Федеральное государственное бюджетное учреждение науки

Вычислительный центр им. А.А. Дородницына

Российской академии наук

На правах рукописи

Абрамов Вадим Игоревич

МЕТОД ОПОРНЫХ ОБЪЕКТОВ ДЛЯ ОБУЧЕНИЯ РАСПОЗНАВАНИЮ ОБРАЗОВ В ПРОИЗВОЛЬНЫХ МЕТРИЧЕСКИХ ПРОСТРАНСТВАХ

05.13.17 – Теоретические основы информатики

диссертация на соискание ученой степени

кандидата физико-математических наук

Научный руководитель

д.т.н., профессор Моттль Вадим Вячеславович

Москва, 2014

# Содержание

[Содержание 2](#_Toc401192151)

[Введение 5](#_Toc401192152)

[1 Реализация гипотезы компактности при построении методов обучения распознаванию образов. Основные задачи исследования 11](#_Toc401192153)

[1.1 Проблема восстановления скрытой зависимости по эмпирическим данным и гипотеза компактности 11](#_Toc401192154)

[1.2 Классический метод опорных векторов 12](#_Toc401192155)

[1.2.1 Концепция оптимальной разделяющей гиперплоскости в пространстве действительных признаков объектов 12](#_Toc401192156)

[1.2.2 Выпуклый критерий обучения и его двойственная формулировка 14](#_Toc401192157)

[1.2.3 Решающее правило распознавания. Подмножество опорных объектов. 18](#_Toc401192158)

[1.3 Беспризнаковое распознавание образов 20](#_Toc401192159)

[1.4 Погружение множества объектов реального мира с пред-евклидовой метрикой в евклидово линейное пространство 28](#_Toc401192160)

[1.5 Основные задачи исследования 31](#_Toc401192161)

[2 Погружение метрического пространства с произвольной метрикой в псевдоевклидово линейное пространство 32](#_Toc401192162)

[2.1 Построение псевдо-евклидова линейного пространства 32](#_Toc401192163)

[2.1.1 Общность пар элементов метрического пространства 32](#_Toc401192164)

[2.1.2 Индефинитное скалярное произведение 36](#_Toc401192165)

[2.1.3 Изометрический образ метрического пространства в псевдоевклидовом линейном пространстве 37](#_Toc401192166)

[2.1.4 Частный случай: Погружение метрического пространства с пред-евклидовой метрикой в евклидово линейное пространство 39](#_Toc401192167)

[2.2 Аффинные операции в псевдоевклидовом линейном пространстве 40](#_Toc401192168)

[2.2.1 Аффинная комбинация элементов псевдоевклидова пространства 40](#_Toc401192169)

[2.2.2 Аффинное псевдоевклидово пространство, натянутое на изометрический образ метрического пространства 43](#_Toc401192170)

[2.2.3 Частный случай пред-евклидовой метрики: Погружение метрического пространства объектов реального мира в непрерывное метрическое пространство с аффинными операциями 44](#_Toc401192171)

[3 Решающее правило различения объектов двух классов без выбора центрального элемента и критерий обучения по методу опорных векторов 46](#_Toc401192172)

[3.1 Диполь в псевдоевклидовом линейном пространстве 46](#_Toc401192173)

[3.1.1 Понятие диполя 46](#_Toc401192174)

[3.1.2 Параметрическое семейство дискриминантных функций в псевдоевклидовом линейном пространстве 49](#_Toc401192175)

[3.1.3 Частный случай пред-евклидовой метрики: Дискриминантная гиперплоскость в евклидовом линейном пространстве 52](#_Toc401192176)

[3.2 Метод опорных объектов для обучения распознаванию образов 53](#_Toc401192177)

[3.2.1 Невыпуклая задача обучения по методу опорных объектов: Максимизация зазора между объектами двух классов 53](#_Toc401192178)

