots \\ C_n^{m_k} & C_n^{m_k-1} & \dots & C_n^{m_k-k} \end{vmatrix} = \begin{vmatrix} C_n^{m_1} & C_{n+1}^{m_1} & \dots & C_{n+k}^{m_1} \\ \dots & \dots & \dots & \dots \\ C_n^{m_k} & C_{n+1}^{m_k} & \dots & C_{n+k}^{m_k} \end{vmatrix}.$$

1.28. Пусть $\Delta_n(k) = |a_{ij}|_0^n$, где $a_{ij} = C_{k+i}^{2j}$. Докажите, что

$$\Delta_n(k) = \frac{k(k+1) \dots (k+n-1)}{1 \cdot 3 \dots (2n-1)} \Delta_{n-1}(k-1).$$

1.29. Пусть $D_n = |a_{ij}|_0^n$, где $a_{ij} = C_{n+1}^{2j-1}$. Докажите, что $D_n = 2^{n(n+1)/2}$.

1.26. Для определителей матриц порядка $n+1$ легко проверить следующие равенства:

$$\begin{vmatrix} s_1 - a_{11} & \dots & s_1 - a_{1n} & 0 \\ \dots & \dots & \dots & \dots \\ s_n - a_{n1} & \dots & s_n - a_{nn} & 0 \\ -1 & \dots & -1 & 1 \end{vmatrix} = \begin{vmatrix} s_1 - a_{11} & \dots & s_1 - a_{1n} & (n-1)s_1 \\ \dots & \dots & \dots & \dots \\ s_n - a_{n1} & \dots & s_n - a_{nn} & (n-1)s_n \\ -1 & \dots & -1 & 1-n \end{vmatrix} =$$

$$= (n-1) \begin{vmatrix} s_1 - a_{11} & \dots & s_1 - a_{1n} & s_1 \\ \dots & \dots & \dots & \dots \\ s_n - a_{n1} & \dots & s_n - a_{nn} & s_n \\ -1 & \dots & -1 & -1 \end{vmatrix} =$$

$$= (n-1) \begin{vmatrix} -a_{11} & \dots & -a_{1n} & s_1 \\ \dots & \dots & \dots & \dots \\ -a_{n1} & \dots & -a_{nn} & s_n \\ 0 & \dots & 0 & -1 \end{vmatrix}$$

(для получения первого равенства нужно к последнему столбцу прибавить сумму всех остальных, а для получения последнего равенства нужно последний столбец вычесть из всех остальных).

1.27. Воспользуемся тождеством $C_p^q + C_p^{q-1} = C_{p+1}^q$. Складывая соответствующим образом столбцы, от матрицы со строками вида $(C_n^m, C_n^{m-1}, \dots, C_n^{m-k})$ можно перейти к матрице со строками $(C_n^m, C_{n+1}^m, C_{n+1}^{m-1}, \dots, C_{n+1}^{m-k+1})$. Прodelывая аналогичную операцию для столбцов с номерами 2, 3, ..., n , получим матрицу со строками $(C_n^m, C_{n+1}^m, C_{n+2}^m, C_{n+2}^{m-1}, \dots, C_{n+2}^{m-k+2})$. После нескольких таких операций в итоге получим матрицу со строками $(C_n^m, C_{n+1}^m, \dots, C_{n+k}^m)$.

1.28. Вычтем в определителе $\Delta_n(k)$ из строки с номером $i+1$ строку с номером i для $i = n, n-1, \dots, 1$. В результате получим $\Delta_n(k) = \Delta'_{n-1}(k)$, где $\Delta'_m(k) = |a'_{ij}|_m$, $a'_{ij} = C_{k+i}^{2j+1}$. А так как $C_{k+i}^{2j+1} = \frac{k+i}{2j+1} C_{k-1+i}^{2j}$, то

$$\Delta'_{n-1}(k) = \frac{k(k+1) \dots (k+n-1)}{1 \cdot 3 \dots (2n-1)} \Delta_{n-1}(k-1).$$

1.29 [Карлиц, 1957]. Согласно задаче 1.27 $D_n = D'_n = |a'_{ij}|_n$, где $a'_{ij} = C_{n+1+i}^{2j}$, т. е. в обозначениях задачи 1.28 получаем

$$D_n = \Delta_n(n+1) = \frac{(n+1)(n+2) \dots 2n}{1 \cdot 3 \dots (2n-1)} \Delta_{n-1}(n) = 2^n D_{n-1},$$

так как $(n+1)(n+2) \dots 2n = (2n)!/n!$, $1 \cdot 3 \dots (2n-1) = (2n)!/2 \cdot 4 \dots 2n$ и $\Delta_{n-1}(n) = D_{n-1}$. А так как $D_0 = 1$, то $D_n = 2^k$, где $k = n + (n-1) + \dots + 1 = n(n+1)/2$.

Топологические абелевы группы и преобразования Фурье

Ранее было замечено, что предельные переходы в дискретном тождестве Фурье для функций определенных на \mathbb{Z}/N приводят к следующим геометрическим объектам: \mathbb{Z} , \mathbb{R} , S^1 , на каждом из которых имеется структура абелевой группы. В дискретном случае речь шла о произвольных комплекснозначных сигналах т.е. функциях $f: \mathbb{Z}/N \rightarrow \mathbb{C}$ и обратимости для них преобразования Фурье, но оба этих свойства становятся неочевидными из-за предельных переходов. С другой стороны на каждый такой объект обладает структурой *топологической абелевой группы*¹. В этом разделе приведены факты о пространствах гомоморфизмов на этих конкретных группах: это описание на самом деле ведет к глубокому обобщению известной теоремы о строении конечно-порожденных абелевых групп. При этом условие конечной порожденности абелевой группы заменяется на условие локальной компактности и возникает новая общая теорема о двойственности (теорема Понтрягина-Ван Кампена). Интересующиеся формальной стороной вопроса могут посмотреть несложную книжку [Моррис (1980)]. Следуя традиции, мы будем использовать обозначение \mathbb{T} , при этом $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ соответствующая проекция $\mathbb{R} \rightarrow \mathbb{T}$ определяется взятием дробной части $\{x\} \in \mathbb{T}$ для $x \in \mathbb{R}$. При этом обозначение S^1 сохраняем для окружности единичного радиуса, которую удобно считать состоящей из комплексных чисел по модулю равных единице вида e^{iy} , $0 \leq y < 2\pi$.

Предложение

Недискретная подгруппа в \mathbb{R} плотна в ней. Действительно, надо показать, что в ε -окрестности любой точки $x \in \mathbb{R}$ есть элемент g из подгруппы G . По недискретности найдем сначала соответствующую точку в $\frac{\varepsilon}{2}$ -окрестности 0, а потом рассматривая элементы $\pm ng$ доберемся и до ε -окрестности любой точки $x \in \mathbb{R}$.

Следствие: Замкнутая подгруппа в \mathbb{R} совпадает либо с \mathbb{R} либо с $a\mathbb{Z}$ для некоторого $a \in \mathbb{R}$

Следствие: Для рационально не зависимых $a, b \in \mathbb{R}$ замыкание порожденной ими подгруппы совпадает с \mathbb{R} .

Следствие: Любая собственная замкнутая подгруппа в \mathbb{T} конечна. Указание: рассмотрите проекцию $\mathbb{R} \rightarrow \mathbb{T}$. Более того, любая собственная замкнутая подгруппа в \mathbb{T} циклическая, поскольку ее прообраз при канонической проекции $\mathbb{R} \rightarrow \mathbb{T}$ циклическая группа.

Характеры и двойственность

Определение Если G — абелева топологическая группа, то непрерывный гомоморфизм $g: G \rightarrow \mathbb{T}$ называется характером. Очевидно, что использование групповой операции в \mathbb{T} превращает набор всех характеров называется в группу, она называется двойственной к G группой и обозначается через G^* . Топология (компактно-открытая) в G^* вводится стандартным для пространства отображений образом.

Рассмотрим группу \mathbb{Z} . Каждый характер определяется значением $g(1)$, который может быть произвольным элементом группы \mathbb{T} . По любому элементу $a \in \mathbb{T}$ построим характер g , для которого $g(1) = a$ — это задает алгебраический изоморфизм группы \mathbb{Z}^* и \mathbb{T} . Проверьте, что этот изоморфизм является вдобавок гомеоморфизмом соответствующих топологических пространств.

Рассмотрим группу \mathbb{T} , тогда любой ее характер g определяется некоторым целым числом m так что $g(x) = mx$. Действительно, рассматриваем ядро гомоморфизма, которое должно быть компактным и замкнутым, то есть конечной подгруппой в \mathbb{T} . Воспользовавшись отождествлением \mathbb{T} с $[0, 1)$ мы выбираем наименьший элемент y ядра, таким образом, для любого $x \in [0, y)$ $g(x) \neq 0$, при этом y является рациональным числом $\frac{1}{m}$ (см. выше: ядро — циклическая подгруппа). Ясно, что умножение на m задает гомеоморфизм $[0, y) \rightarrow [0, 1)$ вдобавок являющийся гомоморфизмом.

1. Сколько элементов порядка 2 есть в \mathbb{T} ?

2. Покажите, что если g — непрерывный взаимно однозначный гомоморфизм группы \mathbb{T} в себя, то либо $\forall x \in \mathbb{T} g(x) = x$, либо $\forall x \in \mathbb{T} g(x) = -x$.

Осталось заметить, что m сдвигов на рациональное число $\frac{1}{m}$ полуинтервала $[0, y)$ дают $[0, 1)$ и $g(x + \frac{1}{m}) = g(x)$. Таким образом, двойственная к \mathbb{T} группа \mathbb{T}^* алгебраически изоморфна \mathbb{Z}

¹Напомним, что топологическое пространство G с групповой структурой называется топологической группой, если групповая операция $G \times G \rightarrow G$ и обращение $G \rightarrow G$ являются непрерывными.

Рассмотрим группу \mathbb{R} . Каждый характер g группы \mathbb{R} можно записать в виде $g(x) = e^{2\pi i c x}$ где $c \in \mathbb{R}$ — фиксированно для данного g и $x \in \mathbb{R}$. Это утверждение означает, что $\mathbb{R}^* = \mathbb{R}$. Доказательство использует каноническую проекцию $\mathbb{R} \rightarrow \mathbb{T}$ и рассуждение из предыдущего раздела. Продумайте его самостоятельно.

Вопросы для самопроверки

1. Покажите, что \mathbb{R} топологически изоморфна двойственной к ней группе.
2. Найдите группы, двойственные к дискретным конечным циклическим группам.

Равномерные последовательности и повороты окружности

Продолжим изучение свойств поворота окружности на иррациональный (в градусах) угол, что соответствует сдвигу $R : \mathbb{T} \rightarrow \mathbb{T} \ x \mapsto x + r$ в группе \mathbb{T} на иррациональный элемент. Очевидным образом сдвиг сохраняет метрику, также из изложенного ранее понятно, что траектория $R(x) = x, R^1(x) = x + r, R^2(x), \dots$ всюду плотна в \mathbb{T} . Аналогично можно изучать последовательности дробных частей $\{k \cdot x\}$ для иррационального вещественного x и $k = 0, 1, 2, \dots$, встречаются в нескольких разделах математики.

Выражение

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(R^k(x))$$

называется *временным средним*².

Определение Последовательность точек x_k на окружности \mathbb{T} называется *равномерно распределенной* если для любой дуги ds окружности предел частоты $\frac{\nu(ds)}{n}$ попаданий в эту дугу среди x_1, x_2, \dots, x_n

$$\lim_{n \rightarrow \infty} \frac{\nu(ds)}{n} = |ds|$$

Определение равномерно распределенной последовательности без труда переформулируется для любой окружности.

Теорема Боля-Серпинского-Вейля: пусть R — поворот окружности $S^1 = \{z \in \mathbb{C}, |z| = 1\}$ на угол, не соизмеримый с 2π : $R(z) = e^{2\pi i \omega} z = \theta \cdot z$, где ω — иррациональное число и f — непрерывная функция на S^1 (на самом деле доказательство проходит и для функции, интегрируемой по Риману). Тогда временное среднее функции f существует и совпадает с пространственным (если нормализовать на длину окружности) средним.

Доказательство 1-й случай: $f(z) = z^p$, $p \in \mathbb{Z}$, тогда

$$\frac{1}{n} \sum_{k=0}^{n-1} f(R^k(z)) = \frac{1}{n} \sum_{k=0}^{n-1} (\theta^k z)^p = \begin{cases} 1 & p = 0 \\ \frac{1}{n} z^p \frac{\theta^{np} - 1}{\theta^p - 1} & p \neq 0 \end{cases}$$

Так как число ω иррационально, мы имеем

$$\theta^p - 1 \neq 0 \quad \text{и} \quad |\theta^{np} - 1| < 2$$

то есть временное и пространственное среднее совпадают и равны 1 при $p = 0$ и 0 при $p \neq 0$.

2-й случай: $f(z) = a_0 + a_1 z + \dots + a_m z^m$ — тригонометрический полином, аналогичными рассуждениями временное и пространственное среднее совпадают и равны соответственно a_0 . В частности утверждение верно для вещественной и мнимой частей тригонометрического полинома. Вспомним тождество Фурье и сошлемся на то, что в курсе анализа оно обосновывается с использованием интегральной характеристики для расстояния: иными словами, каждую такую функцию на S^1 можно для каждого $\varepsilon > 0$ приблизить вещественнозначными тригонометрическими полиномами $P_\varepsilon^-(z)$ и $P_\varepsilon^+(z)$ так, что

$$\forall z \in S^1 \quad P_\varepsilon^-(z) < f(z) < P_\varepsilon^+(z) \quad \int_{S^1} (P_\varepsilon^+(z) - P_\varepsilon^-(z)) dl < \varepsilon$$

²Интеграл $\int_{\mathbb{T}} f dl$ называется пространственным средним. Свойство **эргодичности** состоит в совпадении пространственного и временного средних, мы пока не обсуждаем общее определение эргодичности преобразования пространств с мерой.

Но для вещественнозначных тригонометрических полиномов совпадение пространственных и временных средних вытекает из 2-го случая — поэтому

$$\begin{aligned} \int_{S^1} P_\varepsilon^- dl &= \lim_{n \rightarrow \infty} \inf \frac{1}{n} \sum_{k=0}^{n-1} P_\varepsilon^-(R^k(z)) \leq \lim_{n \rightarrow \infty} \inf \frac{1}{n} \sum_{k=0}^{n-1} f(R^k(z)) \leq \\ &\leq \lim_{n \rightarrow \infty} \sup \frac{1}{n} \sum_{k=0}^{n-1} f(R^k(z)) \leq \lim_{n \rightarrow \infty} \sup \frac{1}{n} \sum_{k=0}^{n-1} P_\varepsilon^+(R^k(z)) = \int_{S^1} P_\varepsilon^+ dl \end{aligned}$$

Поскольку пространственные интегралы у $P_\varepsilon^-(z)$ и $P_\varepsilon^+(z)$ различаются на сколь угодно малое ε , то теорема доказана. ■

Теперь несложно установить свойство равномерной распределенности траектории точки при поворотах: для этого у данной дуги dl надо аппроксимировать непрерывными функциями ее индикаторную функцию.

Практические следствия

Псевдослучайные числа

Алгоритм компьютерного генератора **rand** псевдослучайных чисел (имитирующих равномерное распределение на отрезке $[0, 1]$) основан именно на идее поворота на иррациональный угол.

О законе Бенфорда

Эвристический закон «закон первой цифры» утверждает, что частота появления единицы в реальных данных превосходит частоты появления других цифр. Несколько более детальное исследование (Бенфорд) первой значащей цифры в распределениях величин, взятых из реальной жизни, показывает, что частота появления цифры d , $d = 1, 2, \dots, 9$ близка к величине $\lg(d+1) - \lg(d)$. Объясним это свойство для геометрической прогрессии. Действительно, при q отличном от целой степени числа 10 дробные части десятичных логарифмов от q^k расположены на окружности \mathbb{T} как равномерно распределенная последовательность и, следовательно, частота их появления в заданной дуге в пределе совпадает с длиной дуги. Осталось вычислить длины дуг, отвечающих заданной первой значащей цифре — сделайте это самостоятельно.

