Согласованная фильтрация. Корреляционный прием

30.12.2020, 22 с

В данном пособии объясняются принципы согласованной фильтрации и корреляционного приема, которые дают один и тот же теоретический результат, и используются в большинстве радиотехнических систем: в радиолокации, радионавигации и цифровой связи. Дополнительно, корреляционный прием является «кирпичиком» при извлечении информации с помощью нейросетей. Последние, хоть и основаны на классике согласованной фильтрации, заслуживают отдельного места в «чемоданчике» инженера ввиду наличия определенного искусства построения нейросетей.

Пособие содержит упражнения, ориентированные на самостоятельную работу.

Пособие ориентировано на студентов старших курсов высших технических учебных заведений.

Оглавление

Введение

Каким бы сложным ни казалось название данного пособия, в основе согласованной фильтрации и корреляционного приема лежит такое школьное понятие как **скалярное произведение** двух векторов, которое является вещественным числом, пропорциональным длине этих векторов и степени их сонаправленности (коллинеарности). Эта степень определяется косинусом угла между векторами, однако, косинус можно интерпретировать как коэффициент корреляции $r = \cos\varphi$, откуда и следует понятие корреляционного приема. Смысл «согласованности» будет ясен после изучения всего пособия.

Слово «фильтрация» подразумевает наличие трех сущностей: фильтра, входного воздействия (входного сигнала) и соответствующего отклика (выходного сигнала). Математически согласованная фильтрация и корреляционный прием дают один и тот же результат: корреляцию (меру похожести) принятого сигнала и опорного. Разница между этими методами заключается в технической реализации: корреляционный прием реализуется напрямую (wi-fi и другая цифровая радиосвязь), а согласованная фильтрация, как правило, через быстрое преобразование Фурье (например, в радарах «круиз-контроля»).

Блок, который реализует корреляционный прием, называют **коррелятором.** Согласованную фильтрацию выполняет **согласованный фильтр,** или «конвольвер» от слова convolve – сворачивать.

Принцип согласованной фильтрации следует из практической необходимости извлечения полезной информации из принимаемого сигнала. Например, наблюдая показание термометра (набор цветовых сигналов), можно извлечь некоторое число: 36.7. В качестве согласованного фильтра (коррелятора) здесь выступает мозг, хранящий эталоны цифр и умеющий сравнивать принятый сигнал с набором опорных (эталонных). В результате сравнения принимается два решения: сначала о наборе цифр, а затем о наличии/отсутствии отклонения от нормы. Норма является некоторым **пороговым значением** (порогом).

Рассматривая радиолокационные системы, можно утверждать, что радиолокатор (радар) анализирует результат зондирования (принятый отраженный радиосигнал) и выносит решение о наличии/отсутствии цели в заданной области пространства: так решается **задача обнаружения** полезного сигнала на фоне помехи. В случае радионавигационных систем gps-приемник принимает решение о наличии/отсутствии сигнала со спутников, и по результату решения декодирует принимаемые сигналы для вычисления собственных координат. В случае цифровых систем передачи информации приемник, наблюдая отрезок сигнала на своем входе, принимает решение о значении переданного символа, который в простейшем и важном случае является битом (двоичным символом), 0 или 1: так решается **задача различения** двух сигналов на фоне помехи (или задача приема двоичного сигнала на фоне помехи).

Согласованная фильтрация восходит к работам академика В.А. Котельникова, указавшего на способ оптимальной обработки аддитивной смеси полезного сигнала и белого гауссовского шума. Оптимальность понимается в смысле минимизации полной (итоговой) вероятности ошибки при принятии решения. Одновременно обеспечивается максимум отношения сигнал-шум после обработки (на выходе фильтра). Также известно, что структура согласованного фильтра следует из решения задачи поиска линейного фильтра, дающего максимальное отношение сигнал-шум на своем выходе после обработки аддитивной смеси полезного сигнала и белого шума (требование гауссовости здесь убрано). Линейный фильтр, как известно, полностью характеризуется импульсной характеристикой – реакцией фильтра на единичный импульс: дельта-функцию для случая непрерывного времени, или «единичку» для случая дискретного.

В данном пособии излагаются необходимые для понимания корреляционного приема элементы теории линейной фильтрации, дается понятие производной по вектору, после чего с помощью введенных инструментов элегантно решается задача поиска линейного фильтра, дающего максимальное отношение сигнал-шум после обработки аддитивной смеси полезного сигнала и окрашенного (т. е. не обязательно белого) стационарного шума.

Элементы теории линейной фильтрации

Всякая фильтрация подразумевает преобразование входного воздействия (сигнала) в выходное (в отклик). Линейная фильтрация осуществляет взвешенное суммирование «частичек» сигнала, поступающего на вход фильтра. Линейный фильтр устроен так, что он помнит предыдущие отсчеты сигнала и использует их для формирования выходного сигнала. Насколько долго он помнит предыдущий сигнал, определяется структурой фильтра. В качестве примера фильтра можно привести заряженный конденсатор емкостью C (отключенный от источника питания), к которому подсоединили нагрузочный резистор: в этом случае в течение некоторого времени τ через сопротивление R будет течь ток – отголосок начального напряжения (от источника питания). Такая RC-цепь будет экспоненциально забывать начальное воздействие; время, через которое фильтр забудет, пропорционально емкости конденсатора и сопротивлению резистора, а в более общем случае и начальному напряжению, потому что, применяя этот фильтр в некоторой системе, можно установить пороговый уровень напряжения (тока), который следующее за фильтром звено не чувствует (чувствительность радиотехнического звена).