[3.2.2 Двойственная форма задачи обучения 56](#_Toc401192179)

[3.2.3 Различие произвольной и евклидовой метрик 59](#_Toc401192180)

[3.3 Класс метрических дискриминантных решающих правил возрастающей сложности 60](#_Toc401192181)

[3.3.1 Преобразование исходной метрики 60](#_Toc401192182)

[3.3.2 Обучение во вложенных семействах дискриминантных решающих правил возрастающей сложности 63](#_Toc401192183)

[3.3.3 Частный случай исходной пред-евклидовой метрики 65](#_Toc401192184)

[4 Численная реализация двойственной задачи обучения распознаванию образов в множестве объектов с произвольной метрикой и результаты экспериментальных иследований 66](#_Toc401192185)

[4.1 Верификация личности по подписи для случая пред-евклидовой метрики 66](#_Toc401192186)

[4.2 Верификация личности по подписи для случая псевдо-евклидовой метрики 67](#_Toc401192187)

[5 Заключение 69](#_Toc401192188)

[Приложение: доказательство теорем 70](#_Toc401192189)

[5.1 Доказательство теоремы 1. 70](#_Toc401192190)

[5.2 Доказательство теоремы 2. 70](#_Toc401192191)

[5.3 Доказательство теоремы 3. 71](#_Toc401192192)

[5.4 Доказательство теоремы 4. 71](#_Toc401192193)

[5.5 Доказательство теоремы 5. 72](#_Toc401192194)

[5.6 Доказательство теоремы 6. 73](#_Toc401192195)

[5.7 Доказательство теоремы 7. 73](#_Toc401192196)

[5.8 Доказательство теоремы 8. 74](#_Toc401192197)

[5.9 Доказательство теоремы 9. 74](#_Toc401192198)

[5.10 Доказательство теоремы 10. 75](#_Toc401192199)

[5.11 Доказательство теоремы 11. 76](#_Toc401192200)

[Литература 79](#_Toc401192201)

# Введение

**Актуальность.** Проблема обучения распознаванию образов составляет важнейший аспект современной теории машинного обучения (*Machine Learning*) [[[1]](#endnote-1)]. Пусть в пределах генеральной совокупности объектов реального мира  всякий объект характеризуется скрытой от наблюдателя принадлежностью к одному из конечного множества классов . Природа выбирает некоторый объект и требует, чтобы наблюдатель «угадал» класс объекта, всякий раз «наказывая» за ошибку . В данной работе мы ограничимся рассмотрением задачи распознавания образов для двух классов .

Теория машинного обучения основана на понятии конечной обучающей совокупности, в которой известны истинные значения индексов классов объектов

  ()

и которая является единственным источником информации о скрытых свойствах природы, доступным наблюдателю. Задача обучения заключается в поиске такой аппроксимации  функции , известной лишь в пределах обучающей совокупности, которая позволила бы продолжить эту функцию на все множество объектов  .

В данной диссертации гипотеза компактности, сформулированная Э.М. Браверманом [[[2]](#endnote-2)], рассматривается как основополагающий принцип машинного обучения. В самой общей формулировке для задач обучения распознаванию образов гипотеза компактности заключается в предположении, что объекты, отнесенные природой к одному и тому же классу, имеют, как правило, бóльшее сходство между собой, чем объекты разных классов. Способ измерения сходства либо несходства объектов должен выработать наблюдатель.

В качестве математически корректной модели интуитивного суждения о несходстве объектов реального мира  в данной диссертации рассматривается некоторая метрика, выбранная наблюдателем:

 (2)

Выбор метрики удачен, если выполняется гипотеза компактности: когда , то в большинстве случаев классы объектов совпадают .