На самом деле «закон первой цифры» возникает (на качественном уровне) из наблюдений. В ряде случаев удается количественно объяснить отчего реальные данные — например, данные о населении разных стран, о ценах или о некоторых природных процессах могут быть приближены членами геометрической прогрессии. Если перейти на уровень непрерывных описаний, то подобные явления связаны с решениями дифференциального уравнения $y' = y$ — то есть речь идет о данных экспоненциального роста. Имеются также вполне мистические рассуждения: например, аргументы присутствия фальсификаций на выборах Байден-Трамп 2020г часто использовали закон Бенфорда в анализе данных по разным округам.

Приложение к численному 2D-интегрированию

Выше мы получили заодно одну идею как можно интегрировать функции на окружности, впрочем пока не ясно, чем же такой метод предпочтительней других. Неожиданно идея равномерно распределенных последовательностей всплывает в случае необходимости интегрировать функции на сфере S^2 , где ситуация на первый взгляд совершенно другая: действительно, области любого большого грида не будут конгруэнтными и придется явно вычислять их площади, в аппроксимации интегральных сумм — все это будет совсем не похоже на вычислительно простую конструкцию временного среднего. Имеется, однако, аналогичная временному среднему конструкция для сферы S^2 , а именно, для любого N можно построить конечную последовательность точек s_1, s_2, \dots, s_N , $q_k \in S^2$ так, что для кусочно-непрерывной функции $f : S^2 \rightarrow \mathbb{R}$ «временное» среднее хорошо приближает пространственное среднее

$$\frac{1}{N} \sum_{k=1}^N f(s_k) \sim \int_{S^2} f(s) ds$$

Конструирование этих точек делается так: сначала в прямоугольнике $[0, 2\pi) \times [0, 2] \subset \mathbb{R}^2$ строим конечную последовательность точек с координатами $(\{k\theta\}, \frac{k}{n+1})$ (как и ранее $\theta = e^{2\pi i\omega}$, где ω иррациональное число и $k = 1, 2, \dots$). Далее сворачиваем указанный прямоугольник в цилиндр, внутри которого размещаем сферу единичного радиуса. Как известно, цилиндрическая проекция (проекция Архимеда) сохраняет площадь и потому

несложно убедиться, что указанная конечная последовательность точек аппроксимирует свойство равномерной распределенности как в прямоугольнике, так и на поверхности двумерной сферы. Такой грид известен под названием «спирального» — происхождение названия становится понятным, если написать указанные точки при помощи компьютера.

Связь рядов Фурье и точности метода трапеций

Во многих приложениях требуется оценить интеграл функции на конечном интервале. Перенормируя, мы приходим к оценке интеграла

$$I = \int_0^1 f(t) dt$$

Обычно это делается с использованием конечной последовательности узлов t_n и весов w_n так чтобы

$$I_N = \sum_{n=0}^{N-1} w_n f(t_n) \sim I$$

Числовая точность квадратурного правила характеризуется поведением величины $|I_N - I|$ при увеличении N . Самый простой метод квадратуры — это известный метод трапеций, который использует равномерно расположенные узлы с (почти) однородными весами:

$$I_N = \frac{1}{N} \left[\frac{1}{2} f(0) + f\left(\frac{1}{N}\right) + \dots + f\left(\frac{N-1}{N}\right) + \frac{1}{2} f(1) \right]$$

Предположим, что $f(0) = f(1)$, тогда

$$I_N = \frac{1}{N} \sum_{n=0}^{N-1} f\left(\frac{n}{N}\right)$$

Если функция f достаточно гладкая, то ее можно записать как сумму ряда Фурье (второй случай тождества Фурье):

$$f(t) = \sum_{k \in \mathbb{Z}} \langle f, u_k \rangle u_k(t)$$

то есть как линейную комбинацию частот $u_n(t) = e^{2\pi i n t}$

При $t = k/N$ $e^{2\pi i k n / N}$ является корнем N -й степени из единицы; он равен 1 тогда и только тогда, когда k кратно N . Далее

$$\frac{1}{N} \sum_{n=0}^{N-1} e^{2\pi i k n / N} = \begin{cases} 1 & k \text{ делится на } N \\ 0 & k \text{ не делится на } N \end{cases}$$

Тогда имеем следующее представление для I_N

$$I_N = \sum_{k \in \mathbb{Z}} \langle f, u_k \rangle \left[\sum_{m \in \mathbb{Z}} \delta_{mN}^k \right] = \langle f, u_0 \rangle + \sum_{m \in \mathbb{Z}, m \neq 0} \langle f, u_{mN} \rangle = I + \sum_{m \in \mathbb{Z}, m \neq 0} \langle f, u_{mN} \rangle$$

Тем самым ошибка интегрирования мала, если коэффициенты Фурье функции быстро убывают. Как известно, это убывание связано с гладкостью функции на \mathbb{T} , действительно, рассмотрим f , дифференцируемую p раз, интегрирование по частям дает

$$\begin{aligned} \langle f, u_k \rangle &= \int_0^1 f(t) e^{-2\pi i k t} dt = \frac{1}{2\pi i k} \int_0^1 f'(t) e^{-2\pi i k t} dt = \frac{1}{(2\pi i k)^p} \int_0^1 f^{(p)}(t) e^{-2\pi i k t} dt \\ |I_N - I| &\leq \sum_{m \in \mathbb{Z}, m \neq 0} \frac{\|f^{(p)}\|_\infty}{(2\pi m N)^p} = \frac{2\|f^{(p)}\|_\infty}{(2\pi N)^p} \sum_{m > 0} \frac{1}{m^p} \end{aligned}$$

Периодичность необходима для обращения в нуль граничных членов при интегрировании по частям. Если f гладкая, эта оценка верна для любого p , что означает ошибка убывает с ростом N быстрее, чем N^{-p} для любого p . Такое поведение называется спектральной точностью. Если функция гладкая лишь внутри отрезка, то это рассуждение невозможно и тем самым точность оказывается меньше.

Удивительное следствие

Формула (гауссовой) кривой $y = \exp(x^2)$ на конечном сегменте, скажем, $[-10, 10]$ определяет площадь под этой кривой, которую можно выразить с помощью имеющейся даже в калькуляторах функции erf:

$$\operatorname{erf}(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^x e^{-t^2} dt$$

Вычислите приближенно указанную площадь по формуле трапеций разделив отрезок целыми точками. Сравните полученный ответ с ответом, который дает функция erf.

Список литературы

[Моррис (1980)] С.Моррис , Двойственность Понтрягина и строение локально компактных абелевых групп, *изд. «Мир» 1980.*

Операторы в пространстве функций

Материал раздела целиком позаимствован из стенфордских лекций Emmanuel Candes "Applied Fourier Analysis and Elements of Modern Signal Processing".

Сдвигово-инвариантные линейные операторы в пространстве сигналов

Напомним, что сигналом в дискретном неограниченном времени принято называть функцию $f : \mathbb{Z} \rightarrow \mathbb{R}$, в непрерывном — функцию $f : \mathbb{R} \rightarrow \mathbb{R}$. Линейное преобразование сигналов $L : f \mapsto g$ в дискретном времени можно, например, записать (по аналогии умножением вектора на матрицу)

$$g(n) = \sum_{m \in \mathbb{Z}} K(n, m) f(m)$$

а в непрерывном времени в предыдущей формуле следует рассмотреть переобозначение и потом «предельный переход»¹:

$$g(t) = \int_{\mathbb{R}} K(t, s) f(s) ds$$

С другой стороны, сдвиг на величину τ во времени $f_{\tau}(t) = f(t - \tau)$ также является линейным оператором и стоит разобраться какое отношение формулы, приведенные выше, имеют к этому случаю.

Линейный оператор L перестановочный $g_{\tau} = L(f_{\tau})$ с оператором сдвига называется *инвариантным по времени*, широко употребляется аббревиатура ТИО от *time-invariant operation*.

Свертки

Определение операции свертки в непрерывном времени:

$$(f * g)(t) = \int_{\mathbb{R}} f(u) g(t - u) du$$

достаточно легко доказать (сделайте это!) коммутативность $f * g = g * f$ и ассоциативность $(f * g) * h = f * (g * h)$. Операция свертки с фиксированной функцией h порождает линейный оператор свертки $f \mapsto L(f) = g = h * f$, полезно представлять (пользуясь, например, версией свертки сигналов дискретного времени) этот оператор как усреднение исходной функции f с весами, определяемыми функцией h .

Заметим, что оператор свертки является ТИО:

$$(L f_{\tau})(t) = \int_{\mathbb{R}} f(u - \tau) h(t - u) du = \int_{\mathbb{R}} f(u) h((t - \tau) - u) du = (f * h)(t - \tau) = (L f)_{\tau}(t)$$

Обратное рассуждение для

$$g(t) = \int h(t, u) f(u) du$$

(опять=—таки по модулю обоснования корректности некоторых последующих выражений — с точки зрения математиков изрядные усилия необходимы для этого)². Если известно, что L

¹Здесь мы опять переносим вопрос о том, когда такой предел существует, в обсуждение конкретных примеров.

²Продумайте, например, как устроена функция *двух переменных* $h(t, u)$

является ТПО,

$$\begin{aligned}
g_\tau(t) &= (Lf)_\tau(t) = \int_{\mathbb{R}} h(t - \tau, u) f(u) du \\
(Lf_\tau)(t) &= \int_{\mathbb{R}} h(t, u) f_\tau(u) du = \int_{\mathbb{R}} h(t, u) f(u - \tau) du = \int_{\mathbb{R}} h(t, u + \tau) f(u) du \\
\forall f \quad \int_{\mathbb{R}} h(t - \tau, u) f(u) du &= \int_{\mathbb{R}} h(t, u + \tau) f(u) du \implies \int_{\mathbb{R}} [h(t - \tau, u) - h(t, u + \tau)] f(u) du = 0
\end{aligned}$$

откуда $h(t - \tau, u) - h(t, u + \tau)$ при всех t . Таким образом $h(t, u) = h(0 - (-t), u) = h(0, u - t)$ и тем самым $h(t, u) = \tilde{h}(t - u)$. Поэтому и $(Lf)(t) = \int \tilde{h}(t - u) f(u) du$, где \tilde{h} называется *сверточным ядром*. В приложениях это ядро часто называют *функцией отклика импульса*, а в оптике также называют *функцией рассеяния точки*. Так при оптическом исследовании очень локализованного практически в точку и яркого пятна (по идее похожего на физическую реализацию импульса) мы в реальности наблюдаем несколько размытое отображение; это размытое изображение и есть сверточное ядро оптической системы.

Свертки в случае $T = \mathbb{Z}$ и $T = \mathbb{Z}/N$

Формула свертки для двух сигналов времени $T = \mathbb{Z}$ определяется очевидным образом

$$(f * g)(n) = \sum_{m \in \mathbb{Z}} f(m) g(n - m)$$

В случае же $T = \mathbb{Z}/N$ часто говорят о *циклической свертке*

$$(f * g)(n) = \sum_{m=0}^{N-1} f(m) g((n - m) \bmod N)$$

Доказательство того, что ТПО является сверткой вполне легко трансформируется на эти два случая, что в общем неудивительно ввиду возможности рассматривать предельный переход от сигналов на $T = \mathbb{Z}/N$ к более общим сигналам.

Функция отклика импульса

Дискретное время $T = \mathbb{Z}$

В этом случае (вспоминаем конечный случай \mathbb{Z}/N единичный импульс δ в момент времени 0

$$\delta(t) = \begin{cases} 1 & t = 0 \\ 0 & t \neq 0 \end{cases}$$

Обозначение не случайно выбрано в согласии с обозначением обобщенной функции Дирака, действительно:

$$\begin{aligned}
\langle f, \delta \rangle &= \sum_{s \in \mathbb{Z}} f(s) \delta(s) = f(0) \\
(f * \delta)(t) &= \sum_{s \in \mathbb{Z}} f(s) \delta(t - s) = f(t) \implies f * \delta = f.
\end{aligned}$$

Тем самым, для ТПО L имеем $l\delta = h * \delta = h$

Непрерывное время $T = \mathbb{R}$

По самому определению δ -функции Дирака как линейного функционала и записи функционалов через скалярное произведение в пространстве функций

$$\langle f, \delta \rangle = f(0) = \int_{\mathbb{R}} f(s) \delta(s) ds$$
$$(f * \delta)(t) = \int_{\mathbb{R}} f(s) \delta(t-s) ds = f(t) \implies f * \delta = f.$$

Передаточные функции

В конечномерном случае \mathbb{Z}/N частотный базис возникал как собственные вектора (сигналы) оператора сдвига. Предельные переходы (обоснование которых было эвристическое, но так и не было объяснено в деталях) намекают на то, что и в случаях $T = \mathbb{Z}$, $T = \mathbb{R}$ частоты должны быть собственными векторами — на самом деле в разделе функционального анализа все это построение проводится вполне строго, но в нашем курсе воспроизводить все определения теории линейных операторов представляется нереальным. Поскольку речь идет о приложениях Математики, представляется более важным продемонстрировать идейную связь формул в разных ситуациях и ограничиться ссылками для желающих восстановить все детали. Например, для случая $T = \mathbb{R}$, для ТПО L и элемента $u_\omega(t) = e^{i\omega t}$ из «базиса частот»:

$$L(u_\omega)(t) = \int_{\mathbb{R}} h(s) e^{i\omega(t-s)} ds = e^{i\omega t} \underbrace{\int_{\mathbb{R}} h(s) e^{-i\omega s} ds}_{\hat{h}(\omega)} = \hat{h}(\omega) u_\omega(t)$$

выделение интеграла — выделение преобразования Фурье, которое можно **пока что** считать определенным для интегрируемых на \mathbb{R} функций³ Таким образом, для ТПО L элементы частотного базиса являются собственными векторами. Функция $\hat{h}(\omega)$ называется *передаточной функцией* — название возникло в эпоху развития радиосвязи и соответствующих электронных схем обработки сигналов. Тогда же возник термин *фильтр* для обработки сигналов схемами, которые реализуют некоторый ТПО оператор, а поскольку все производилось в реальном времени, то очень важно было различие между *текущим* значением, *прошлыми* и *будущими* значениями сигнала: физическая реализуемость требовала, чтобы работа фильтра не затрагивала «будущее». На инженерном языке это свойство называется *казуальностью фильтра*, а на математическом — означает, что $\forall s < 0 \ h(s) = 0$.

- Вопрос для самопроверки. Вернемся к дискретному преобразованию Фурье при $T = \mathbb{Z}/n$. Является ли полосовой фильтр ТПО?

Некоторые вычисления

1. Преобразование Фурье $\hat{f}(\omega)$ функции

$$f(t) = \begin{cases} 0 & t \notin [-\frac{1}{2}, \frac{1}{2}] \\ 1 & t \in [-\frac{1}{2}, \frac{1}{2}] \end{cases}$$

³Существует некоторая вариативность нормировок для преобразования Фурье и обратного к нему. Так, например, преобразование Фурье иногда определяют и так:

$$\hat{f}(\omega) = \int_{\mathbb{R}} f(s) e^{-2\pi i \omega s} ds$$

Преобразование Фурье $\hat{f}(\omega)$ этой функции вычисляется явно и ответ выражается через «почти элементарную» функцию sinc:

$$\text{sinc}(x) = \begin{cases} \frac{\sin \pi x}{\pi x} & x \neq 0 \\ 1 & x = 0 \end{cases}$$

Восстановите вычисление и найдите точное выражение для $\hat{f}(\omega)$ самостоятельно. Следует отметить важную особенность ответа: функция $f(t)$ имела компактный носитель в то время как носитель ее образа Фурье не ограничен. На самом деле, сходное свойство лежит также в основе известного соотношения неопределенностей для квантовой теории, суть которого заключена в том, что квантовые наблюдаемые (в теории) связаны преобразованием Фурье!