Процесс взвешенного суммирования описывается интегралом

$$v(t) = \int_{-\infty}^{\infty} h(\tau) s(t-\tau) d\tau,$$

который называется **интегралом свертки** (или интегралом Дюамеля). При подаче на вход единичного импульса (дельта-функции) $s(t) = \delta(t)$ интегрирование дает следующий результат

$$v(t) = h(t),$$

т. е. $h(t)$ является откликом фильтра на единичное воздействие, а это по определению есть **импульсная характеристика** фильтра. Здесь использовалось **фильтрующее (отсчетное) свойство** дельта-функции

$$\int_{-\infty}^{\infty} f(t) \cdot \delta(t-\tau) dt = f(\tau).$$

Дельта-функция, хотя и не являющаяся обычной функцией, по определению имеет единичный образ, который является обычным числом

$$\int_{-\infty}^{\infty} 1 \cdot \delta(\tau) d\tau = 1.$$

Это означает, что площадь под этим очень коротким импульсом (в идеале нулевой длительности) равна единице. Дельта-импульс можно аппроксимировать выражением типа синус Котельникова (но не только им)

$$\delta_{approx.}(t) = \frac{\sin \pi a t}{\pi t} = a \operatorname{sinc}(at)$$

при большом параметре $a \gg 1$, Рисунок 1.

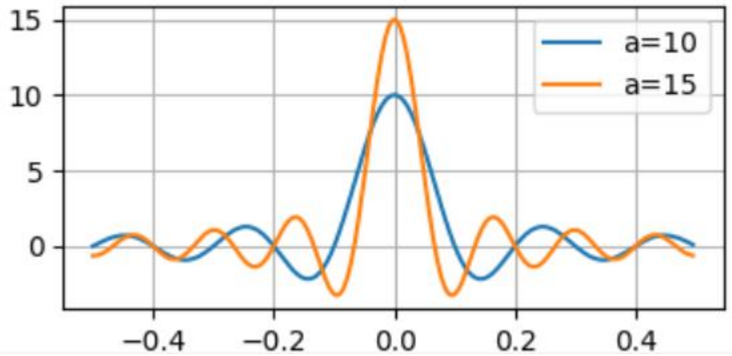


Рисунок 1 Аппроксимация дельта-функции масштабированным синусом Котельникова

Структуру интеграла свертки поясняет Рисунок 2.

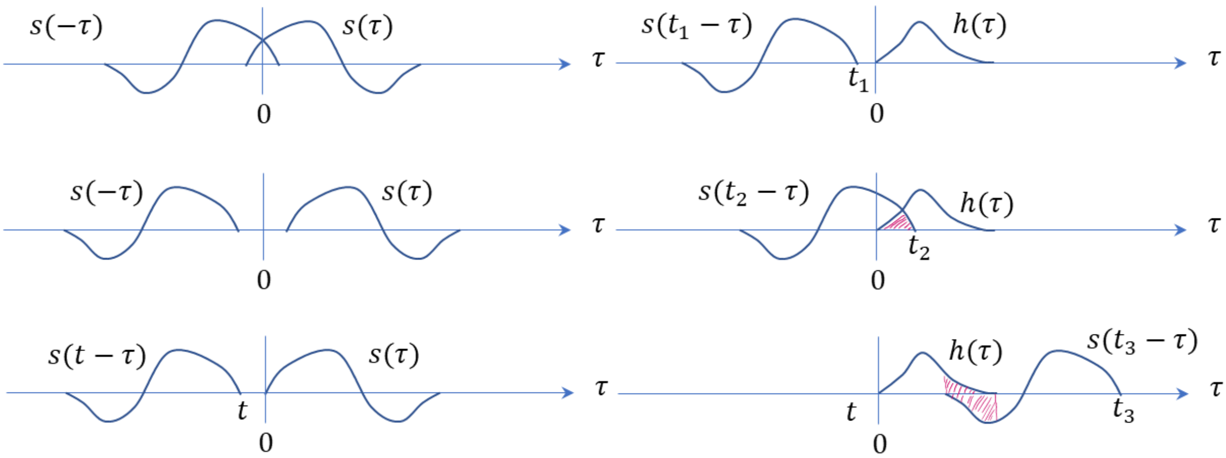


Рисунок 2 К пониманию зеркального отражения входного сигнала $s(\tau)$, его сдвига и свертки с импульсной характеристикой фильтра $h(\tau)$

Здесь в левом нижнем углу сигнал $s(\tau)$ начинается в нулевой момент времени (это удобно), а $t_1 < t_2 < t_3$ справа означают моменты времени, иллюстрирующие движение зеркально-отраженного входного сигнала вправо по оси времени τ. Заштрихованным показана область взаимодействия (т. е. ненулевого произведения), дающая вклад в формирование отсчета выходного сигнала в момент времени t. Так как у любого практического сигнала есть начало, то из рисунка ясно, что интегрировать имеет смысл до момента t, поэтому чаще интеграл Дюамеля записывают так

$$v(t) = \int_{-\infty}^{t} h(\tau)s(t-\tau)d\tau.$$

Это отражает **свойство причинности:** отклик не может появиться до начала действия сигнала (нулевой уровень – это не сигнал!). Вдобавок, если принять, что импульсная характеристика фильтра начинается с нулевого момента времени, то интеграл становится гораздо удобнее с практической точки зрения (однако при этом следует быть осторожным, и в случае необходимости корректно изменять нижний предел интегрирования)

$$v(t) = \int_{0}^{t} h(\tau)s(t-\tau)d\tau, \text{ это частный случай!}$$

В целом ситуация такова, что для практики удобнее конечные пределы интегрирования (программы численного моделирования работают корректнее), а для теории – бесконечные (упрощается доказательство свойств, например, связанных с преобразованием Фурье).