В диссертации показано, что именно гипотеза компактности, выражаемая некоторой метрикой (2), лежит в основе чрезвычайно популярного метода машинного обучения, получившего в мировой литературе название *кернельного метода* (*Kernel Based Approach*) [[[3]](#endnote-3)], являющегося развитием известного метода потенциальных функций [[[4]](#endnote-4)]. Под кернелом (потенциальной функцией) понимается всякая симметричная числовая функция, определенная на множестве объектов , образующая неотрицательно определенную матрицу значений  для всякой конечной совокупности объектов . Известно, что всякий кернел определяет погружение исходного множества объектов  в некоторое гипотетическое линейное пространство, быть может, бесконечномерное, на которое эта функция естественным образом продолжается и в котором она играет роль скалярного произведения, определяя тем самым евклидову метрику.

Наиболее популярный метод обучения распознаванию двух классов объектов, получивший название *метода опорных векторов* (*Support Vector Machine*) основан на идее поиска такой гиперплоскости в этом воображаемом линейном пространстве (дискриминантной гиперплоскости), которая обеспечила бы наиболее правильную классификацию объектов обучающей совокупности в смысле максимизации их евклидовых расстояний до гиперплоскости с нужной стороны от нее [3].

**Первая проблемная ситуация**, определившая выбор темы данного диссертационного исследования, заключается в том, что существует континуум разных кернелов, отличающихся друг от друга только выбором нулевого элемента в соответствующих гипотетических линейных пространствах, но определяющих одну и ту же метрику на множестве объектов реального мира. Кроме того, опыт показывает, что держатели прикладных задач, эксперты в соответствующих предметных областях знания, хорошо понимают предложение сформировать метрику на множестве объектов, удовлетворяющую, по их мнению, гипотезе компактности, но плохо понимают просьбу предложить соответствующий кернел.

**Для разрешения этой проблемной ситуации** в настоящей диссертации предлагается метод обучения распознаванию образов в множествах объектов, в которых определена лишь метрика вместо кернела, существенно более сложного по своей математической структуре. Метод практически полностью аналогичен методу опорных векторов и требует лишь, чтобы метрика на множестве объектов обладала некоторыми специальными свойствами. Метрики этого вида предложено называть *пред-евклидовыми* (*proto-Euclidean metrics*), поскольку каждая из них обеспечивает континуум естественных погружений исходного множества объектов в разные линейные пространства с разными скалярными произведениями, отличающиеся друг от друга только выбором нулевого элемента, но с одной и той же евклидовой метрикой. Предложенный метод поиска решающего правила распознавания по обучающей совокупности объектов приводит к задаче выпуклого квадратичного программирования. В отличие от классического метода опорных векторов, разработанный метод обучения назван *методом опорных объектов*, поскольку он не использует векторы признаков объектов в явном виде.

**Вторая проблемная ситуация** заключается в том, что существует широкий класс прикладных задач обучения распознаванию образов, в которых естественные метрики на множестве объектов принципиально не относятся к классу пред-евклидовых метрик. В частности, это задачи классификации биологических полимеров, компьютерного анализа речи, распознавания динамических подписей, целый рад задач распознавания изображений. В результате возникает необходимость создания методологии обучения распознаванию образов, опирающейся на понимание множества объектов реального мира как метрического пространства с произвольной метрикой.

**Для разрешения этой проблемной ситуации** в настоящей диссертации разработан способ погружения произвольного метрического пространства в линейное пространство с индефинитным скалярным произведением, называемое также *псевдоевклидовым пространством*. В отличие от обычного евклидова линейного пространства, в псевдоевклидовом линейном пространстве метрика определена не для всех пар элементов, и существенно искажается само понятие дискриминантной гиперплоскости. Разработанный метод обучения распознаванию образов в множествах объектов с произвольной метрикой, аналогичный классическому методу опорных векторов, основан на доказанном в диссертации факте, что для элементов метрического пространства, погружаемого в псевдоевклидово линейное пространство как подмножество изолированных точек, сохраняется корректное значение метрики.