2. Преобразование Фурье $\hat{h}(\omega)$ «треугольной функции»

$$h(t) = \begin{cases} 0 & t \notin [-1, 1] \\ 1 - |t| & t \in [-1, 1] \end{cases}$$

Приведем вычисление, которое необходимо еще пояснить самостоятельно. Во-первых, при $\omega = 0$ имеем $\hat{h}(0) = 1$ (см. также замечание к следующему пункту). Во-вторых, при $\omega \neq 0$

$$\begin{aligned} \hat{h}(\omega) &= \int_{-1}^1 f(t) \cos(\omega t) dt = 2 \int_0^1 (1 - t) \cos(\omega t) dt \\ &= 2(1 - t) \frac{\sin(\omega t)}{\omega} \Big|_0^1 + \frac{2}{\omega} \int_0^1 \sin(\omega t) dt = \frac{2}{\omega^2} (1 - \cos \omega) = \text{sinc}^2 \left(\frac{\omega}{2\pi} \right) \end{aligned}$$

3. Вычислите самостоятельно свертку $f * f$. Отметим, что в теории вероятностей функция f является функцией плотности равномерно распределенной на отрезке $[-\frac{1}{2}, \frac{1}{2}]$ случайной величины, а функция h из следующего пункта является функцией плотности суммы двух независимых случайных величин с одинаковой плотностью f . Это техническое наблюдение на самом деле дает основу метода доказательства Центральной Предельной Теоремы для независимых и одинаково распределенных случайных величин с некоторыми ограничениями на их функцию плотности.

4. Преобразование Фурье от функции Гаусса, нормировку которой выберем также в соответствии с Теорией Вероятности:

$$g(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

Наиболее короткое и изящное вычисление сводится к дифференцированию под знаком интеграла и решению простейшего дифференциального уравнения.

$$\frac{d}{d\omega} (\hat{g}(\omega)) = \frac{1}{\sqrt{2\pi}} \int (-it) e^{-\frac{t^2}{2}} e^{-i\omega t} dt \stackrel{\text{по частям}}{=} -\frac{\omega}{\sqrt{2\pi}} \int e^{-\frac{t^2}{2}} e^{-i\omega t} dt = -\omega \hat{g}(\omega)$$

Решением ОДУ $\frac{d(\hat{g}(\omega))}{d\omega} = -\omega \hat{g}(\omega)$ является, очевидно, функция вида $\text{const} \cdot e^{-\frac{\omega^2}{2}}$. Осталось разобраться (самостоятельно) чему, исходя из формулы преобразования Фурье, равна эта константа.

Вычисления преобразования Фурье

Существуют как минимум три различия интерпретаций быстрых преобразований Фурье, соответствующих случаям либо конечной, либо компактной, либо некомпактной группы.

- Конечные группы. В этом случае мы рассматриваем f как функцию от элементов циклической группы \mathbb{Z}/N порядка N ; дискретное преобразование Фурье — отвечает процедуре замены базиса Прямое применение формулы

$$\hat{f}(n) = \sum_s f(s) e^{-\frac{2\pi i s n}{N}}$$

в алгоритме вычисления всех коэффициентов Фурье $\hat{f}(k)$ использует $\sim N^2$ операций, быстрое преобразование Фурье FFT, о котором речь пойдет далее, — другой алгоритм с меньшим числом операций $\sim N \ln N$

- Компактная топологическая абелева группа — окружность S^1 (комплексные числа единичного модуля), коэффициенты Фурье в этом случае

$$\hat{f}(n) = \int_0^1 f(e^{2\pi i t}) e^{-2\pi i t n} dt$$

В этом варианте интересен приближенный алгоритм вычисления коэффициентов Фурье в весьма специальном случае, когда среди них лишь конечное число отлично от нуля. В этом случае также используют FFT.

- Некомпактная топологическая абелева группа \mathbb{R} , преобразование Фурье определено интегральным оператором

$$\hat{f}(x) = \int_{-\infty}^{\infty} f(y) e^{-2\pi i y x} dy$$

Как и в предыдущем случае, дискретное FFT можно использовать для аппроксимации интеграла, и снова важен случай, когда преобразование Фурье имеет только конечный носитель). В этом случае действительная функция f определяется своими значениями только на дискретной решетке в \mathbb{R} — это так называемая теорема Котельникова-Шеннона, о которой речь впереди.

- Некоммутативный случай. Напомним, что в случае \mathbb{Z}/N формула преобразования Фурье появилась из замены координат такой, что координатные вектора были выбраны как собственные для оператора сдвига $T_y \mathbf{v}(x)$ в пространстве \mathbb{C}^N . В случае конечной некоммутативной группы G из конечного числа N элементов сдвиги устроены сложнее и возникает действие сдвигов на аргументах и тем самым представление группы сдвигов T_y в группе конечных матриц преобразований пространства \mathbb{C}^N).

В некоммутативном случае нет собственных функций одновременных для всех сдвигов. Это приводит к поиску базисных функций $\mathbf{v}_k(x)$ $\mathbf{v}_k \in \mathbb{C}^N$, обладающих свойством

$$T_y \mathbf{v}_k(x) = \mathbf{v}_k(xy^{-1}) = \sum_l a_y(k, l) \mathbf{v}_l(x)$$

Видим, что собственные векторы и собственные значения заменяются на векторы и действия матричных элементов $a_y(k, l)$; это и задает обобщение случая коммутативной группы. Все вместе, они задают матричное представление группы G — то есть элемент двойственной группы (обозначается \hat{G}).

Итак, для любой функции f , определенной на конечной группе G , имеем аналог разложения Фурье:

$$f(x) = \sum_{r \in \hat{G}} c_r \sum_{k,l} \hat{f}(k,l) a_x(k,l)$$

где c_r — некоторая константа, зависящая от неприводимого представления r , $\hat{f}(k,l)$ — «коэффициенты Фурье», а (зависящие от x) матричные элементы $a_x(k,l)$ теперь порождают аналог спектрального пространства. Преобразование Фурье в этом общем случае вычисляет коэффициенты Фурье посредством скалярного произведения функции с базисом из матричных элементов неприводимых представлений. Если группа G не является конечной, а компактна, то такая сумма бесконечна (по аналогии с суммой над целыми числами в случае окружности, но опять встает вопрос: для каких функций f бесконечная сумма имеет смысл). Если же G некомпактна, то выражение с бесконечными суммами, является своего рода интегралом (но эти понятия уже требуют ссылки на серьезное изложение теории представлений).

Теорема отсчетов (Котельникова-Шеннона)

Рассмотрим интегральный вариант преобразования Фурье в случае хорошей (как обычно, мы пока не проговариваем до конца условия на функцию, для которой выполнено тождество Фурье) функции $f(t)$ такой, что ее преобразование Фурье \hat{f} равно нулю всюду, вне отрезка $[-\pi, \pi]$. Обозначим $g = 2\pi\hat{f}$. Обычные формулы дают следующее:

$$f(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(s) e^{its} ds \quad g(s) = \sum_{n=-\infty}^{\infty} c_n e^{-ins}$$

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(s) e^{its} ds = f(n)$$

Подставим в первый интеграл вместо g ее ряд Фурье и после почленного интегрирования имеем:

$$f(t) = \sum_{n=-\infty}^{\infty} c_n \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(t-n)s} ds = \sum_{n=-\infty}^{\infty} f(n) \frac{\sin \pi(t-n)}{\pi(t-n)} = \frac{\sin \pi t}{n} \sum_{n=-\infty}^{\infty} f(n) \frac{(-1)^n}{t-n}$$

Здесь условие о локализации спектра именно на $[-\pi, \pi]$ тривиально обобщается на случай сегмента $[-a, a]$ и потому получаем следующее представление функции с финитным спектром, известное под названием формулы Котельникова, или формулы Котельникова—Шеннона, или теоремы отсчетов:

$$f(t) = \sum_{n=-\infty}^{\infty} f(n\pi/a) \frac{\sin a(t - n\pi/a)}{a(t - n\pi/a)}$$

Эта формула была получена Котельниковым в 1933 году и переоткрыта Шенноном в 1949 году. Она показывает, что регулярный сигнал с финитным спектром частот, лежащих в промежутке $[-a, a]$, может быть полностью восстановлен по набору его дискретных значений, отсчитываемых с интервалом $\Delta = \frac{\pi}{a}$. (Отсюда название «теорема отсчетов»). Более того, сигнал восстанавливается посредством комбинации, составленной из сдвигов простейшего фильтра низких частот с верхней граничной частотой a , т. е. это ещё и относительно просто технически реализуемая формула дискретного кодирования и передачи сигнала.

Алгоритм Дж. У. Кули и Джона Тьюки FFT

Алгоритм, который вычисляет дискретное преобразование Фурье (DFT, а в русскоязычной литературе ДПФ) последовательности или ее обратное преобразование (IDFT), может быть оптимизирован и в этом случае он называется *быстрым* преобразованием Фурье (FFT). Алгоритм Дж. У. Кули и Джона Тьюки, является наиболее распространенным алгоритмом FFT, он основан на некоторой простой рекурсивной формуле когда длина сигнала N – четное число. Существует также метод PFA (algorithm Good-Thomas), который рекурсивно выражает дискретное преобразование Фурье размера $N = N_1 \cdot N_2$, но только в том случае, когда N_1 и N_2 взаимно простые. Применив PFA рекурсивно можно получить другой алгоритм FFT и вообще говоря с другой эффективностью вычислений. Также имеются алгоритмы, использующие рекурсии, построенные по разложению целого числа N в произведение простых чисел (например, алгоритм Винограда).

Выяснилось, что основная идея (после интенсивного рекламирования метода FFT) применения рекурсии была изобретена около 1805 года Карлом Фридрихом Гауссом, который однако не стал публиковать соответствующую статью, и не анализировал асимптотическое время вычислений. Различные ограниченные формы алгоритма также несколько раз переоткрывались заново, но оставались малоизвестными на протяжении XIX и в начале XX вв, пока наконец Кули и Тьюки в 1965 г. самостоятельно заново его открыли и, — что весьма важно, — популяризировали, описав через 160 лет после К-Ф Гаусса, как нужно алгоритм выполнять на компьютере. Как сообщают, идея был придумана во время заседания Научно-консультативного комитета при президенте Джоне Кеннеди: обсуждались способы обнаружения ядерных испытаний в Советском Союзе с помощью сейсмометров, расположенных по всему миру. Эти датчики должны доставлять временные ряды наблюдений за колебаниями почвы, анализ этих сигналов потребовал бы быстрых алгоритмов расчета из-за количества датчиков и ограничений времени. Алгоритм Кули-Туки для $N = 2^m$ является наиболее простой и распространенной формой FFT и входит в стандартные библиотеки многих языков программирования; в оптимизированных реализациях на уровне архитектуры процессоров, как правило, используют другие версии ускорения расчета преобразования Фурье.

Разберем детали алгоритма: основная идея состоит в разделении дискретного преобразования Фурье для сигнала длины N — будем употреблять заодно с DFT и обозначение F — на пару чередующихся F размера $N/2$ и дальнейшем итерировании такого представления (программист реализует это через рекурсивный вызов соответствующей функции). Таким образом, (в случае, когда N - это степень двойки) общее число операций уменьшается до $O(N \log N)$. Количество N изначально обычно можно подобрать, например, путем дополнения сигнала нулями: это часто не является важным ограничением.

Дискретное преобразование Фурье DFT определяется формулой:

$$\hat{f}_s = \sum_{t=0}^{N-1} f_t e^{-\frac{2\pi i}{N} ts}$$

где s - целое число от 0 до $N - 1$.

Вычислим DFT по четным индексам f_0, f_2, \dots, f_{N-2} и по нечетным индексам f_1, f_3, \dots, f_{N-1} , а затем объединим эти два результата для получения DFT всей последовательности. Более конкретно:

$$\hat{f}_s = \sum_{t=0}^{N/2-1} f_{2t} e^{-\frac{2\pi i}{N} (2t)s} + \sum_{t=0}^{N/2-1} f_{2t+1} e^{-\frac{2\pi i}{N} (2t+1)s}$$

Из второй суммы можно вынести общий множитель $e^{-\frac{2\pi i}{N}s}$. Ясно, что получившимися двумя суммами являются DFT четно-индексированной части и DFT нечетно-индексированной части. Обозначим DFT четно-индексированной части F_k^e и DFT нечетно-индексированной части F_k^o , получим:

$$\hat{f}_s = \underbrace{\sum_{t=0}^{N/2-1} f_{2t} e^{-\frac{2\pi i}{N/2}ts}}_{\text{четная часть}} + e^{-\frac{2\pi i}{N}s} \underbrace{\sum_{t=0}^{N/2-1} f_{2t+1} e^{-\frac{2\pi i}{N/2}ts}}_{\text{нечетная часть}} = F_s^e + e^{-\frac{2\pi i}{N}s} F_s^o$$

Благодаря периодичности комплексных экспонент выражение для $\hat{f}_{s+\frac{N}{2}}$ также получается из F_s^e и F_s^o :

$$\begin{aligned} \hat{f}_{s+\frac{N}{2}} &= \sum_{t=0}^{N/2-1} f_{2t} e^{-\frac{2\pi i}{N/2}t(s+\frac{N}{2})} + e^{-\frac{2\pi i}{N}(s+\frac{N}{2})} \sum_{t=0}^{N/2-1} f_{2t+1} e^{-\frac{2\pi i}{N/2}t(s+\frac{N}{2})} \\ &= \sum_{t=0}^{N/2-1} f_{2t} e^{-\frac{2\pi i}{N/2}ts} e^{-2\pi it} + e^{-\frac{2\pi i}{N}s} e^{-\pi i} \sum_{t=0}^{N/2-1} f_{2t+1} e^{-\frac{2\pi i}{N/2}ts} e^{-2\pi it} \\ &= \sum_{t=0}^{N/2-1} f_{2t} e^{-\frac{2\pi i}{N/2}ts} - e^{-\frac{2\pi i}{N}s} \sum_{t=0}^{N/2-1} f_{2t+1} e^{-\frac{2\pi i}{N/2}ts} \\ &= F_s^e - e^{-\frac{2\pi i}{N}s} F_s^o \end{aligned}$$

Окончательно:

$$\begin{aligned} \hat{f}_s &= F_s^e + e^{-\frac{2\pi i}{N}s} F_s^o \\ \hat{f}_{s+\frac{N}{2}} &= F_s^e - e^{-\frac{2\pi i}{N}s} F_s^o \end{aligned}$$

Тем самым DFT сигнала длины N рекурсивно выражено в терминах двух DFT размера $N/2$, итерируя эту конструкцию при $N = 2^m$, получим окончательный алгоритм, состоящий из $\text{const} \cdot N \log N$ операций. Несколько иное изложение см. у Emmanuel Candes в лекции 7 раздел 4.