Почему линейная фильтрация так популярна? Дело в том, что радиотехнические системы устроены так, что информация кодируется и передается с помощью набора определенных гармонических составляющих (это основано на рядах Фурье), и образно процесс передачи можно сравнить с размазыванием битов (абстрактных объектов) по некоторой полосе частот (по диапазону частот – физическому объекту). Линейная фильтрация примечательна тем, что сохраняет гармонический состав входного воздействия, т. е. не создает новых (комбинационных) частот, при этом ролью линейной фильтрации в приемнике является анализ определенной полосы частот и максимально достоверный сбор исходных битов – своего рода «вытягивание» информации путем накопления сигнала и соответствующего снижения влияния помехи (шума).

Глядя на интеграл свертки, можно сказать, что линейный фильтр единовременно выполняет три базовых операции: задержку, умножение и суммирование. В цифровом мире эти операции реализуются напрямую в виде отдельных блоков (регистры сдвига, умножители и сумматоры), а в аналоговом мире – в виде емкостей, индуктивностей и сопротивлений (конденсаторов, катушек и резисторов соответственно).

Рассмотрим пример фильтрации прямоугольного импульса некоторым линейным фильтром с заданной импульсной характеристикой, Рисунок 3. В момент, когда происходит резкое нарастание фронта входного импульса, фильтр начинает откликаться и выходной сигнал плавно нарастает. Наличие постоянного входного уровня (+4 в данном случае) равнозначно наличию множества подряд идущих единичных импульсов (с очень малым интервалом по времени Δt), которые возбуждают фильтр, и он, вследствие этого, формирует импульсные характеристики, разнесенные на этот малый момент времени Δt и суммирующиеся друг с другом (с весом Δt). Далее наступает момент, когда выходной сигнал фильтра не изменяется – фильтр выходит на стационарный режим, при этом сигнал пропорционален сумме уровней импульсной характеристики, или интегралу от нее – площади под кривой с учетом знака.

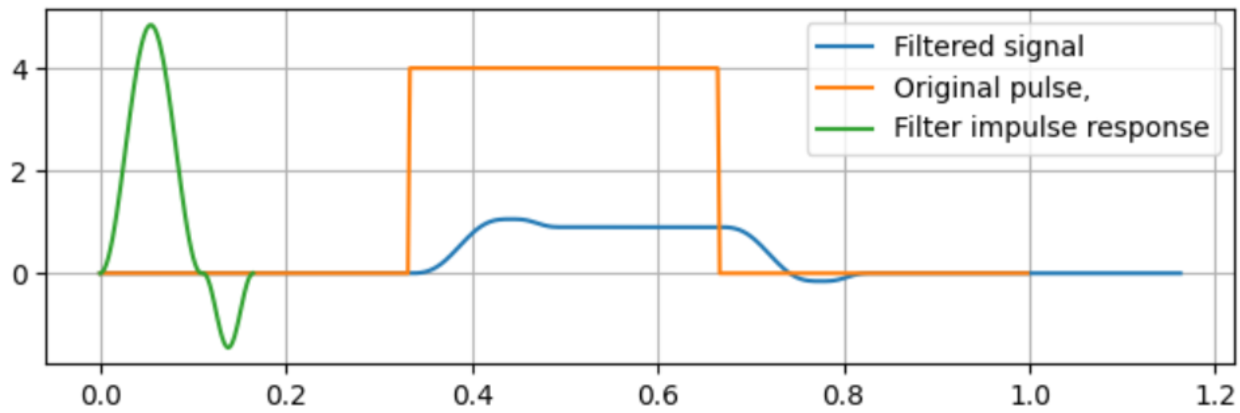


Рисунок 3 Результат фильтрации прямоугольного импульса Original pulse линейным фильтром с импульсной характеристикой Filter impulse response

Так как среднее значение импульсной характеристики в данном случае положительно, то стационарный уровень (отклик на «ступеньку» при большом времени, т. е. при $t \to \infty$) также положителен. При этом говорят, что фильтр пропускает постоянную составляющую с некоторым ненулевым коэффициентом передачи $K(0) = \left(\frac{1}{T}\right)\int_0^T h(\tau)d\tau$; также говорят, что фильтр имеет открытый вход (не имеет разделительных конденсаторов на входе). Если же коэффициент передачи меньше нуля, то фильтр имеет склонность к инверсии низкочастотных компонент входного сигнала; в этом случае коэффициент передачи нагляднее представить в виде комплексного числа в экспоненциальной форме, допустим, $K(0) = -\frac{1}{4} = \left(\frac{1}{4}\right)e^{j\pi}$. Аргумент π говорит о повороте фазы соответствующих компонент на $180°$ – так себя ведут некоторые транзисторные усилители мощности: там при нарастании входного напряжения базы транзистора он постепенно открывается, вследствие чего постепенно понижается выходное (коллекторное) напряжение, и наоборот.

В момент, когда начинается резкий спад фронта входного импульса, Рисунок 3, выходной сигнал начинает постепенно уменьшаться, и в итоге наступает момент последнего «вздоха» фильтра: появляется отрицательный «лепесток» импульсной характеристики и отклик постепенно затухает (переходный процесс заканчивается).