**Третья проблемная ситуация** заключается в том, что, в отличие от частного случая пред-евклидовой метрики на множестве реальных объектов, метод обучения для произвольной метрики приводит к задаче квадратичного программирования с невыпуклой квадратичной формой в качестве целевой функции, матрица которой содержит как положительные, таки и отрицательные собственные числа. В результате квадрат нормы направляющего вектора искомой дискриминантной гиперплоскости в псевдоевклидовом пространстве может оказаться отрицательным, что разрушает саму идею разделения образов объектов двух классов. Такое явление характерно для обучающих совокупностей, для которых используемая метрика не удовлетворяет гипотезе компактности, что аналогично линейной неразделимости образов объектов разных классов в псевдоевклидовом пространстве.

**Для разрешения этой проблемной ситуации**, во-первых, в критерий обучения введена специальная штрафная функция, ограничивающая корректную область поиска разделяющей гиперплоскости. Во-вторых, разработано параметрическое семейство преобразований исходной метрики, аналогичное идее классической потенциальной функции [4], в котором нулевое значение параметра сохраняет исходную метрику, а его достаточно большое значение обеспечивает полное разделение классов в любой обучающей совокупности. Конечно, разделимость классов на этапе обучения не гарантирует малой ошибки распознавания на генеральной совокупности, поэтому выбор такого структурного параметра критерия обучения должен быть основан на соответствующей процедуре кросс-валидации.

**Целью диссертации** является разработка методологии обучения распознаванию объектов двух классов в множествах объектов, рассматриваемых как метрические пространства с произвольной метрикой.

**Методы исследования**. Теоретическое исследование базируется на общих принципах теории метрических пространств, аналитической геометрия и линейной алгебры, методе опорных векторов, методах выпуклой и невыпуклой оптимизации. Экспериментальное исследование проводилось с использованием программно-алгоритмического комплекса, разработанного автором.

**Научная новизна.** В данной работе впервые сформулирована задача обучения распознаванию образов в множестве объектов, рассматриваемом как метрическое пространство. Разработан математический аппарат погружение конечного метрического пространства с произвольной метрикой в псевдоевклидово линейное пространство. Предложено параметрическое семейство решающих правил различения точек двух классов в псевдоевклидовом линейном пространстве. Разработан критерий обучения распознаванию объектов двух классов в произвольном метрическом пространстве как обобщение классического метода опорных векторов. Разработаны численные методы и алгоритмы обучения распознаванию объектов двух классов в произвольном метрическом пространстве.

**Положения, выносимые на защиту:**

1. Математический аппарат погружения метрического пространства с произвольной метрикой в псевдоевклидово линейное пространство.
2. Параметрическое семейство решающих правил линейного различения точек двух классов в псевдоевклидовом линейном пространстве.
3. Критерий обучения распознаванию объектов двух классов в произвольном метрическом пространстве, являющийся обобщением классического критерия опорных векторов.
4. Численные методы и алгоритмы обучения распознаванию объектов двух классов в произвольном метрическом пространстве.

**Достоверность полученных результатов** подтверждается доказательствами сформулированных теорем и экспериментальной проверкой полученных результатов.

**Практическая значимость.** Разработанные алгоритмы позволяют решать широкий класс прикладных задач обучения распознаванию образов, для которых естественно понимание множества объектов как метрического пространства с произвольной метрикой. В частности, к этому классу относятся задачи классификации биологических полимеров, компьютерного анализа речи, распознавания динамических подписей, целый рад задач распознавания изображений.

**Связь с плановыми научными исследованиями.** Работа выполнена при поддержке грантов Российского фонда фундаментальных исследований №№ 11-07-00409-a, 14-07-00661-а, 13-07-13132 офи\_м\_РЖД.