Свертка и преобразование Фурье

Если все функции f, h лежат в классе $L^1(\mathbb{R})$ интегрируемых, то и $g = f * h$ принадлежит $L^1(\mathbb{R})$, а также

$$\hat{g}(\omega) = \hat{f}(\omega) \hat{h}(\omega)$$

Первое утверждение вполне просто (поскольку произведение $f(u)h(t-u)$ интегрируемых на \mathbb{R} функций дает функцию, интегрируемую на \mathbb{R}^2 и, значит, можно применить теорему Фубини):

$$\int |g(t)| dt \leq \iint |f(u)| |h(t-u)| du dt = \left(\int |f(u)| du \right) \cdot \left(\int |h(t)| dt \right) < \infty$$

Второе утверждение также использует теорему Фубини

$$\begin{aligned} \hat{g}(\omega) &= \int \left(\int f(u) h(t-u) du \right) e^{-i\omega t} dt = \iint f(u) h(t-u) e^{-i\omega t} du dt = \iint f(u) h(t) e^{-i\omega(u+t)} du dt \\ &= \left(\int f(u) e^{-i\omega u} du \right) \cdot \left(\int h(t) e^{-i\omega t} dt \right) = \hat{f}(\omega) \hat{h}(\omega) \end{aligned}$$

сигнал	спектр	название
$(f * g)(t)$	$\hat{f}(\omega)\hat{g}(\omega)$	свертка сигналов
$f(t)g(t)$	$\frac{1}{2\pi}(\hat{f} * \hat{g})(\omega)$	умножение сигналов
$f(t-s)$	$e^{-i\omega s}\hat{f}(\omega)$	запаздывание
$e^{i\omega_0 t}f(t)$	$\hat{f}(\omega - \omega_0)$	амплитудная модуляция
$f(\frac{t}{a})$	$ a \hat{f}(a\omega)$	масштабирование
$f^{(k)}(t)$	$(i\omega)^k \hat{f}(\omega)$	дифференцирование
$(-it)^k f(t)$	$\hat{f}^{(k)}(\omega)$	дифференцирование спектра

Таблица 1: Таблица свойств *интегрального* преобразования Фурье, предполагается, что для сигналов в левой колонке оно определено

Формулы и их использование

Разумеется, при преобразовании Фурье имеет место связь свертки и произведения не только в случае сигналов $\mathbb{R} \rightarrow \mathbb{R}$, но и в других версиях выбора временной области: \mathbb{T} , \mathbb{Z} , \mathbb{Z}/N — в последнем случае речь идет о циклической свертке. Полезно научиться формулировать соответствующее утверждение в каждом из указанных случаев.

Достаточно ясно, что в случае дискретного времени \mathbb{Z}/N произведение сосчитать гораздо проще и быстрее, нежели по исходной формуле вычислять *циклическую* свертку. С другой стороны, далее будет рассказано, что вычисление в этом случае преобразования Фурье возможно за гораздо меньшее, чем N^2 число операций (алгоритм Быстрого Преобразования Фурье или FFT) и потому появляется возможность вычисления свертки с помощью быстрого алгоритма fft. Неожиданное применение этот факт находит при создании алгоритмов быстрого умножения многозначных чисел, вот одна из идей.

Алгоритм Schönhage–Strassen быстрого умножения

Идея Schönhage–Strassen алгоритма быстрого умножения многозначных чисел с помощью FFT основана на использовании формулы *циклической* свертки. Действительно, пусть $(f_0, f_1, f_2) = (123)$, $(g_0, g_1, g_2) = (456)$. Рассмотрим операцию умножения в столбик:

$$\begin{array}{r}
 \begin{array}{ccc} 1 & 2 & 3 \end{array} \\
 \times \quad \begin{array}{ccc} 4 & 5 & 6 \end{array} \\
 \hline
 \begin{array}{ccc} 6 & 12 & 18 \end{array} \\
 \begin{array}{ccc} 5 & 10 & 15 \end{array} \\
 \begin{array}{ccc} 4 & 8 & 12 \end{array} \\
 \hline
 \begin{array}{ccc} 4 & 13 & 28 & 27 & 18 \end{array}
 \end{array}$$

Нетрудно видеть, например, что $f_0g_2 + f_1g_1 + f_2g_0 = h_2 = 28$, где $h = f * g$. Аналогично можно объяснить и другие числа в нижней строке, если дополнить исходные сигналы f, g нулями. Таким образом в последней строке записана свертка сигналов, а от результата 560088 истинного умножения ее отличает только перенесение разрядов многозначных чисел — операции сложения, повторенные несколько (пропорционально длине последней строки) раз.

Вернемся к общим формулам свойств преобразования Фурье, Итак, у нас есть четыре разных области U определения комплекснозначных функций: множество \mathbb{Z}/N дискретных точек на окружности, множество целых чисел \mathbb{Z} , окружность S длины $2L$, числовая прямая \mathbb{R} . Каждую такую область мы называем либо импульсной областью и тогда ее элементы называются моментами времени, либо спектральной областью — тогда ее элементы называются частотами. Комплекснозначная функция на импульсной области обычно называется сигналом¹, комплекснозначная функция на спектральной области обычно называется спектром. При этом для *некоторых* комплекснозначных функций на выбранной области имеются сходные меж собой конструкции позволяющие по сигналу f строить спектр \hat{f} (определенный, вообще говоря, на другой области) и наоборот — эти конструкции называются соответственно преобразованием Фурье² и обратным преобразованием Фурье. Сводка результатов, которые мы успели обсудить ранее, для случая непрерывного времени \mathbb{R} дана в Таблице 1.

Контрольный вопрос

Сформируйте аналогичные таблицы свойств для остальных вариантов сигнал-спектр.

¹В традиционных приложениях сигнал обычно вещественнозначный

²Напомним, что в стандартных начальных курсах математического анализа обычно сосредотачиваются на свойствах вещественнозначных функций и потому формулы преобразований Фурье часто адаптируют к этому случаю

Взаимосвязь непрерывных и дискретных сигналов

Во многих реальных приложениях интересны измерения физических величин, которые (*в теории*) непрерывно изменяются во времени, они называются *аналоговыми сигналами или сигналами непрерывного времени* и их моделью служит непрерывная функция $f(t)$. Часто однако из-за специфики и ограничений процесса измерения, часто невозможно получить все множество значений $f(t)$, $t \in \mathbb{R}$ для функции и потому приходится работать с дискретным набором данных. Традиционный метод получения дискретных данных состоит в измерении значений f через интервалы времени длины Δ , в этом случае возникает последовательность значений $\{f(n\Delta)\}$, $n \in \mathbb{Z}$ в которой и заключена доступная информация о непрерывном сигнале. Мы называем $1/\Delta$ частотой дискретизации. Соответствующая электронная схема в аналоговых измерительных приборах называется аналого-цифровым преобразователем — АЦП.

Вообще говоря, нет надежды восстановить аналоговый сигнал $f(t)$ по его дискретизации $\{f(n\Delta)\}$ — это достаточно очевидно. Попробуем однако отследить связь спектров исходного сигнала и дискретного, для этого сравнения нам понадобится обратиться к обобщенным функциям.

Рассмотрим обобщенную функцию, которая называется *гребенка Дирака*, а также ее преобразование Фурье, которое является $\frac{2\pi}{\Delta}$ -периодической функцией:

$$f(t) = \sum_{n \in \mathbb{Z}} \delta(t - n\Delta) \Rightarrow \hat{f}(\omega) = \sum_{n \in \mathbb{Z}} e^{-in\omega\Delta}$$

Формула суммирования Пуассона

Для гладкой быстроубывающей со всеми своими производными (то есть взятой из пространства пробных функций) $\varphi(t)$ имеем

$$\int_{\mathbb{R}} \left(\sum_{n \in \mathbb{Z}} e^{-in\omega\Delta} \right) \varphi(\omega) d\omega = \frac{2\pi}{\Delta} \sum_{n \in \mathbb{Z}} \varphi\left(\frac{2\pi n}{\Delta}\right)$$

иными словами, преобразование Фурье от гребенки Дирака с периодом Δ , это снова гребенка Дирака (вспомните для сравнения конечный дискретный пример из предыдущей лекции), но уже с периодом $\frac{2\pi}{\Delta}$.

Доказательство. В силу свойства масштабирования для $g(t) = f(\Delta t)$ имеем $\hat{g}(\omega) = \frac{1}{\Delta} \hat{f}\left(\frac{\omega}{\Delta}\right)$; поэтому можно рассматривать лишь случай $\Delta = 1$. В этом случае (в силу явного представления через комплексные экспоненты) \hat{f} имеет период 2π и, значит, достаточно будет проверить, что ее ограничение на отрезок $[-\pi, \pi]$ (с продолжением нулями вне) совпадет с обычной δ -функцией Дирака (что, кстати, даст еще одно представление функции Дирака через последовательность функционалов):

$$\lim_{N \rightarrow \infty} \int_{-\pi}^{\pi} \left(\sum_{n=-N}^N e^{-ins} \right) \varphi(s) ds = 2\pi \varphi(0)$$

Ранее, при обсуждении сходимости частичных сумм ряда Фурье был сформулирован Принцип Локализации, несложно видеть, что он дает нужный ответ для предела частичных сумм коэффициентов c_n разложения в ряд Фурье гладкой внутри $[-\pi, \pi]$ функции $\varphi(s)$.

$$\lim_{N \rightarrow \infty} \sum_{n=-N}^N \left(\int_{-\pi}^{\pi} e^{-ins} \varphi(s) ds \right)$$

Действительно, тождество Фурье для гладкой функции выполнено и потому значение в нуле функции φ выражается через сумму (двустороннего) ряда $\sum_{n \in \mathbb{Z}} c_n \exp(i \cdot 0)$; коэффициент 2π пришел из тождества Фурье.

Алиасинг и формула отсчетов

Термин алиасинг (калька английского термина *aliasing*, в переводе «наложение») отвечает ситуации, когда дискретизация сигнала связана с искажениями (аддитивного характера) в спектре дискретизованного сигнала по сравнению с исходным.

Рассмотрим обобщенную функцию $f_d(t)$

$$f_d(t) = \sum_{n \in \mathbb{Z}} f(n\Delta) \delta(t - n\Delta)$$

связанную с дискретизацией $\{f(n\Delta)\}_{n \in \mathbb{Z}}$ сигнала $f(t)$. Преобразование Фурье обобщенной функции f_d

$$\hat{f}_d(\omega) = \sum_{n \in \mathbb{Z}} f(n\Delta) e^{-in\Delta\omega}$$

Основной вопрос: как связаны \hat{f} и \hat{f}_d ?

Формула алиасинга

$$\hat{f}_d(\omega) = \frac{1}{\Delta} \sum_{n \in \mathbb{Z}} \hat{f}\left(\omega - \frac{2\pi}{\Delta}n\right)$$

Таким образом формула алиасинга показывает, что спектр дискретизации с шагом Δ отвечает суммированию сдвинутых на $2\pi/\Delta$ спектров самой функции с последующей нормализацией на $1/\Delta$. Интуитивно понятно, что низкие частоты дискретизации (т. есть при больших значениях Δ) вызовут существенные перекрытия, что затмит реальный спектр \hat{f} . С другой стороны, поскольку для обычных функций \hat{f} должен затухать на высоких частотах, частая дискретизация уменьшит это перекрытие и лучше сохранит форму \hat{f} в спектре дискретизированного сигнала.

Объяснение формулы алиасинга: как и было выше, для простоты положим $\Delta = 1$, тогда с использованием формулы свертки и доказанной выше формулы суммирования Пуассона имеем

$$\begin{aligned} f_d(t) &= \sum_{n \in \mathbb{Z}} f(n) \delta(t - n) = f(t) \sum_{n \in \mathbb{Z}} \delta(t - n) = f \cdot \sum_{n \in \mathbb{Z}} \delta_n \\ \hat{f}_d(\omega) &= \frac{1}{2\pi} \left[\hat{f} * \left(2\pi \sum_{n \in \mathbb{Z}} \delta_{2\pi n} \right) \right] (\omega) = \sum_{n \in \mathbb{Z}} \hat{f}(\omega - 2\pi n) \end{aligned}$$

В ситуации, когда носитель $|\text{supp}| \hat{f}$ компактен перекрытия можно полностью избежать. Поэтому особо интересен специальный класс сигналов — сигналы с ограниченным спектром, то есть те, у которых преобразование Фурье имеет компактный носитель $\hat{f}(\omega) \equiv 0$ при $|\omega| > \omega_b$, в этом классе связь спектров исходного сигнала и дискретного такова, что восстановление непрерывного сигнала по дискретизации возможно.

Формула отсчетов. Повторение

Если спектр регулярного (например, непрерывного) сигнала f сосредоточен на сегменте $[-\frac{\pi}{\Delta}, \frac{\pi}{\Delta}]$, то

$$\boxed{f(t) = \sum_{n=-\infty}^{\infty} f(n\Delta) \text{sinc}\left(\frac{t}{\Delta} - n\right)}. \text{ В частности, при } \Delta = 1$$

$$f(t) = \sum_{n=-\infty}^{\infty} f(n) \frac{\sin \pi(t-n)}{\pi(t-n)} = \frac{\sin t\pi}{\pi} \sum_{n=-\infty}^{\infty} \frac{(-1)^n}{t-n} f(n)$$

Действительно, при $\Delta = 1$ по формуле алиасинга сдвиги спектра (сосредоточенного *внутри* сегмента $[-\pi, \pi]$) на $2\pi k$ при $k \neq 0$ вообще не пересекаются, и потому $\hat{f}(\omega) = \hat{f}_d(\omega) \mathbf{1}_{\{|\omega| \leq \pi\}}$. Произведение при преобразовании Фурье возникает из свертки, а значит

$$f(t) = \left[\left(\sum_{n \in \mathbb{Z}} f(n) \delta_n \right) * \text{sinc} \right] (t) = \sum_{n \in \mathbb{Z}} f(n) \text{sinc}(t - n)$$

Замечание Обратите внимание, что частота ω , используемая в условии про ограниченность спектра сегментом длины $\frac{2\pi}{\Delta}$, отличается множителем от частоты дискретизации с шагом Δ — ее естественно было бы ассоциировать с величиной $\nu = \frac{1}{\Delta}$. Поэтому ω называют иногда *круговой частотой*, имея в виду естественное соотношение $\omega = 2\pi\nu$, она получает размерность «радианы в секунду» в отличие от размерности в герцах sec^{-1} , характерных для ν . Формула Котельникова-Шеннона в единицах ν (при этом подразумевается, что спектр ограничен $|\nu| < W$) имеет несколько иной вид:

$$f(t) = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2W}\right) \frac{\sin \left[2\pi W \left(t - \frac{n}{2W} \right) \right]}{2\pi W \left(t - \frac{n}{2W} \right)}$$

С названием формулы история такая: Котельников опубликовал эту теорему в 1933 году, а Шеннон в 1949, кроме того похожий результат иногда называют еще теоремой Найквиста. Как интерполяционная формула (специальный случай формулы Лагранжа для целых функций конечной степени) она уже была известна и ранее. Но Котельников и Шеннон интересовались интерпретацией этой формулы для теории кодирования сигнала и передачи информации по каналу связи.

Пример

Рассмотрим (не принадлежащую $L^2(\mathbb{R})$) функцию $f(t) = \cos\left(\frac{3\pi}{2}t\right)$, то есть $\hat{f}(\omega) = \frac{1}{2} [\delta_{-3\pi/2} + \delta_{3\pi/2}]$, то есть спектр сосредоточен на $[-3\pi/2, 3\pi/2]$ и, согласно формуле Котельникова-Шеннона удобное представление дискретным рядом возникает при $\Delta = 2/3$. Однако, можно рассмотреть свойство ограниченности спектра на отрезке $[-\pi, \pi]$ и выбрать $\Delta = 1$. Тогда для дискретизации $f_d = \sum_{k \in \mathbb{Z}} f(k)\delta_k$ преобразование Фурье

$$\hat{f}_d(\omega) = \sum_{k \in \mathbb{Z}} \hat{f}(\omega - 2\pi k) = \frac{1}{2} \sum_{k \in \mathbb{Z}} \left[\delta\left(\omega - \frac{3\pi}{2} - 2\pi k\right) + \delta\left(\omega + \frac{3\pi}{2} - 2\pi k\right) \right]$$

на отрезке частот $[-\pi, \pi]$ не есть \hat{f} , но

$$\hat{f}_d(\omega) \cdot \mathbf{1}_{\{|\omega| \leq \pi\}} = \frac{1}{2} \left[\delta\left(\omega - \frac{\pi}{2}\right) + \delta\left(\omega + \frac{\pi}{2}\right) \right]$$

что соответствует равенству в выражениях связанных с формулах Котельникова-Шеннона

$$\sum_{k \in \mathbb{Z}} f(k) \text{sinc}(t) = \sum_{n \in \mathbb{Z}} \cos\left(\frac{3\pi}{2}n\right) \text{sinc}(t) = \sum_{n \in \mathbb{Z}} \cos\left(\frac{\pi}{2}n\right) \text{sinc}(t) = \cos\left(\frac{\pi}{2}t\right)$$

(но, разумеется, спектр для $\cos\left(\frac{\pi}{2}t\right)$ сосредоточен на $[-\pi, \pi]$, а для $\cos\left(\frac{3\pi}{2}t\right)$ это не так.