Процесс, при котором фильтром сглаживаются фронты прямоугольного импульса, называется **завалом фронта,** при этом говорят, что фильтр **заграждает** (сильно ослабляет) верхние частоты, т. е. препятствует чрезмерно быстрым изменениям выходного сигнала. В реальности любое радиотехническое звено, будь то антенна, волновод, кабель, транзистор и т. п., заграждает верхние частоты, вопрос только в диапазоне этих частот: килогерцовый, мегагерцовый, гигагерцовый и т. д. Частота среза, в целом, пропорциональна обратной длительности импульсной характеристики $\frac{1}{\tau_{IR}}$. Если наложить на импульсную характеристику фильтра некоторый гармонический процесс достаточно высокой частоты $f \gg \frac{1}{\tau_{IR}}$, то очевидно, что среднее значение произведения этих двух сигналов $K(f) =$

$\frac{1}{T}\int_0^T h(\tau)\sin(2\pi f\tau+\varphi)\,d\tau$ будет близко к нулю, какой бы высокой частота ни была – это и говорит о подавлении высоких частот данным фильтром.

Если рассматривать обработку периодических сигналов $s(t+nT)=s(t)$, $n\in\mathbb{N}$, $t\geq 0$, то отклик в установившемся режиме будет иметь тот же период T. В установившемся – это когда входной сигнал заполнил всю достаточно длинную (возможно, что больше периода сигнала) импульсную характеристику. Радиосистемы с непрерывными (или квазинепрерывными) сигналами работают в установившемся режиме, а анализ переходных процессов необходим для того, чтобы проектируемые устройства корректно инициализировались при включении и успокаивались при выключении (корректно отрабатывали амплитудные перегрузки по току и/или напряжению). В радиосистемах с импульсными сигналами анализ переходных процессов необходим по самому существу работы системы.

Оказывается, что h и s в интеграле свертки можно менять местами

$$\int_{-\infty}^{\infty} h(\tau)s(t-\tau)d\tau=\int_{-\infty}^{\infty} s(\tau)h(t-\tau)d\tau,$$

потому что не важно кто на кого набегает: сигнал на импульсную характеристику или она на сигнал, Рисунок 2. Здесь полезна общая запись интеграла, когда пределы интегрирования бесконечные: после доказательства данного свойства их можно ограничить в соответствии с текущим раскладом времени. Вообще, интеграл свертки легко запоминается правилом «трех тау»: тау, тэ минус тау, дэ тау.

Согласованный фильтр

Если линейный фильтр настроить на сигнал длительностью T, т. е. принять, что импульсная характеристика является зеркальной копией сигнала $h(t)=s(T-t)$, то фильтрация даст следующий результат

$$v(t)=\int_{-\infty}^{\infty} h(\tau)s(t-\tau)d\tau=\int_{-\infty}^{\infty} s(T-\tau)s(t-\tau)d\tau=\int_{0}^{T} s(T-\tau)s(t-\tau)d\tau.$$

Здесь пределы интегрирования сведены к отрезку, соответствующему импульсу $s(t)$ длительностью T. В каждом конкретном случае пределы следует сужать до своего минимума. В результате такой настройки фильтра на сигнал, отклик получился пропорциональным **автокорреляционной функции** (АКФ) сигнала $s(t)$

$$R_{ss}(\tau) = \int_{-\infty}^{\infty} s(t)s(t + |\tau|)dt.$$

Иногда перед интегралом АКФ ставят нормировочный множитель $1/T$, не влияющий на ее форму. Модуль $|\tau|$ подчеркивает четность АКФ (попробуйте графически понять почему АКФ всегда четная, даже без модуля).

В момент времени $t_3 = T$, Рисунок 4, происходит фазировка внутри фильтра (когерентное суммирование) и формируется максимум, пропорциональный энергии импульса

$$v(T) = \int_{0}^{T} s^2(\tau)d\tau = E_s.$$

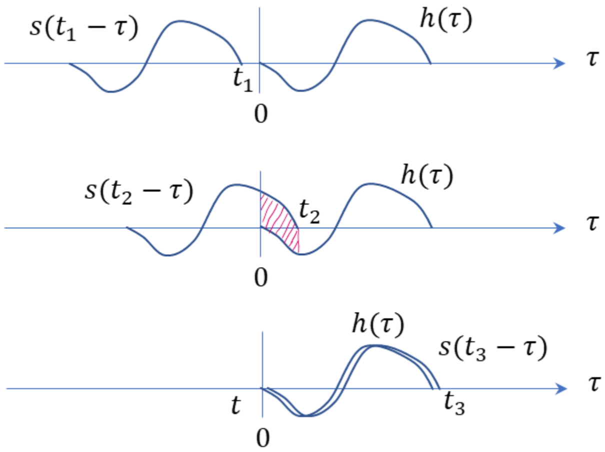


Рисунок 4 К формированию отклика согласованного фильтра на собственный сигнал. Иллюстрация фазировки – когерентного суммирования

Отсчет $v(T)$ и является сутью согласованной фильтрации или корреляционного приема, и отражает результат накопления информации о наличии полезного импульса на входе фильтра. Также разумно утверждать, что принятый импульс после обработки сворачивается в точку – число.

Линейный фильтр с импульсной характеристикой

$$h(t) = s(T - t),$$

называется **согласованным фильтром.** Согласование понимается в плане согласованности с импульсом $s(t)$ длительностью T (в практических ситуациях интервал наблюдения T конечен). Перед сигналом теоретически можно ставить ненулевой масштабирующий множитель; практически же этот множитель ограничен динамическим диапазоном обрабатывающего звена (разрядностью при цифровой реализации, т. е. количеством битов на слово).