Тем самым, при заниженных частотах дискретизации мы пропускаем разницу между $\cos\left(\frac{\pi}{2}t\right)$ и $\cos\left(\frac{3\pi}{2}t\right)$ и формула Котельникова-Шеннона реконструирует целый класс функций (совпадающих в данных узлах с точностью до тех функций, у которых нули в данных узлах).

Замечание Рассуждения последних разделов использовали преобразование Фурье в дискретном и непрерывном случаях. Для включения в этот ряд рядов Фурье сформулируем еще раз тождество Фурье в смысле равенства в пространстве L^2 с выбором традиционных коэффициентов для этого случая. Для суммируемых с квадратом функций f , определенных на $[-\pi, \pi]$ таких, что $f(-\pi) = f(\pi)$

$$\begin{aligned} f &= \sum_{k \in \mathbb{Z}} c_k \cdot u_k & c_k &= (u_k, f) & c_k &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(t) e^{-ikt} dt \\ f(t) &= \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} c_k e^{ikt} & \|f\|_{L^2}^2 &= \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} |c_k|^2 \end{aligned}$$

Сигнал и его спектр

Мы начали изучение преобразования Фурье с дискретного варианта тождества Фурье. В этом дискретном случае имело место соответствие сигналов и спектров, что было несложным результатом из линейной алгебры. Преобразование Фурье в других рассмотренных случаях за счет предельного перехода с одной стороны нарушает это удобное свойство и нам придется рассмотреть по крайней мере несколько вариантов дополнительных условий на сигнал и/или спектр, чтобы соответствующая «обратимость преобразования» была бы возможна. Необходимость таких рассмотрений связана с тем, что вычислительные модели все равно дискретны и тем самым трудности обратимости так или иначе связаны с точностью приближений непрерывных сигналов с помощью дискретных моделей.

С другой стороны, реальные данные и так дискретны и вопрос об их соответствии какой-либо непрерывной модели важен. Несмотря на некоторую некорректность постановки такой задачи, в ряде случаев можно предложить методы, проясняющие аргументы в пользу или против того, что исследуемый сигнал близок к «хорошему непрерывному» и его исследование с помощью спектра не внесет грубых ошибок. Забегая вперед, сообщим, что появление «плохих сигналов» означает лишь то, что их исследование следует проводить иными методами, например, привлекая вероятностные модели. И, наконец, объединение этих подходов дает совершенно неожиданный эффект, известный под названием «принципа равномерной неопределенности» в практическом смысле означая, что с большой вероятностью существенная информация о сигнале (например об изображении или звуке) может быть воссоздана по совсем небольшой части сигнала или его спектра, это сравнительно новый подход, появившийся в начале XXI века. Мы постараемся объяснить (на несложных примерах) этот подход несколько позже, а пока вернемся к задаче обратимости преобразования Фурье то есть к соответствию сигнал-спектр в смысле поточечной сходимости в имеющемся тождестве Фурье.

Начнем здесь со сравнительно легкозапоминающегося варианта для периодических сигналов, тем более, что по традиции этот случай обычно входит в стандартные курсы математического анализа.

Периодический сигнал, случай C^2 -функций и поточечная сходимость

Напомним, что для функции $f \in C^p$ интегрирование по частям дает

$$\langle f, u_k \rangle = \int_0^1 f(t) e^{-2\pi i k t} dt = \frac{1}{2\pi i k} \int_0^1 f'(t) e^{-2\pi i k t} dt = \frac{1}{(2\pi i k)^p} \int_0^1 f^{(p)}(t) e^{-2\pi i k t} dt$$

и потому уже при $p = 2$ коэффициенты ряда Фурье убывают как сходящийся числовой ряд, а значит функциональный ряд Фурье равномерно сходится к некоторой непрерывной функции g . Если предположить, что разность $f - g = h \neq 0$, тогда в силу непрерывности квадратичная норма функции h ненулевая и равна некоторому $\delta > 0$, но ряд Фурье этой функции h тем не менее состоит из одних нулей $\forall k \langle h, e^{-ikx} \rangle = 0$, в частности функция h перпендикулярна любому тригонометрическому моному. Из стандартного курса мат.анализа (теорема Стоуна-Вейерштрасса) известно, что любая непрерывная функция может быть сколь угодно точно приближена по квадратичной норме (порожденной скалярным произведением) тригонометрическим полиномом $\sum c_k e^{-ikx}$ и тут возникает противоречие:

$$\|h - \sum c_k e^{-ikx}\|^2 = \|h\|^2 + \|\sum c_k e^{-ikx}\|^2 \geq \|h\|^2 = \delta$$

Периодический сигнал, разрывные функции

В случае разрывной периодической функции тождество Фурье вообще говоря неверно. Даже для кусочно непрерывных функций описание связи функции и ее спектра достаточно сложно. Например, на упражнениях мы рассматривали сумму ряда Фурье для кусочно-непрерывной с разрывом первого рода в точке x_0 функции $r(x)$ (эффект Гиббса). Приведем для этого случая замечание из книги В.И. Арнольда "Лекции об уравнениях с частными производными".

Замечание Можно рассмотреть в \mathbb{R}^2 непрерывные графики соответствующих последовательных приближений, которые сходятся поточечно, но не к графику функции, а к кривой, где над точкой x_0 вставлен вертикальный отрезок, который *длиннее* чем величина скачка в точке разрыва

$$|\lim_{x \rightarrow x_0+} r(x) - \lim_{x \rightarrow x_0-} r(x)|$$

При этом длины дополнительных вертикальных хвостиков, выступающих вверх и вниз за график исходной функции, составляют всегда одинаковую долю (примерно по 9% каждый) от величины скачка.

Для доказательства этого удивительного факта в эффекте Гиббса достаточно явно провести вычисления максимума и минимума частичной суммы ряда Фурье простейшей 2π периодической разрывной функции (скажем, равной x при $|x| < \pi$. Общий случай можно свести к этому, уничтожая разрывы вычитанием подходящей линейной комбинации сдвигов этой простейшей разрывной функции.

Явление Гиббса приводит к интересным артефактам в рентгеновской томографии (где функция ... рентгеновская оптическая плотность плоского сечения человеческого тела ...).

Поскольку плотность — разрывная функция (например, из-за существования костей), то возникает явление Гиббса. Оно проявляется здесь в виде дополнительных к линиям разрыва прямых: двойных касательных и касательных перегиба реальных границ тканей различной плотности.

Периодический сигнал, общие свойства

Предложение

Абсолютно интегрируемая на \mathbb{R} функции $f(x)$ непрерывна в среднем, то есть

$$\lim_{h \rightarrow 0} \int_{\mathbb{R}} |f(x+h) - f(x)| dx = 0$$

Доказательство Это заведомо верно для кусочно-постоянных функций. Интеграл от разницы между кусочно-постоянной функцией и $f(x)$ можно сделать сколь угодно малым (это, собственно и есть определение интегрируемости).

Теорема Римана об осцилляции

Пусть функция $f(x)$ абсолютно интегрируема на конечном или бесконечном интервале (a, b) . Тогда

$$\lim_{\omega \rightarrow \infty} \int_a^b f(x) \cos(\omega x) dx = \lim_{\omega \rightarrow \infty} \int_a^b f(x) \sin(\omega x) dx = 0$$

Доказательство Добавлением нулевых значений функции можно считать область интегрирования бесконечным интервалом. Тогда при $\omega \rightarrow +\infty$

$$I[\omega] = \int_{\mathbb{R}} f(x) \cos \omega x \, dx = - \int_{\mathbb{R}} f\left(x + \frac{\pi}{\omega}\right) \cos \omega x \, dx = -\frac{1}{2} \int_{\mathbb{R}} \left(f\left(x + \frac{\pi}{\omega}\right) - f(x)\right) \cos \omega x \, dx \rightarrow 0$$

Следствие

Коэффициенты ряда Фурье абсолютно интегрируемой на отрезке функции при увеличении (абсолютного) номера частоты стремятся к нулю. Для разложения 2π -периодической функции (длина периода здесь выбрана лишь для упрощения дальнейших формул) напомним

Определение Ядром Дирихле называется (очевидно, четная, непрерывная и 2π -периодическая) функция¹

$$D_n(x) = \frac{1}{2} + \sum_{k=1}^n \cos kx = \frac{\sin\left(\frac{2n+1}{2}x\right)}{2 \sin \frac{x}{2}}$$

ясно, что $\max_x |D_n(x)| = D_n(0) = n + 1/2$. Также

$$\frac{2}{\pi} \int_0^\pi D_n(t) dt = 1$$

¹Элементарное доказательство этого равенства получается последовательным применением формулы разности синусов

$$\sin A - \sin B = 2 \sin \frac{A-B}{2} \cos \frac{A+B}{2}$$

к ряду

$$\sin \frac{x}{2} + \left(\sin \frac{3x}{2} - \sin \frac{x}{2} \right) + \dots + \left(\sin \frac{(2n+1)x}{2} - \sin \frac{(2n-1)x}{2} \right)$$

Частичную сумму $S_n[f](x)$ ряда Фурье 2π -периодической функции можно выразить (восстановите пропущенную выкладку) через ядро Дирихле:

$$S_n[f](x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \left[\frac{1}{2} + \sum_{k=1}^n \cos k(t-x) \right] dt = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) D_n(t-x) dt$$

Производя в последнем интеграле (называемом интегралом Дирихле) замену переменной t на $t+x$ и сдвиг отрезка интегрирования, получим

$$S_n[f](x) = \frac{1}{\pi} \int_{-\pi}^{\pi} D_n(t) f(x+t) dt = \frac{1}{\pi} \left(\int_{-\pi}^0 + \int_0^{\pi} \right) D_n(t) f(x+t) dt = \frac{1}{\pi} \int_0^{\pi} D_n(t) [f(x+t) + f(x-t)] dt$$

Последнее выражение можно переписать с использованием разбиения отрезка интегрирования на участки $[0, \delta]$ и $[\delta, \pi]$, причем на втором участке подынтегральное выражение не имеет особенностей и по теореме Римана об осцилляции стремится при $n \rightarrow \infty$ к нулю. Тем самым вопрос о сходимости ряда Фурье сводится к вопросу о поведении первого интеграла при $n \rightarrow \infty$, это утверждение называется **принцип локализации**.

Принцип локализации

2π -периодическая абсолютно интегрируемая на (окружности) $[-\pi, \pi)$ функция f задает в каждом аргументе x_0 частичные суммы $S_n[f](x_0)$ своего ряда Фурье, причем пределы при $n \rightarrow \infty$ выражений

$$S_n[f](x_0) \quad \text{и} \quad \frac{1}{\pi} \int_0^{\delta} \frac{f(x_0+t) + f(x_0-t)}{2 \sin \frac{t}{2}} \sin \left(\left(n + \frac{1}{2} \right) t \right) dt$$

существуют или не существуют одновременно и совпадают в случае существования.

Поточечная сходимость ряда Фурье для непрерывной функции

Таким образом, ответ на вопрос о поточечной сходимости ряда в конкретной точке x_0 (коэффициенты, ряда, заметим, определяются по поведению функции *на всей окружности*) зависит от поведения функции *в сколь угодно малой* δ -окрестности точки x_0 , $\delta > 0$.

Условия на поточечную сходимость могут быть сформулированы по-разному, для практики довольно полезны условия², состоящие в существовании такого $\alpha \in (0, 1]$ и такой константы $C > 0$, что :

$$\begin{aligned} |f(x_0+h) - f(x_0)| &\leq C \cdot h^{\alpha}, & \forall h \in (0, \delta) \\ |f(x_0-h) - f(x_0)| &\leq C \cdot h^{\alpha}, & \forall h \in (0, \delta) \end{aligned}$$

Эти условия (порознь) называются *односторонними условиями Гельдера с показателем α для непрерывной функции*. Покажем, что их хватает для поточечной «правильной» сходимости.

$$\begin{aligned} S_n[f](x_0) - f(x_0) &= \frac{1}{\pi} \int_0^{\pi} D_n(t) (f(x_0+t) + f(x_0-t)) dt - f(x_0) \frac{2}{\pi} \int_0^{\pi} D_n(s) ds = \\ &= \frac{1}{\pi} \int_0^{\pi} \frac{f(x_0+t) + f(x_0-t) - 2f(x_0)}{2 \sin \frac{t}{2}} \sin \left(\left(n + \frac{1}{2} \right) t \right) dt = \\ &= \frac{1}{\pi} \int_0^{\pi} \left[\frac{f(x_0+t) - f(x_0)}{t} + \frac{f(x_0-t) - f(x_0)}{t} \right] \frac{t}{2 \sin \frac{t}{2}} \sin \left(\left(n + \frac{1}{2} \right) t \right) dt \end{aligned}$$

В силу свойства Гельдера в последнем выражении обе дроби в квадратных скобках являются абсолютно интегрируемыми (а почему? Проверьте.), следующий множитель также абсолютно интегрируем и потому по теореме об осцилляции выражение в целом стремится к нулю при $n \rightarrow \infty$. Что и гарантирует «правильную» сходимость.

В случае дискретного сигнала, даже и предполагая, что он является ограничением непрерывного, разбираться с его гипотетическими производными не очень ясно как; в то же время формулировка условия Гельдера вполне позволяет придать всему вполне наглядный смысл: зависимость $f(x_0+h) - f(x_0)$ от h иногда можно попробовать оценить на графике.

²Обозначения $f(x_0 \pm 0)$ отвечают как всегда односторонним пределам

Пример

При помощи генератора равномерно распределенной на $(-1, 1)$ случайной величины создадим 1000 значений x_k , а из них приготовим данные вида $y_m = 0.01 \cdot \sum_{k=1}^m x_k$. Что можно сказать об односторонних показателях Гельдера в точке с номером 500?

Информация

Тождество Фурье не обязательно выполнено в случае непрерывных функций, например, ряд Фурье может сходиться, но не к значениям исходной функции.

Сигнал непрерывного времени \mathbb{R} и тождество Фурье в частном случае

Сначала укажем на не вполне очевидное свойство: преобразование Фурье в случае \mathbb{R} переводит интегрируемые функции из $L^1(\mathbb{R})$ в непрерывные $C^0(\mathbb{R})$. Действительно, для параметра $\omega_0 \in \mathbb{R}$ и $\omega \sim \omega_0$

$$|\hat{f}(\omega) - \hat{f}(\omega_0)| = \left| \int_{\mathbb{R}} f(t) (e^{-i\omega t} - e^{-i\omega_0 t}) dt \right| \leq \int_{\mathbb{R}} |f(t)| \cdot |e^{-i\omega t} - e^{-i\omega_0 t}| dt \xrightarrow{\omega \rightarrow \omega_0} 0$$

Аналогично, обратное преобразование Фурье переводит интегрируемые функции из $L^1(\mathbb{R})$ в непрерывные $C^0(\mathbb{R})$.

Таким образом, можно проверить справедливо ли интегральное тождество Фурье для непрерывных функций

$$f(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f(u) e^{-i\omega u} du \right] e^{i\omega t} d\omega$$

если и прямое и обратное преобразование связаны с интегрируемыми (а значит и непрерывными) на \mathbb{R} функциями. Смена порядка интегрирования не применима: условия для этого³ в \mathbb{R}^2 не выполняются.

Поэтому рассмотрим двумерный интеграл с «локализацией» по аргументу ω :

$$I_\varepsilon(t) = \frac{1}{2\pi} \iint f(u) e^{-i\omega(t-u)} e^{-\varepsilon^2 \omega^2 / 2} d\omega du$$

для него уже выполнены условия перемены порядка интегрирования и потому возможно использовать теорему Фубини о приведении двойных интегралов к повторным. Если сначала проинтегрировать по u , то получим выражение

$$I_\varepsilon(t) = \frac{1}{2\pi} \int \hat{f}(\omega) e^{i\omega t} e^{-\varepsilon^2 \omega^2 / 2} d\omega \xrightarrow{\varepsilon \rightarrow 0} \frac{1}{2\pi} \int \hat{f}(\omega) e^{i\omega t} d\omega$$

Для функции $g(t) = 1/\sqrt{2\pi} \cdot e^{-t^2/2}$ рассмотрим положительную функцию $g_\varepsilon(u) = \frac{1}{\varepsilon} g\left(\frac{u}{\varepsilon}\right)$ интеграл которой равен единице, а значения вне ε - окрестности нуля очень малы. Мы ранее уже рассматривали интегральное преобразование Фурье от гауссианы, поэтому если интегрировать сперва по ω , то

$$I_\varepsilon(t) = \int f(u) g_\varepsilon(t-u) du$$

В силу непрерывности $f(\cdot)$ при $\varepsilon \rightarrow 0$ имеем $I_\varepsilon(t) \rightarrow f(t)$, то есть как раз тождество Фурье. Последнее соображение объясняет заодно конструкцию как можно «приблизить» δ -функцию Дирака с помощью последовательности гладких интегрируемых функций из $C^\infty(\mathbb{R})$.

Замечание

Несмотря на простоту и изящество сформулированного свойства очень простые сигналы не попадают под условия непрерывности, что порождает проблему: например, функция $\mathbf{1}_{[a,b]}(x)$ не является непрерывной, а ее преобразование Фурье не является в обычном смысле интегрируемым на всей прямой.

³условия Теоремы Фубини требуют, чтобы $f(s)e^{i\omega(t-s)} \in L^1(\mathbb{R}^2)$, а это не так

Сигнал и его спектр. Обобщения

Мы видели, что понимание тождества Фурье в смысле выполнения поточечного равенства приводит к необходимости формулировать сильные ограничительные условия на рассматриваемую функцию. Попытка избавиться от ограничений естественно приводит к рассмотрению функциональных пространств, в которых равенство между функциями понимается как равенство между классами *эквивалентных в подходящем смысле функций*. Мы начнем с пространства $L^2(\mathbb{R})$ функций (комплекснозначных) с конечным значением интеграла $\int_{\mathbb{R}} |f(t)|^2 dt$ по всей прямой от квадрата функции. Как известно, вместе со скалярным произведением и нормой

$$\|f\|_{L^2} = \sqrt{(f, f)} = \sqrt{\int_{\mathbb{R}} |f(t)|^2 dt}$$

Сходимость интеграла от квадрата функции вовсе не означает интегрируемости самой функции (а это было основное условие для рассмотрения преобразования Фурье), тем не менее, можно продолжить определение преобразования Фурье на элементы⁴ из L^2 .

Ключевое соображение состоит в равенстве Парсеваля, которое в дискретной модели выводится из того, что вектора базиса Фурье (как и вектора импульсного базиса) взаимно ортогональны и, значит, евклидова норма сигнала и его преобразования Фурье может быть очевидным образом выражена в обоих базисах. Приведем формулировку и доказательство в случае интегрального преобразования Фурье, оставив случай рядов Фурье для самостоятельного размышления.

Равенство Парсеваля

Для $f, g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ преобразование Фурье (и его обратное) является с точностью до множителя 2π изометрией:

$$(f, g) = \frac{1}{2\pi} (\hat{f}, \hat{g})$$

Для интегрируемых функций положим $h(t) = \overline{f(-t)}$

$$h * g(0) = \int h(-t)g(t)dt = \int \overline{f(t)}g(t)dt = (f, g)_{L^2}$$

А теперь используем свойство преобразования Фурье переводить свертки в произведения и наоборот

$$h * g(0) = \frac{1}{2\pi} \int \overline{\hat{f}(\omega)} \hat{g}(\omega) d\omega = \frac{1}{2\pi} (\hat{f}, \hat{g})_{L^2}$$

Распространение преобразования Фурье на L^2

Идея здесь проста: взять представителя из $f \in L^2$ и приблизить его последовательностью $\{f_n\}$, например, последовательностью «урезанных» функций f_n из $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ вида $\mathbf{1}_{\{-n \leq t \leq n\}} \cdot f(t)$. Надо проверить лишь, что последовательность преобразований Фурье \hat{f}_n (корректно определенных так как $f_n \in L^1$) имеет предел в L^2 , который и будет считаться преобразованием Фурье от f . Это технически не слишком сложно, поскольку в силу полноты пространства L^2 надо только проверить свойство Коши для этой последовательности⁵

Вспомогательные утверждения для обоснования:

1. $\|f_n\|_{L^1} \leq \sqrt{2n} \|f\|_{L^2}$ — по неравенству Коши-Шварца (проверить!).
2. Так как $f_n, f_m \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, то в силу равенства Парсеваля при $n \geq m$

$$\|\hat{f}_n - \hat{f}_m\|_{L^2}^2 = 2\pi \|f_n - f_m\|_{L^2}^2 \leq 2\pi \|\mathbf{1}_{\{m \leq |t| \leq n\}} f(t)\|_{L^2}^2 \rightarrow 0$$

и значит последовательность $\{\hat{f}_n\}$ фундаментальная и в силу полноты L^2 имеет предел.

⁴такой элемент предполагает неразличение некоторых функций, не совпадающих поточечно на множествах меры нуль

⁵На самом деле, выбор конкретных функций $f_n \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ для сходящейся к f последовательности Коши не критичен.

Обобщенные функции и преобразование Фурье

Здесь надо вспомнить, что обобщенные функции g определяют как линейные функционалы $g() = L_g()$ на пространстве так называемых пробных функций (т.е. бесконечно дифференцируемых и быстроубывающих со всеми своими производными, так что на пробной функции φ преобразование Фурье $\hat{\varphi}$ корректно определено и является также пробной функцией). По определению преобразование Фурье \hat{g} обобщенной функции g , определяют как линейный функционал $L_{\hat{g}}\varphi = L_g\hat{\varphi}$

Пример: преобразование Фурье от δ -функции Дирака

$$\hat{\delta}(\varphi) = \delta(\hat{\varphi}) = \hat{\varphi}(0) = \int \varphi(t)dt = \int 1 \cdot \varphi(t)dt = (\bar{1}, \varphi)$$

Тем самым, $\hat{\delta} = 1$

Традиционные функции также являются и обобщенными (линейный функционал определен через интеграл от произведения).

Пример: преобразование Фурье от константы

$$\hat{1}(\varphi) = \int 1 \cdot \hat{\varphi}(\omega)d\omega = 2\pi\varphi(0)$$

Поэтому $\hat{1} = 2\pi\delta$.

Легкое упражнение: преобразование Фурье от функции e^{ist}

Указание: для сдвинутой δ -функции $\delta_s(\varphi) = \varphi(s)$ имеем

$$\hat{\delta}_s(\varphi) = \delta_s(\hat{\varphi}) = \hat{\varphi}(s) = \int \varphi(t)e^{-ist}dt$$

Вот ответ $\widehat{e^{ist}} = 2\pi\delta_s$, восстановите рассуждение самостоятельно.

Трудное упражнение: преобразование Фурье от функции Хевисайда

Похожая на функцию sign функция Хевисайда, это «ступенька»:

$$H(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}$$

Найдите \hat{H} .

Преобразование Фурье и соотношение неопределенности, непрерывный сигнал

Речь снова пойдет о сравнении свойств сигнала и его спектра, однако сравнивать будем не $|\text{supp } f|$, а некоторую иную характеристику $\sigma^2(f)$ функции $f \in L^2(\mathbb{R})$ — «размах» (*spread* в оригинале), по сути относящуюся к теории вероятностей. Действительно, отношение $|f(t)|^2/\|f\|^2$ — неотрицательная функция, интеграл которой равен единице — тем самым это плотность распределения некоторой случайной величины ξ . Если дополнительно предположить, что у этой случайной величины существуют первый $E(\xi)$ и второй $E(\xi^2)$ моменты, то размах — это просто дисперсия $E((\xi - E(\xi))^2)$

$$\sigma^2(f) = \sigma^2(\xi) = \frac{1}{\|f\|^2} \int_{\mathbb{R}} (t - E(\xi))^2 |f(t)|^2 dt$$

Аналогично определяется $\sigma(\hat{f})$ и справедлив

Принцип неопределенности Гейзенберга-Вейля

Для $f \in L^2(\mathbb{R})$ такой, что $\sigma(f)$, $\sigma(\hat{f})$ корректно определены,

$$\sigma(f)\sigma(\hat{f}) \geq \frac{1}{2}$$

При этом равенство достигается в том и только в том случае, когда f пропорционально (сдвинутой) гауссиане. Заметим, что если в задействованных формулах интегралы расходятся, то принцип неопределенности по тривиальным основаниям следует считать также справедливым.

Доказательство

По свойствам преобразования Фурье (см. ранее приведенную таблицу свойств) шкалирование и сдвиг аргумента не меняют размаха, поэтому без ограничения общности можно предполагать (здесь мы, в отличие от дискретного случая выше не трогаем коэффициент $\frac{1}{2\pi}$)

$$\begin{aligned} \int_{\mathbb{R}} |f(t)|^2 dt = 1 &\Rightarrow \int_{\mathbb{R}} |\hat{f}(\omega)|^2 d\omega = 2\pi \\ \int_{\mathbb{R}} t |f(t)|^2 dt = \int_{\mathbb{R}} \omega |\hat{f}(\omega)|^2 d\omega = 0 \end{aligned}$$

Существенным упрощением будет рассматривать не произвольные, но гладкие функции, поскольку они плотны в $L^2(\mathbb{R})$. Продолжим теперь использование свойств преобразования Фурье теперь уже для гладких функций:

$$\sigma^2(\hat{f}) = \frac{1}{\|\hat{f}\|^2} \int_{\mathbb{R}} \omega^2 |\hat{f}(\omega)|^2 d\omega \Rightarrow \frac{1}{2\pi} \int_{\mathbb{R}} |i\omega \hat{f}(\omega)|^2 d\omega = \int_{\mathbb{R}} |f'(t)|^2 dt$$

что превращает соотношение неопределенностей между сигналом и его спектром в соотношение между сигналом и его производной:

$$\int_{\mathbb{R}} t^2 |f(t)|^2 dt \cdot \int_{\mathbb{R}} |f'(t)|^2 dt \geq \frac{1}{4}$$

Теперь применим неравенство Коши, к $g(t) = tf(t)$ и $h(t) = f'(t)$ с учетом того, что выше показано $\|g\| = \sigma(f)$ и $\|h\| = \sigma(\hat{f})$:

$$\begin{aligned} \left| \int_{\mathbb{R}} \overline{tf(t)} f'(t) dt \right| &\leq \sigma(f)\sigma(\hat{f}) \quad \text{и} \quad \left| \int_{\mathbb{R}} tf(t) \overline{f'(t)} dt \right| \leq \sigma(f)\sigma(\hat{f}) \\ \left| \int_{\mathbb{R}} \overline{tf(t)} f'(t) dt + \int_{\mathbb{R}} tf(t) \overline{f'(t)} dt \right| &\leq 2\sigma(f)\sigma(\hat{f}) \end{aligned}$$

Последнее выражение с использованием интегрирования по частям:

$$\int_{\mathbb{R}} \overline{tf(t)} f'(t) dt + \int_{\mathbb{R}} tf(t) \overline{f'(t)} dt = \int_{\mathbb{R}} t \frac{d}{dt} |f(t)|^2 dt = t |f(t)|^2 \Big|_{-\infty}^{+\infty} - \int_{\mathbb{R}} |f(t)|^2 dt = -1$$

что и дает $\sigma(f)\sigma(\hat{f}) \geq \frac{1}{2}$. Неравенство Коши превращается в равенство, если g пропорционально h , то есть $tf(t)$ пропорционально $f'(t)$. Последняя пропорциональность дает дифференциальное уравнение

$$f'(t) = -\frac{t}{\zeta^2} f(t) \quad \Rightarrow \quad f(0)e^{-t^2/2\zeta^2}$$

Обратная проверка того, что при гауссиане получается равенство можно проверить и непосредственно:

$$\begin{aligned} f(t) = e^{-t^2/2} &\Rightarrow |f(t)|^2 = e^{-t^2} \Rightarrow \sigma^2(f) = \frac{1}{2} \\ \hat{f}(\omega) = \sqrt{2\pi}e^{-\omega^2/2} &\Rightarrow |\hat{f}(\omega)|^2 = 2\pi e^{-\omega^2} \Rightarrow \sigma^2(\hat{f}) = \frac{1}{2} \end{aligned}$$

Дискретные интерпретации

Указанный вариант неопределенности можно перенести и на дискретный сигнал. Действительно, рассмотрим f и \hat{f} при единичной длине вектора (т.е. $\|f\| = 1$), квадраты координат $|f_k|^2$ задают величины вероятности на базисных ортах в любом ортонормированном базисе. Если сопоставить базисным ортам числовые значения (как это делается в следующем разделе про квантовомеханические интерпретации), то в каждом таком базисе будут определены дискретные случайные величины, в частности можно изучать соотношение неопределенностей для стандартного базиса и базиса Фурье в зависимости от размерности N и способа сопоставить базисным ортам числовые значения

Например, возьмем трехмерное пространство \mathbb{C}^3 , базису можно сопоставить «нумерацию» числами $-1, 0, +1$, понимаемыми как собственные значения (самосопряженного) оператора

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Половина разности двух сдвигов «вперед-назад» будет переводить $f_k \mapsto \frac{1}{2}(f_{k+1} - f_{k-1})$ и в дискретном времени такой оператор может быть соотнесен с вычислением «скорости», заметим, что с введением дополнительного множителя $\sqrt{-1}$ этот оператор станет самосопряженным.

$$\frac{1}{2} \begin{pmatrix} 0 & i & -i \\ -i & 0 & i \\ i & -i & 0 \end{pmatrix}$$

Найдите его собственные векторы и собственные значения. Эта конструкция лишь поясняет связь вещественнозначных случайных величин, базисными векторами эрмитова пространства и самосопряженными операторами. Если рассмотреть еще самосопряженный оператор проектирования на вектор базиса, то интуитивно можно считать, что такие операторы – описывают «дискретные координаты», а указанный оператор – приращение координаты за дискретное время, то есть своего рода «скорость», причем эта модель отвечает конечному множеству возможных значений.

Квантовомеханическая интерпретация

В нашем курсе не имеет смысла здесь повторять известные изложения как устроена математическая модель квантовой механики — другие люди уже сделали это гораздо лучше. Здесь достаточно напомнить, что одномерная модель квантовой частицы использует алгебру неограниченных самосопряженных линейных операторов в бесконечномерном евклидовом пространстве $L^2(\mathbb{R})$, а численные результаты физических измерений всегда излагаются в терминах меры на спектре¹ такого оператора. Подобный материал присутствует в множестве учебников, выделим здесь книгу Л.Фаддеев, О.Якубовский "Лекции по квантовой механике для студентов-математиков".