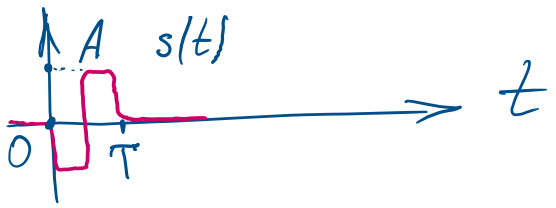
<u>Упражнение 1</u>

А) Отыскать отклик согласованного фильтра на прямоугольный импульс амплитудой A и длительностью T.

Б) Отыскать отклик согласованного фильтра на импульс типа «Манчестер» амплитудой A и длительностью T (смотри рисунок ниже).

В) Определить отклики на противоположные импульсы, т. е. на $-s(t)$.

Г) Сделать выводы относительно возможности приема двоичного сигнала $\{-s(t),\ s(t)\}$ согласованным фильтром (построить функциональную схему приемника).



Упражнение полезно запрограммировать в среде *Mathcad*, пользуясь встроенным определенным интегралом.

При наличии шума (помехи) в принимаемой (наблюдаемой) реализации $v(t) = s(t) + n(t)$, вычисляемый отсчет $v(T)$ будет содержать две составляющие – сигнальную и шумовую

$$v(T) = \int_0^T s^2(\tau)d\tau + \int_0^T s(\tau)n(\tau)d\tau = E_s + E_{sn}.$$

Компонента E_{sn} (ее называют кросс-компонентой или взаимной корреляцией) является **скалярным произведением** сигнала и шума.

Сигнальная компонента E_s известна заранее, а фактор, препятствующий извлечению полезной информации, заложен в шумовой компоненте – в ее случайности, и качество приема будет зависеть не от формы импульса $s(t)$, а от отношения сигнальной компоненты к шумовой, только не самих величин, а их квадратов, чтобы убрать влияние знака шумовой компоненты. Дабы нивелировать случайную природу шума, квадрат шумовой компоненты усредняется, в результате чего определяется **отношение сигнал-шум после обработки**

$$q^2 \overset{\text{def}}{=} \frac{E_s^2}{\overline{E_{sn}^2}}.$$

Существует такое понятие как **белый шум.** Это стационарный шум с нулевым средним и игольчатой автокорреляционной функцией

$$R_{nn}(\tau) = \overline{n(t)n(t+|\tau|)} = \frac{N_0}{2}\delta(\tau),$$

где $\frac{N_0}{2}$ – спектральная плотность мощности шума[1], Вт/Гц. Собственный шум электронных устройств, ограничивающий производительность радиосистем, хорошо аппроксимируется моделью белого шума в рабочей полосе частот. Достоинством этой модели является простота вычисления интегралов ввиду фильтрующего свойства дельта-функции, при этом отношение сигнал-шум после согласованной фильтрации будет следующим

$$q^2 = \frac{2E_s}{N_0}.$$

Это следует из цепочки равенств

$$\begin{aligned}\overline{E_{sn}^2} &= \overline{\int_0^T s(\tau)n(\tau)d\tau \int_0^T s(\tau)n(\tau)d\tau} = \int_0^T \int_0^T s(\tau_1)s(\tau_2)\overline{n(\tau_1)n(\tau_2)}d\tau_1\, d\tau_2 \\ &= (N_0/2)\int_0^T \int_0^T s(\tau_1)s(\tau_2)\delta(\tau_2-\tau_1)d\tau_1\, d\tau_2 = (N_0/2)\int_0^T s^2(\tau)d\tau \\ &= E_s N_0/2,\end{aligned}$$

[1] $N_0/2$ – двухсторонняя, а N_0 – односторонняя спектральная плотность мощности

которые выводятся в курсе «Статистической радиотехники» (иногда их можно найти в курсах «Радиотехнические цепи и сигналы» и «Цифровая связь»). Из этого выражения следует, что шумовая компонента (амплитудное значение, RMS, или среднеквадратическое значение, СКЗ) зависит не только от плотности шума, но и от энергии принятого импульса, однако эта зависимость пропорциональна корню квадратному

$$E_{sn\,RMS} = \sqrt{\overline{E_{sn}^2}} = \sqrt{E_s N_0/2},$$

что связано с дельта-коррелированной природой белого шума: если полезный сигнал в согласованном фильтре складывается когерентно (как сумма квадратов s^2), то белый шум – некогерентно (как сумма произведений sn). Случайная природа шума, в среднем, не позволяет интегралу (сумме) расти максимально быстро. Несмотря на пропорциональность шумовой компоненты энергии импульса, отношение сигнал-шум, все-таки, растет с ростом энергии принятого импульса, и чем больше эта энергия, тем лучше качество приема. Энергия импульса зависит не только от его мощности, но и от длительности, поэтому в цифровых радиосистемах большую роль играет длительность импульса, с помощью которой можно выставить баланс между скоростью передачи информации и вероятностью ошибки при заданной средней мощности формируемого импульса. Конечно, дополнительно существуют разные виды модуляции и кодирования, которые позволяют более тонко отрегулировать качество передачи информации, однако, это выходит за рамки данного пособия.