Известно, что в модели квантового осциллятора неограниченный самосопряженный оператор координаты \mathbf{Q} выбирается как $f(x) \mapsto xf(x)$, а неограниченный самосопряженный оператор импульса² это $f \mapsto -i\hbar \frac{\partial}{\partial x} f(x)$. Этот выбор задает обобщенные «собственные функции» этих операторов: соответственно $\delta(x - x_0)$ и e^{-ipx} ; мы ранее уже проверяли, что обобщенное преобразование Фурье действительно переводит эти функции друг в друга. Но эти обобщенные функции не лежат в $L^2(\mathbb{R})$ и потому матричное представление операторов \mathbf{P}, \mathbf{Q} выписывают в $L^2(\mathbb{R})$ -базисе из собственных векторов считающего оператора \mathbf{N} , экспонента $\exp\left(\frac{i\pi}{2}\mathbf{N}\right)$ которого задает в точности преобразования Фурье в пространстве $L^2(\mathbb{R})$. Также имеет место так называемое *каноническое коммутационное соотношение* со скобкой Пуассона $[\mathbf{Q}, \mathbf{P}] = i\mathbf{E}$, из которого можно вывести соотношение

¹здесь слово «спектр оператора» употреблено в общем смысле функционального анализа, а термины «собственные значения» и «собственные функции» имеют ограниченное применение

²здесь термин «импульс» выбран из описания механики движения, в первой лекции про конечномерные дискретные сигналы мы использовали термин иначе

неопределенностей Гейзенберга для дисперсий в любом состоянии *без отсылки к преобразованию Фурье*. Известно, что набор указанных свойств модели невозможно получить самосопряженными операторами в конечномерном евклидовом пространстве и потому использование $L^2(\mathbb{R})$ здесь существенно. Тем не менее конечномерные модели алгебры самосопряженных операторов также весьма полезны³ для физических приложений, хотя при этом и нарушаются указанные соотношения между \mathbf{P}, \mathbf{Q} в модели квантового осциллятора.

Тем самым, рассмотренная выше связь свойств функции и свойств ее преобразования Фурье имеет отношение к квантовому наблюдаемому, причем разброс отвечает понятию точности измерения. Попытка прояснить на конечномерном примере структуру операторов \mathbf{Q} и \mathbf{P} хотя и была предпринята выше, но этот конкретный пример не имел физического смысла.

Обратимся теперь к идее рассматривать возникающие меры на спектре оператора как распределения вероятностей (в смысле прикладных интерпретаций, использующих повторения физических экспериментов) возникающая при этом так называемая «теория квантовых вероятностей» или «квантовая теория вероятностей» имеет следующее, отличающееся от традиционного, описание:

1. Рассматривается линейное пространство L со скалярным произведением, для физических применений существенно также, чтобы все эти линейные пространства рассматривалось над полем комплексных чисел \mathbb{C} вместе с эрмитовым скалярным произведением, но здесь можно не обращать внимания на эту деталь, поскольку наша цель лишь в объяснении принципиальной разницы с традиционной теорией вероятностей.
2. Событиями называют теперь *произвольные линейные подпространства* $V \subset L$. Сумма событий $U \oplus V$ в квантовой теории вероятностей определяется как линейная оболочка объединения соответствующих подпространств, произведение $U \cdot V$ отвечает пересечению подпространств, противоположное событие \bar{U} отвечает ортогональному дополнению U^\perp .
3. Вероятность события U в квантовой теории вероятностей определяется через вектор состояния w — вектор единичной длины в L , а именно: вероятность равна квадрату длины $|\mathfrak{P}_U(w)|^2$ ортогональной проекции $\mathfrak{P}_U(w) \in U$ вектора состояния w на линейное подпространство U . Физики, рассматривая комплексный случай, называют вектор $\mathfrak{P}_U(w)$ *комплексной амплитудой события* U .

Ключевая разница между квантовой схемой и традиционной видна из решения следующей (и очень простой!) задачи, которую предлагается решить самостоятельно:

Задача

Проверьте для событий квантовой теории вероятностей тождества сложения и умножения событий (справедливые для булевой алгебры событий традиционной теории вероятностей):

$$A \cdot \left(\bigoplus_i B_i \right) = \bigoplus_i (A \cdot B_i) \quad \text{и} \quad A + \prod_i B_i = \prod_i (A + B_i)$$

Таким образом в квантовой модели вероятностей события булеву алгебру относительно операций сложения, умножения и взятия противоположного события *не образуют*. В частности, для квантового случая невозможно указать полную группу событий — разбиение на «альтернативные гипотезы» и потому рассуждать в терминах условных вероятностей становится затруднительно: мы имеем дело с неколомгоровской теорией вероятностей.

Однако физический микромир оказался в удивительном согласии именно с квантовой схемой вычисления вероятностей, хотя некоторые свойства таких событий будут выглядеть парадоксально, если неправильно интерпретировать события как элементы булевой алгебры. Кроме того, помимо обычной зависимости и независимости событий (определяемых так же, как и в традиционной теории вероятностей) имеет место еще совершенно иная связь событий, называемая *квантовой спутанностью*.

Повторения эксперимента и квантовая спутанность

Если L служит квантовой вероятностной моделью для одиночного эксперимента, то его квантовая модель повторений составляет линейное пространство $L^{\otimes m} = L \otimes L \otimes \dots \otimes L$ — тензорное произведение пространств.

- Какую размерность имеет пространство $L^{\otimes 4}$, если L двумерно?

³Hermann Weil "Theory of Groups Quantum Mechanics", Chapter 4, Section D, §14 Quantum Kinematics as an Abelian Group of Rotations, Dover Publications, Year: 1950

Векторы состояния в пространстве $L^{\otimes m}$, отличные от вида $w \otimes w \otimes \dots \otimes w$, порождают для квантовой теории вероятности новые эффекты зависимостей между квантовыми событиями в модели повторений. Эти зависимости называются *квантовой спутанностью событий*. Такие связи, вообще говоря, никак не могут быть проинтерпретированы в терминах традиционной теории вероятностей. Популярны сегодня идеи квантовой защиты каналов информации и квантовой криптографии используют именно квантовую спутанность.

Квантовые и традиционные вероятности вокруг нас

В курсах физики правила исчисления вероятностей в квантовом случае приводятся как правило без мотивирующих объяснений и употребляется вроде бы знакомая по традиционному подходу вероятностная терминология. Это порождает много известных из популярной литературы «парадоксов» микромира. Интересующиеся теорией исчисления квантовых вероятностей могут обратиться к книге А.Холево "Вероятностные и статистические аспекты квантовой теории".

Преобразования Радона и Фурье.

Интегральный случай и важное наблюдение

Для двумерного сигнала в \mathbb{R}^2 имеем интегральное преобразование Фурье; мы можем считать, что \mathbf{x} и $\boldsymbol{\nu}$ даны в прямоугольных координатах \mathbb{R}^2 , причем координатные оси удобно отождествить: $x_2 = 0$ с $\nu_2 = 0$ и $x_1 = 0$ с $\nu_1 = 0$. В зависимости от того, какой именно тип частот мы рассматриваем, **показатель у экспоненты и соответственно коэффициент при интеграле может варьироваться на дополнительный множитель, связанный с π или 2π — это обычное дело, когда речь идет о преобразовании Фурье.** Далее в нашем тексте также появятся эти доп.коэффициенты, единообразного определения мы придерживаться не станем.

$$\hat{f}(\boldsymbol{\nu}) = \int_{\mathbb{R}^2} f(\mathbf{x}) e^{-i(\mathbf{x}, \boldsymbol{\nu})} d\mathbf{x}$$

- Запишите обратное преобразование Фурье для этого случая

Как и ранее справедливо (но доказательство этого нетривиально) тождество Фурье в случае, когда обе функции $f(\mathbf{x})$, $\hat{f}(\boldsymbol{\nu})$ интегрируемы на \mathbb{R}^2 (и, стало быть, они и непрерывны в \mathbb{R}^2), именно этим случаем мы далее ограничимся. Основные свойства интегрального преобразования Фурье почти очевидно можно перенести на многомерный случай, отметим здесь версию свойства масштабируемости: для невырожденного линейного оператора $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^n$

$$F : f(\mathcal{A}\mathbf{x}) \mapsto |\det \mathcal{A}|^{-1} \hat{f} \left((\mathcal{A}^T)^{-1}(\boldsymbol{\nu}) \right)$$

Преобразование Радона

Интегральное преобразование Радона в его классическом смысле сопоставляет дифференцируемой и интегрируемой в \mathbb{R}^2 функции $f(\mathbf{x})$ семейство \mathcal{C} интегралов от ограничений этой функции на всевозможные (аффинные) прямые из \mathbb{R}^2 . У этого семейства аффинных прямых имеется естественная параметризация (t, θ) , с $-\infty < t < +\infty$ и $0 \leq \theta < \pi$ (r, ϕ), см. Рис.1: Результат преобразования Фурье в указанной параметризации называется *синограммой*.

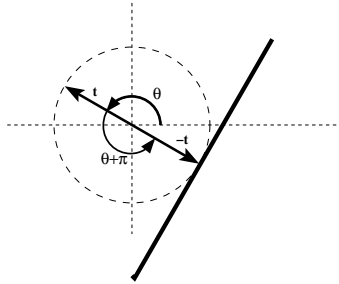


Рис. 1: Параметризация областью $\mathbb{R} \times [0, \pi)$ аффинных прямых в \mathbb{R}^2

граммой.

В качестве упражнения рассмотрим явное аналитическое вычисление преобразования Радона для простой функции $\mathbf{1}_{D^2}$ двух переменных

$$\mathbf{1}_{D^2}(x_1, x_2) = \begin{cases} 1 & x_1^2 + x_2^2 \leq 1 \\ 0 & x_1^2 + x_2^2 > 1 \end{cases}$$

- Запишите явно результат преобразования Радона $\mathfrak{R}[\mathbf{1}_{D^2}]$ от функции $\mathbf{1}_{D^2}$, с использованием указанной выше параметризации (t, θ) .
- Все ли верно на синограмме Fig.2 Lecture 9, §4 (лекции Elements of Modern Signal Processing).

Таким образом возникает преобразование \mathfrak{R} из пространства дифференцируемых и интегрируемых функций двух переменных $f(x_1, x_2)$ в функции $\check{f}(t, \theta)$ (каждой прямой сопоставляется интеграл по ней от f), это частный пример более общей конструкции преобразования, разработанного Дж. Радонем в 1917.

Связь с преобразованием Фурье: важное наблюдение

Рассмотрим ограничение $\hat{f}(\nu_1)$ непрерывной функции $\hat{f}(\nu_1, \nu_2)$ на прямую $\nu_2 = 0$; соответствующие значения совпадают с интегральным преобразованием Фурье $\int_{\mathbb{R}} \hat{f}(x_1) e^{-i\pi x_1 \nu_1} dx_1$ от функции $\check{f}(x_1) = \mathcal{P}[f](x_1) = \int_{\mathbb{R}} f(x_1, x_2) dx_2$ x_1 , которую уместно называть *интегральным проектированием* на координатную ось. Можно записать это символически как тождество операторов, используя обозначение \mathcal{K} для оператора ограничения функции двух переменных на координатную прямую: $\mathcal{K}F = F\mathcal{P}$. Исходя из того, что интегральное двумерное преобразование Фурье определено через скалярное произведение и потому не зависит от поворота систем координат, можно сделать вывод, что указанное тождество применимо к любой паре ортогональных прямых, проходящих через начало координат.

Теоретическая обратимость преобразования Радона

Заметим, что значения \check{f} задают на проходящей через начало координат прямой значения интегральных проекций $f(x_1, x_2)$ на эту прямую, а в силу «важного наблюдения» выше они задают и ограничение двумерного преобразования Фурье $\hat{f}(\boldsymbol{\nu})$ на эту же прямую. Таким образом, преобразование Радона позволяет восстановить сужение $\hat{f}(\boldsymbol{\nu})$ на любую прямую, проходящую через начало координат. Поскольку вся плоскость \mathbb{R}^2 содержится в объединении таких «центральных» прямых, то тем самым из $\check{f}(t, \theta)$ в аналитической версии можно восстановить целиком всю функцию $\hat{f}(\boldsymbol{\nu})$, а следовательно, благодаря двумерному тождеству Фурье, восстановить и исходную функцию $f(x_1, x_2)$. Однако реализация связана с некоторыми трудностями, например, аналитические формулы преобразования Радона не подсказывают очевидного дискретного грида на плоскости, на котором бы удобно вести вычисления. Другие трудности для реальных вычислений связаны с рассмотрением «не самых хороших» функций двух переменных.

Обратное к преобразованию Радона

Выше была установлена принципиальная теоретическая обратимость преобразования Радона. Более того, даже в языке python имеется функция, которая в принципе соответствует восстановлению образа по синограмме, однако в реальности эта функция часто «отказывается» указать ответ — на примере интегрального преобразования Радона на плоскости попробуем понять, какие вполне фундаментальные вопросы скрываются за этим.

Учитывая, что параметризацию множества афинных прямых в \mathbb{R}^2 определяли с помощью перпендикулярного направления \mathbf{n}_θ и параметра t . Для начала запишем двумерное обратное преобразование Фурье с использованием этих же обозначений, используя в интеграле полярные координаты r, φ :

$$f(\mathbf{x}) = \frac{1}{(2\pi)^2} \iint_{\mathbb{R}^2} \hat{f}(\boldsymbol{\nu}) e^{i(\boldsymbol{\nu}, \mathbf{x})} d\boldsymbol{\nu} = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{+\infty} \hat{f}(r\mathbf{n}_\varphi) e^{ir(\mathbf{n}_\varphi, \mathbf{x})} r dr d\varphi$$

Учитывая, что параметризацию множества афинных прямых в \mathbb{R}^2 определяли с помощью перпендикулярного направления \mathbf{n}_θ и параметра t перепишем последнее представление в этих же параметрах с небольшим трюком, относящимся к изначальной разнице в геометрии множеств допустимых параметров (r, φ) и (t, θ) .

$$f(\mathbf{x}) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{+\infty} \hat{f}(r\mathbf{n}_\varphi) e^{ir(\mathbf{n}_\varphi, \mathbf{x})} r dr d\varphi = \frac{1}{2\pi} \int_0^\pi \left[\frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(t\mathbf{n}_\theta) e^{it(\mathbf{n}_\theta, \mathbf{x})} |t| dt \right] d\theta$$

Как уже обсуждалось значения преобразования Фурье на проходящей через начало координат прямой $t\mathbf{n}_\theta$, $-\infty < t < +\infty$ можно связать с интегральными проекциями (значениями преобразования Радона $\mathfrak{R}[f](t, \theta)$) функции $f(\mathbf{x})$ вдоль параллельных \mathbf{n}_θ прямых.

Предположим на минуту, что мы знаем, что $|t|$ есть образ Фурье какой-то функции $h(s)$ одного переменного, Тогда выражение в квадратных скобках в последней формуле есть обратное преобразование Фурье от произведения двух функций (одного переменного s) в одномерной спектральной области (в которой параметром служит t), а значит, при каждом фиксированном \mathbf{x} и $\theta = \theta_0$ это свертка функций $\mathfrak{R}[f](s, \theta_0) = \mathfrak{R}[f] \Big|_{\theta_0}(s)$ и $h(s)$. Значение результата этой свертки вычисляют в аргументе $(\mathbf{n}_{\theta_0}, \mathbf{x})$. Окончательно: Таким образом,

$$f(\mathbf{x}) = \frac{1}{2\pi} \int_0^\pi \left(\mathfrak{R}[f] \Big|_{\theta} * h \right) (\mathbf{n}_\theta, \mathbf{x}) d\theta$$

Последнее выражение называется *профильтрованной обратной проекцией* и по сути представляет усреднение значений соответствующих сверток по всем прямым, проходящим через фиксированную точку $(x_1, x_2) = \mathbf{x} \in \mathbb{R}^2$ плоскости. Именно эта формула положена в основу процедуры `iradon` языка `python`. Более общим образом, *обратной проекцией* \mathcal{R}^* функции двух переменных $g(\mathbf{x})$ в точке \mathbf{x}_0 называется усреднение всех интегралов $\frac{1}{2\pi} \int_l g dl$ по всем прямым l , проходящим через точку \mathbf{x}_0 .

Но что же скрыто за обозначением $h(s)$? Совершенно ясно, что это никак не может быть обычная функция из $L^2(\mathbb{R})$, то есть ее надо искать среди обобщенных функций. Некоторая идея как она может устроена вытекает из формулы дифференцирования для преобразования Фурье «хороших» функций: $f^{(k)}(t) \mapsto (i\omega)^k \hat{f}(\omega)$, то есть соответствующий свертке фильтр будет связан с дифференцированием. Мы, однако, видели, что компьютерная реализация с приличной точностью для дифференцирования весьма затруднительна и потому, соответствующая процедура `iradon` для некоторых синограмм попросту прерывается из-за недостатка точности промежуточных вычислений. Второй метод (решения систем линейных уравнений) формально не использует формулы обратной проекции, но поскольку в основе теории обратного преобразования Радона лежит все-таки оператор типа дифференцирования, решение линейной системы также связано с влиянием неизбежных ошибок. Практическая реализация теоретической схемы обращения преобразования Радона хотя и не единственна, но любая будет чрезвычайно чувствительна к качеству дискретной синограммы (которая, как очевидно) также зависит от свойств исходного образа.

То, что оператор дифференцирования вообще говоря увеличивает L^2 -норму функции полезно сформулировать явно — это так называемое

неравенство Виртингера

Для хорошей 2π -периодической функции $u(t)$ такой, что $\int_0^{2\pi} u(t) dt = 0$

$$\int_0^{2\pi} [u(t)]^2 dt \leq \int_0^{2\pi} [u'(t)]^2 dt$$

при этом равенство возможно лишь при $u(t) = C_1 \cos(t) + C_2 \sin(t)$.

Доказательство использует разложение в ряд Фурье самой функции, формулы дифференцирования, выражающей ряд Фурье для производной и равенство Парсеваля, что приводит к

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} [u(t)]^2 dt &= \sum_{n \in \mathbb{Z}, n \neq 0} |c_n|^2 \\ \frac{1}{2\pi} \int_{-\pi}^{\pi} [u'(t)]^2 dt &= \sum_{n \in \mathbb{Z}, n \neq 0} n^2 |c_n|^2 \geq \sum_{n \in \mathbb{Z}, n \neq 0} |c_n|^2 \end{aligned}$$

неравенство, очевидно, строгое, если только при $|n| \geq 2$ все коэффициенты c_n не обращаются в ноль.

Замечание Равенство Парсеваля означает условия на L^2 -нормы в импульсном и спектральных областях. Естественный вопрос о версии для L^p -норм имеет ответом неравенства Юнга и Хаусдорфа, однако их доказательства достаточно длинные и мы ограничимся лишь формулировками¹. Итак, пусть для норм и на последовательностях и на функциях, в случае $1 < p \leq 2$ и $1/q + 1/p = 1$

$$\left[\frac{1}{2\pi} \int_{-\pi}^{\pi} |f|^q dx \right]^{\frac{1}{q}} \leq \left[\sum_{n \in \mathbb{Z}} |c_n|^p \right]^{\frac{1}{p}} \quad \text{и} \quad \left[\sum_{n \in \mathbb{Z}} |c_n|^q \right]^{\frac{1}{q}} \leq \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} |f|^p dx \right]^{\frac{1}{p}}$$

Возможная трудность использования программ вычислений

Для вычислений нет иного выхода, кроме как использовать «дискретную модель мира» — в нашем случае это означало схему дискретного преобразования Фурье. Поскольку реальные задачи описания действительности обычно сформулированы на аналитическом языке, мы вынужденно приходим к необходимости дискретной интерпретации аналитических формул, на этом пути есть масса сложностей и сейчас мы отметим еще одну: аналитические формулы могут включать выражения, связанные с обобщенными функциями. Из прошлого занятия известно, что в принципе скрывается за «обобщенностью»: линейный (сверточный) оператор в функциональном пространстве, линейный функционал или предел по подходящей норме последовательности линейных функционалов. Полезный практический навык заключается в том, чтобы уметь опознавать в (не всегда корректно выписанных) аналитических формулах их области определения. Упомянем сравнительно простой пример.

¹ Доказательства частных случаев можно найти в классической книге Г.Харди, Дж.Литлвуд и Г.Пойа «Неравенства»

Сингулярные функционалы и сингулярные фильтры. Пример регуляризации

Интерпретация несобственных интегралов как функционалов на пространстве основных (то есть C^∞ -функций $\varphi(x)$ с компактными носителями) может представлять трудность. Так, например, функционал, заданный интегралом² с функцией $f(x) = 1/|x|$ не определен для функций $\varphi(x) \in C^\infty\mathbb{R}$, имеющих ненулевые значения в точке $0 \in \mathbb{R}$; действительно, в этом случае соответствующий интеграл расходится. Стандартная техника устранения изолированных особенностей предлагает замену (заметим, не единственную) этого функционала на *регуляризованный*. Принятый подход (за более подробной информацией стоит заглянуть в книгу И.М.Гельфанд, Г.Е.Шиллов "Обобщенные функции и действия над ними") состоит в рассмотрении линейного функционала, заданного на $\varphi(x)$ другой формулой, а именно

$$(f, \varphi) = \lim_{\varepsilon \rightarrow 0} \left[\int_{-\infty}^{-\varepsilon} \frac{\varphi(x)}{|x|} dx + \int_{-\varepsilon}^{\varepsilon} \frac{\varphi(x) - \varphi(0)}{|x|} dx + \int_{\varepsilon}^{\infty} \frac{\varphi(x)}{|x|} dx \right]$$

Очевидна неоднозначность регуляризации и здесь и в более общем случае: например, числитель в среднем члене можно бы заменить на $\varphi(x) - \left[\varphi(0) + \frac{\varphi'(0)}{1!}x + \dots + \frac{\varphi^{(m)}(0)}{m!}x^m \right]$. Таким образом, имеются, вообще говоря, разные регуляризации для расходящихся интегралов, а, стало быть и несколько вычислительных версий, связанных с дискретизацией соответствующей аналитической формулы. В силу этой глобально неистребимой неоднозначности ценится умение придумать удобный вычислительный алгоритм для *конкретного типа данных*.

Пример. А можно ли верить компьютерным вычислениям?

Может сложиться впечатление, что применение понятных алгоритмов линейной алгебры не несет в себе никаких опасностей. Например, рассматривая задачу поиска собственных значений невырожденной самосопряженной матрицы небольшой размерности, большинство пользователей не станут вникать в смысл ответа, тем более, что такая задача обычно является промежуточным этапом в последовательности вычислений какого-то большого проекта. Рассмотрим однако сильно демотивирующий пример: попробуйте найти функцию `eig`, которая дает (легко понятный изначально) ответ.

Рассмотрим две матрицы — нижнетреугольную **C** и верхнетреугольную **T**:

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \quad \mathbf{T} = \begin{pmatrix} 1 & 2048 & 256 & 128 & 64 & 32 & 16 \\ 0 & -2 & 1024 & 512 & 256 & 128 & 32 \\ 0 & 0 & 4 & 512 & 1024 & 256 & 64 \\ 0 & 0 & 0 & 8 & 512 & 512 & 128 \\ 0 & 0 & 0 & 0 & -4 & 1024 & 256 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2048 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

Очевидно, что они обе невырождены, причем собственные числа матрицы **C** единичные, а собственные числа **T** суть $\pm 1, \pm 2, \pm 4, 8$. Рассмотрим матрицу **R**:

$$\mathbf{R} = \mathbf{C}^{-1}\mathbf{TC} = \begin{pmatrix} 289 & 2064 & 336 & 128 & 80 & 32 & 16 \\ 1152 & 30 & 1312 & 512 & 288 & 128 & 32 \\ -29 & -2000 & 756 & 384 & 1008 & 224 & 48 \\ 512 & 128 & 640 & 8 & 640 & 512 & 128 \\ 1053 & 2256 & -504 & -384 & -756 & 800 & 208 \\ -287 & -16 & 1712 & -128 & 1968 & -30 & 2032 \\ -2176 & -287 & -1565 & -512 & -541 & -1152 & -289 \end{pmatrix}$$

очевидно, что ее собственные числа такие же как и у **T**. Однако, если попробовать рассчитать эти числа непосредственно, подавая на вход (например функции `numpy.linalg.eig`) матрицу **R** ответ получится несколько неожиданный — проверьте какой.

Линейные примеры

На самом деле, предыдущий пример всего лишь показывает опасность использования готового алгоритма «вслепую». Откуда вообще может возникать проблема? Напомним на двумерном примере общеизвестные факты о принципах решений систем линейных уравнений в тех случаях, когда простой подход не работает.

²См. также также функционал, определенный преобразованием Фурье функции $\text{sign}(x)$

Недоопределенная система

Речь идет о ситуации, когда уравнений меньше, чем неизвестных, например, об уравнении $ax_1 + bx_2 = c$, которое определяет прямую в \mathbb{R}^2 . Дополнительным условием, позволяющим избавиться от бесконечного числа решений, обычно служит условие близости к конкретной точке³ $\hat{\mathbf{x}} \in \mathbb{R}^2$; геометрически это означает ортогональную проекцию точки $\hat{\mathbf{x}}$ на прямую $ax_1 + bx_2 = c$

Несовместная система

Матрица системы $\mathbf{A} \cdot \mathbf{x} = \mathbf{c}$ имеет ранг, меньший размерности вектора неизвестных, например $\begin{pmatrix} 1 & 2 \\ 1.05 & 2.1 \end{pmatrix}$, а $\mathbf{c} = \begin{pmatrix} 10 \\ 20 \end{pmatrix}$. Основная идея: рассматривать *невязку* $|\mathbf{A} \cdot \mathbf{x} - \mathbf{c}|$, выбирать то множество *псевдорешений* \mathbf{x} , при которых достигается минимум невязки, а далее использовать соображение о минимизации нормы псевдорешения,

Плохо обусловленная система

Здесь речь идет о системе уравнений с матрицей \mathbf{A} с $\det \mathbf{A} \neq 0$ и $|\mathbf{A}| \cdot |\mathbf{A}^{-1}| \gg 1$ (*число обусловленности*). Несмотря на то, что плохо обусловленные системы имеют единственное решение, на практике искать это решение чаще всего не имеет смысла. Вот конкретный пример: Рассмотрим свойства плохо обусловленных систем линейных уравнений на двух конкретных примерах очень близких матриц и с одинаковой правой частью:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 6.01 \end{pmatrix} \quad \mathbf{T} = \begin{pmatrix} 1 & 2.01 \\ 3 & 6 \end{pmatrix} \quad \mathbf{c} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

Несмотря на близость этих систем, их точные решения оказываются очень далекими друг от друга, а именно:

$$\begin{pmatrix} 803 \\ -400 \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} -265 \\ \frac{400}{3} \end{pmatrix}$$

Прямые, соответствующие плохо обусловленной системе, располагаются в непосредственной близости друг от друга (почти параллельны). В связи с этим, малые ошибки в расположении каждой из прямых могут привести к значительным погрешностям локализации точки их пересечения (решения) в противоположность случаю хорошо обусловленной системы, когда малые погрешности в наклоне прямых мало повлияют на место точки их пересечения. Речь не идет о нарушении непрерывной зависимости от ошибок, но вычислительные действия ведутся с некоторой достижимой точностью и потому на ε - δ языке зависимость δ от ε в условиях фиксированной точности играет большую роль.

Для решения некорректных задач, в частности, вырожденных и плохо обусловленных систем линейных уравнений, разработан эффективный метод, называемый регуляризацией, на основе учета дополнительной априорной информации о структуре решения, которая очень часто имеется в практических случаях.

Регуляризация, линейный случай

Метод, называемый регуляризацией основан на привлечении дополнительной априорной информации о решении, которая может быть как качественной, так и количественной. Достаточно часто встречающийся вид некорректной задачи, в которой (неявно) предполагается слагаемое ε как правило допускающее лишь вероятностное описание

$$\mathbf{A} \cdot \mathbf{x} + \varepsilon = \mathbf{c}$$

Например, можно искать решение, максимально близкое к некоторому вектору $\hat{\mathbf{x}}$ и тогда исходную задачу заменяют на поиск минимума квадратичного функционала, зависящего от параметра: λ .

$$\Psi(\mathbf{x}, \lambda) = |\mathbf{A} \cdot \mathbf{x} - \mathbf{c}|^2 + \lambda \cdot |\mathbf{x} - \hat{\mathbf{x}}|^2$$

Достаточно легко понять, что значение $\tilde{\mathbf{x}}$ в минимуме функционала (которое и принимают за регуляризованное решение) зависит от λ , подбор которого в значительной степени основывается на предположениях о характере добавки ε . В указанном линейном случае поиск минимума сводится к системе линейных уравнений

$$(\mathbf{A}'\mathbf{A} + \lambda\mathbf{E}) \cdot \mathbf{x} = \mathbf{A}' \cdot \mathbf{c} + \lambda\hat{\mathbf{x}}$$

а в более общем нелинейном случае поиск минимума можно вести, например, методом градиентного спуска.

³Часто выбирают $\hat{\mathbf{x}} = 0$

Пример некорректно поставленной задачи

Рассмотрим задачу обращения оператора сглаживания гауссовским фильтром, иными словами дана функция $f : \mathbb{R} \rightarrow \mathbb{R}$ и имеются данные о ее сглаживании $g = f * \varphi$ гауссовским фильтром

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}; \quad \hat{\varphi}(\omega) = e^{-\frac{\omega^2}{2}}$$

Искомую функцию f легко выразить через преобразование Фурье от данных: $\hat{f}(\omega) = \hat{g}(\omega) e^{\frac{\omega^2}{2}}$. Понятно, что любая аддитивная шумовая добавка $\varepsilon(x)$ дает аддитивную добавку, вклад которой в формулу обращения уже не обязательно мал: $\hat{\varepsilon}(\omega) e^{\frac{\omega^2}{2}}$. Поскольку обратное преобразование Фурье с точностью до множителя — изометрия, ошибка для найденного решения может быть также велика. Грубо говоря, обратная задача некорректна, когда решение не является непрерывным по отношению к исходным данным.

Поиск регуляризованных решений для обратного преобразования Радона

В качестве решения $\mathbf{u} \in L_2$ задачи обращения (дискретных, естественно,) данных g преобразования Радона \mathfrak{R} нам нужен физически осмысленный дискретный сигнал $f \in L_2$, у которого $\|f\| < \infty$ то есть можно вести поиск **регуляризованного решения** из выражения, включающего настроечные параметры λ и функционал $P(f)$. например

$$\min_f [\|g - \mathfrak{R}[f]\| + \lambda P(f)]$$

где функционал $P(f)$ называется регуляризацией и вообще говоря, определяется контекстом. Если говорить о недискретном сигнале, то распространены следующие варианты регуляризаций.

1. $P(f) = \|f\|$ — L_2 - норма, а минимум $\min_f \|f\|$ ищут при условии $\|g - \mathfrak{R}[f]\| \leq \varepsilon$ — это часто применяемая регуляризация в смысле А.Н.Тихонова.
2. Вариант предыдущего, но $P(f)$ — соболевская k - норма
3. Вариант предыдущего, но $P(f) = L_1$ - норма.
4. Вариант предыдущего, но $P(f) = L_1$ и/или L_2 - норма не самого, но предварительно профильтрованного (например, полосовым фильтром) сигнала.

На практике априорное знание свойств решения (например, в медицинской томографии) служит аргументом для выбора функционала, при этом универсального подхода не существует.