Упражнение 2

А) Ввести в модель *Mathcad* (смотри Упражнение 1) белый гауссовский шум, пронаблюдать несколько реализаций и понять природу некогерентного суммирования. Проверить равенство $E_{sn\,RMS} = \sqrt{E_s N_0/2}$. Учесть, что $N_0 = \sigma^2/\Delta f$, $\Delta f = 1/(2\Delta t)$, где σ^2 – среднеквадратическое значение шума (его мощность), Δt – интервал дискретизации по времени (при моделировании выбирается произвольно, удобно брать равным единице), Δf – односторонняя полоса частот (полоса обработки, половинная частота дискретизации).

Б) Подобрать пороговый уровень, минимизирующий вероятность ошибки при приеме двоичных сигналов

Вариант 1: биполярные импульсы

$$S_{bip} = \{-s(t), \qquad s(t)\}.$$

Вариант 2: ортогональные импульсы

$$S_{ort} = \{0, \qquad s(t)\}.$$

В) Подобрать уровень двоичных сигналов S_{bip} и S_{ort} так, чтобы средняя энергия на один бит была одинаковой

$$E_b = p_0 E_0 + p_1 E_1,$$

где p_0 и p_1 – вероятности битов, E_i – энергия i-го импульса.

Г) Для нескольких отношений сигнал-шум q^2 измерить вероятность ошибки $P_{\text{ош}}$ при приеме

1. Биполярных
2. Ортогональных

двоичных сигналов. Соблюсти условие постоянства средней энергии на один бит E_b для двух вариантов.

Производная по вектору как мощный инструмент решения оптимизационных задач

Рассмотрим скалярную функцию векторного аргумента

$$y = f(\vec{x}),$$

где $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ – вектор-столбец (матрица размером $n \times 1$) вещественных чисел.

Примерами таких функций могут быть квадрат евклидовой нормы

$$\|\vec{x}\|^2 = \sum_i {x_i}^2,$$

скалярное произведение (это функция двух аргументов)

$$(\vec{x},\vec{y}) = \sum_i x_i \cdot y_i,$$

квадрат евклидова расстояния между двумя векторами

$$d^2(\vec{x},\vec{y}) = \|\vec{x}-\vec{y}\|^2 = \sum_i (x_i - y_i)^2.$$

Элегантной записью введенных функций является векторная запись

$$\|\vec{x}\|^2 = \vec{x}^T\vec{x} = (x_1 \quad \cdots \quad x_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

$$(\vec{x},\vec{y}) = \vec{x}^T\vec{y} = \vec{y}^T\vec{x} = (\vec{y},\vec{x}).$$

Производной функции $y = f(\vec{x})$ по вектору $\vec{x}$ называется вектор

$$\frac{dy}{d\vec{x}} = \begin{pmatrix} \dfrac{\partial y}{\partial x_1} \\ \vdots \\ \dfrac{\partial y}{\partial x_n} \end{pmatrix},$$

состоящий из всех частных производных функции нескольких аргументов $f(x_1, \dots, x_n)$. В качестве примера найдем производную от квадрата нормы

$$\|\vec{x}\|^2 = x_1^2 + x_2^2 + \cdots + x_n^2,$$

$$\frac{d\|\vec{x}\|^2}{d\vec{x}} = \begin{pmatrix} 2x_1 \\ \vdots \\ 2x_n \end{pmatrix} = 2\vec{x}.$$

Символически результат совпадает с результатом обычной производной. Однако, здесь надо быть осторожным и все делать вручную, потому что производная по вектору имеет свои особенности (из-за транспонирования векторов и некоммутативности умножения), а готовые таблицы найти трудно (необходимо углубляться в матричный анализ).

Упражнение 3

Проверить равенства

А)

$$\frac{d(\vec{x},\vec{y})}{d\vec{x}} = \vec{y},$$

$$\frac{d(\vec{x},\vec{y})}{d\vec{y}} = \vec{x}.$$

Б)

$$\frac{d\|\vec{x}-\vec{y}\|^2}{d\vec{x}} = 2(\vec{x}-\vec{y}),$$

$$\frac{d\|\vec{x}-\vec{y}\|^2}{d\vec{y}} = 2(\vec{y}-\vec{x}).$$

В) Подобрать примеры более сложных скалярных функций векторного аргумента (логарифм квадрата нормы и т. п.) и найти производные по вектору. Результат дифференцирования по возможности привести к векторной форме.

Распространены также векторные функции векторного аргумента

$$\vec{y} = f(\vec{x}).$$

Такие возникают после дифференцирования по вектору скалярных функций (смотри предыдущий материал), или как исходные данные некоторой задачи.

Часто распространены линейные формы вида[2]

$$\vec{y} = R\vec{x} = \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ \cdots & \cdots & \cdots \\ r_{n1} & \cdots & r_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

$$\vec{y}^T = \vec{x}^T R^T = (x_1 \quad \cdots \quad x_n) \begin{pmatrix} r_{11} & \cdots & r_{n1} \\ \cdots & \cdots & \cdots \\ r_{1n} & \cdots & r_{nn} \end{pmatrix},$$

позволяющие выразить квадрат нормы преобразованного вектора через исходный вектор

$$\|\vec{y}\|^2 = \vec{x}^T R^T R \vec{x}.$$

Таким способом записывается квадратичная форма. Если же

$$\vec{v} = V\vec{x},$$

$$\vec{w} = W\vec{x},$$

то скалярное произведение векторов запишется как квадратичная форма в обобщенном виде

$$(\vec{v},\vec{w}) = \vec{x}^T V^T W \vec{x}.$$

---

[2] Фактически, это модель одномерного интеграла свертки в виде конечной суммы

Отсюда, кстати, следует правило транспонирования

$$(V\vec{x})^T = \vec{x}^T V^T.$$

<u>Упражнение 4</u>

Выполните транспонирование

- $(\vec{x} + R\vec{y})^T$,
- $\left(\vec{a}^T B^T C \vec{d}\right)^T$.

Проверьте результаты на матрицах размером 2×2.

Производной вектор-функции $\vec{y} = f(\vec{x})$ по вектору $\vec{x}^T$ является матрица

$$\frac{\partial \vec{y}}{\partial \vec{x}^T} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_n} \end{pmatrix},$$

состоящая из всех частных производных всех функций. Здесь берется производная по вектору-строке от вектора-столбца, поэтому частные производные группируются соответствующим образом.

Производной вектор-функции $\vec{y}^T = f(\vec{x})$ по вектору $\vec{x}$ является матрица

$$\frac{\partial \vec{y}^T}{\partial \vec{x}} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_1} \\ \cdots & \cdots & \cdots \\ \frac{\partial y_1}{\partial x_n} & \cdots & \frac{\partial y_n}{\partial x_n} \end{pmatrix}.$$

Эта матрица является транспонированной предыдущей матрицей, поэтому верно равенство

$$\left(\frac{\partial \vec{y}}{\partial \vec{x}^T}\right)^T = \frac{\partial \vec{y}^T}{\partial \vec{x}}.$$

Заметим, что здесь рассматриваются вектор-функции той же размерности n, что и вектор-аргумент, и поэтому все матрицы квадратные. Для решения задачи поиска оптимального фильтра этого достаточно, а для доказательства метода наименьших квадратов – нет.

В качестве примера найдем производную линейных функций

$$\frac{d(R\vec{x})}{d\vec{x}^T} = \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ \cdots & \cdots & \cdots \\ r_{n1} & \cdots & r_{nn} \end{pmatrix} = R,$$

$$\frac{d(\vec{x}^T R)}{d\vec{x}} = \begin{pmatrix} r_{11} & \cdots & r_{n1} \\ \cdots & \cdots & \cdots \\ r_{1n} & \cdots & r_{nn} \end{pmatrix} = R^T.$$

<u>Упражнение 5</u>

Проверьте полученные выше результаты на матрицах размером 2 × 2.

<u>Упражнение 6</u>

Вычислите выражения

А)

$$\frac{d(\vec{x}^T)}{d\vec{x}} =?$$

Б)

$$\frac{d(\vec{x}^T R\vec{x})}{d\vec{x}} =?$$

Попытайтесь связать Б) с А).

Решение задачи поиска оптимального линейного фильтра

Распространена задача минимизации (максимизации) нормы или скалярного произведения. Одним из примеров решения такой задачи является метод наименьших квадратов. В качестве второго примера разумно привести задачу поиска линейного фильтра, извлекающего информацию о наличии полезного сигнала $s(t)$ в наблюдаемой реализации

$$v(t) = b \cdot s(t) + n(t), \qquad b = \{0, 1\},$$

с наибольшим выходным отношением сигнал-шум q^2. При рассмотрении помехи в виде белого шума найденный оптимальный фильтр будет называться согласованным, а в общем случае коррелированного (окрашенного) шума $n(t)$ – просто оптимальным линейным фильтром. Решением последней задачи с использованием производной по вектору и займемся.

Рассмотрим линейный фильтр с некоторой импульсной характеристикой $h(t)$. Подадим на вход фильтра аддитивную смесь полезного сигнала с шумом $v(t) = s(t) + n(t)$, определенную на отрезке $[0, T]$. В момент времени T, когда реализация $v(t)$ полностью зашла в фильтр (и он тем самым может извлечь всю информацию), определим отношение сигнал-шум

$$q^2 \overset{\text{def}}{=} \frac{v_{\text{с}}^2}{\overline{v_{\text{ш}}^2}} = \frac{\left(\int_0^T s(\tau)h(T-\tau)d\tau\right)^2}{\iint_0^T h(T-\tau_1)h(T-\tau_2)\overline{n(\tau_1)n(\tau_2)}d\tau_1 d\tau_2}$$

$$= \frac{\left(\int_0^T s(\tau)h(T-\tau)d\tau\right)^2}{\iint_0^T h(T-\tau_1)h(T-\tau_2)R_{nn}(\tau_2-\tau_1)d\tau_1 d\tau_2}.$$

Здесь $R_{nn}(\tau_2 - \tau_1)$ – автокорреляционная функция шума $n(t)$. Как видим, при линейной фильтрации аддитивной смеси оптимальность фильтра при фиксации полезного сигнала зависит только от корреляционных свойств шума (шум стационарный: среднее равно нулю, дисперсия – постоянная величина, автокорреляционная функция зависит только от разности двух моментов времени).

Стоит задача поиска импульсной характеристики $h(t)$ линейного фильтра, дающей наибольшее отношение сигнал-шум q^2 в момент накопления сигнала, при этом известны полезный сигнал, а также автокорреляционная функция шума.

Запишем определенное отношение сигнал-шум в векторно-матричном виде, используя конечные разности при представлении интегралов (здесь $d\tau$ в числителе и знаменателе сокращаются, поэтому такой переход корректен)

$$q^2 = \frac{\left(\vec{s}^T\vec{h}\right)^2}{\vec{h}^T R \vec{h}}.$$

Здесь

$\vec{s}$ – вектор полезного сигнала $s(t)$,

$\vec{h}$ – вектор зеркально-отраженной импульсной характеристики $h(T-t)$,

R – корреляционная матрица, образованная автокорреляционной функцией $R_{nn}(\tau_2 - \tau_1)$ шума $n(t)$. Следует учесть, что эта функция всегда четная, поэтому $R = R^T$. В справедливости полученного выражения следует убедиться самостоятельно. Также полезно представлять, где число (скаляр), где вектор, а где – матрица.

Таким образом, в явном виде записан функционал от $\vec{h}$, требующий максимизации. Для поиска максимума следует найти производную, приравнять ее к нулю и решить уравнение, тем самым найдя вектор импульсной характеристики, по которой можно восстановить непрерывную функцию, т. к. шаг дискретизации, который был выбран при аппроксимации интегралов, не ограничен снизу, а границы индексов отсутствуют и не влияют на результат (в этом и достоинство векторной записи).

Производная

$$\frac{dq^2}{d\vec{h}} = \frac{2\left(\vec{s}^T\vec{h}\right)\cdot\vec{s}\cdot\vec{h}^T R\vec{h} - \left(\vec{s}^T\vec{h}\right)^2\cdot 2R\vec{h}}{\left(\vec{h}^T R\vec{h}\right)^2}.$$

Приравнивая числитель к нулю и сокращая лишние скаляры, получим следующее уравнение

$$\vec{s}\cdot\vec{h}^T R\vec{h} = \vec{s}^T\vec{h}\cdot R\vec{h}.$$

Как найти $\vec{h}$? Ответ кроется в структуре уравнения

Вектор $n\times 1$ умножить на скаляр = скаляр умножить на вектор $n\times 1$

Перемещая левый скаляр $\vec{h}^T R\vec{h}$ левее, получим

$$\vec{h}^T R\vec{h}\cdot\vec{s} = \vec{s}^T\vec{h}\cdot R\vec{h}.$$

Приравняем векторы отдельно

$$\vec{s} \leftrightarrow R\vec{h},$$

и проверим, в связи с этим, равенство соответствующих скаляров

$$\vec{h}^T R\vec{h} = \vec{h}^T\vec{s},$$

$$\vec{s}^T\vec{h} = \left(R\vec{h}\right)^T\vec{h} = \vec{h}^T R^T\vec{h}.$$

Видим, что равенство выполняется (напомним про то, что $R = R^T$). Значит найденное равенство

$$\vec{s} = R\vec{h}$$

является решением поставленной задачи, пусть даже и в неявном виде. Для явного решения требуется обращение корреляционной матрицы

$$\vec{h} = R^{-1}\vec{s},$$

что зачастую содержит разного рода особенности; плюс переход к непрерывному виду в явном виде затруднителен. Однако, для неявного решения переход довольно-таки очевиден

$$s(t) = W \int_{-\infty}^{\infty} R_{nn}(t-\tau)h(T-\tau)d\tau.$$

Здесь W – нормировочный множитель, такой, что при конечно-разностном представлении $W\Delta t = 1$.

Для белого шума $R_{nn}(\tau) = \frac{N_0}{2}\delta(\tau)$, а интеграл вычисляется явно

$$s(t) = \frac{WN_0}{2}h(T-t).$$

Максимальное отношение сигнал-шум

$$q^2 = \vec{s}^T\vec{h} = \frac{1}{\Delta t}\vec{s}^T\vec{h}\,\Delta t = \frac{2}{WN_0\Delta t}\int_0^T s^2(t)dt = \frac{2E_s}{N_0}.$$

Видно, что сигнал-шум не зависит ни от формы сигнала, ни тем более от нормировочного множителя, а зависит лишь от энергии принятого сигнала и спектральной плотности мощности шума (проще говоря, от уровня шума).

Качество извлечения информации с помощью найденного фильтра зависит от отношения сигнал-шум, которое также называют отношением сигнал-шум после обработки, или ОСШ на выходе **оптимального приемника**, поэтому важна не форма сигнала, а его длительность, т. е. время накопления; также важна не мощность шума, которая зависит от рабочей полосы частот, а спектральная плотность, которая зависит от качества и условий работы приемника. Однако, встречающиеся в реальности искажения полезного сигнала (мультипликативные в отличие от аддитивного наложения шума) вносят зависимость от формы сигнала (разные формы распространяются в канале по-разному), и в итоге снижают наблюдаемое отношение сигнал-шум в оптимальном приемнике. Эталонное ОСШ $\frac{2E_s}{N_0}$ полезно тем, что оно ни при каких условиях не может быть превышено и является ориентиром. Также распространена практика пересчета действующей помехи, не обязанной быть белым шумом, к эквивалентному уровню белого шума.

Переходя в частотную область, можно показать, что частотная характеристика оптимального фильтра должна быть такой

$$H(f) = \frac{S^*(f)}{G_{nn}(f)} e^{-j2\pi fT},$$

где $S^*(f)$ – комплексно-сопряженный спектр полезного сигнала, $G_{nn}(f)$ – вещественная спектральная плотность мощности шума (преобразование Фурье от АКФ шума). Экспоненциальный множитель иллюстрирует задержку отклика фильтра на длительность импульса – фильтр накапливает сигнал, и только после накопления выдает результат. Таким образом, последняя формула говорит о том, что оптимальный фильтр можно организовать как каскадное соединение двух звеньев: фильтра-корректора, который выравнивает спектр принимаемой реализации (выбеливает шумовую часть), и непосредственно согласованного фильтра, который выделяет полезную часть.