**Апробация работы.** Основные положения и результаты диссертации докладывались на конференциях «Интеллектуализация обработки информации ИОИ - 2010» (Республика Кипр, г. Пафос, 2010 г.), «Интеллектуализация обработки информации ИОИ - 2012» (Черногория, г. Будва, 2012 г.), «Интеллектуализация обработки информации ИОИ - 2014» (Греция, о. Крит, 2014 г.), «Математические методы распознавания образов ММРО - 2013» (г. Казань, 2013 г.).

**Публикации.** По тематике работы опубликовано 4 статьи, в том числе 2 статьи в журналах, рекомендованных ВАК.

**Структура и объем работы.** Диссертация состоит из введения, 4 глав основного содержания, заключения и библиографии. Работа содержит страниц основного текста.

# Реализация гипотезы компактности при построении методов обучения распознаванию образов. Основные задачи исследования

## Проблема восстановления скрытой зависимости по эмпирическим данным и гипотеза компактности

Сущность проблемы восстановления скрытой зависимости по эмпирическим данным, составляющей важнейший аспект современной информатики, заключается в следующем [[[5]](#endnote-5)]. Пусть в пределах некоторой генеральной совокупности объектов реального мира  всякий объект  характеризуется значениями двух переменных – доступной наблюдателю  и скрытой . Природа «случайно» выбирает некоторый объект и требует, чтобы наблюдатель «угадал» значение скрытой характеристики по наблюдаемой, всякий раз «наказывая» его за ошибку . Такую задачу называют задачей оценивания числовой (обычно говорят – регрессионной) зависимости, если множество значений скрытой характеристики есть множество действительных чисел , и задачей распознавания образов, если скрытая характеристика принимает значения из конечного неупорядоченного множества . В данной работе рассматривается задача распознавания образов, причем только для двух классов объектов .

Предполагается, что единственным объективным источником информации о скрытых свойствах природы, доступным наблюдателю, является конечная обучающая совокупность

, , (3)

в которой известны истинные значения обеих характеристик объектов. Принятый наблюдателем метод выбора решающего правила , , применимого ко всякому объекту, в том числе не представленному в обучающей совокупности  (3), называется методом обучения.

## Классический метод опорных векторов

### Концепция оптимальной разделяющей гиперплоскости в пространстве действительных признаков объектов

Пожалуй, наиболее популярным в мировой литературе методом обучения распознаванию образов с двумя классами объектов является метод опорных векторов В.Н. Вапника и А.Я. Червоненкиса (SVM – Support Vector Machine) [3], в основе которого лежит ими же ранее предложенный метод обобщенного портрета [[[6]](#endnote-6)].

Пусть изучается некоторое множество объектов реального мира . Никакой реальный объект не может быть воспринят компьютером иначе, как через результат измерения того либо иного его свойства, выражаемый значением некоторой формальной переменной . Будем полагать, что у объектов реально мира измеряются  свойств, выражаемых действительными числами , . Совокупность результатов этих измерений будем называть вектором действительных признаков объекта .

С другой стороны, допустим, что все множество объектов  разбито на два класса индикаторной функцией , вообще говоря, неизвестной наблюдателю. Целью наблюдателя является определение скрытого класса предъявленного объекта , зная лишь доступный для непосредственного наблюдения вектор признаков . Иными словами, желание наблюдателя сводится к построению дискриминантной функции .

Очевидно, что всякая дискриминантная функция определяет некоторое разбиение реально существующих объектов на два класса , не совпадающее, вообще говоря, с объективно существующим разбиением . Естественно желание найти такую дискриминантную функцию, легко вычисляемую на компьютере, для которой это несовпадение было бы как можно меньше.

В качестве исходной информации для выбора дискриминантной функции будем рассматривать обучающую совокупность объектов , представленных и векторам их признаков , и фактическими значениями индикаторной функции класса . Таким образом, обучающая совокупность представлена конечным множеством пар .

Будем искать подходящую дискриминантную функцию в виде линейной действительной функции векторного аргумента:

 ()

где  – вектор действительных признаков объекта,  – индекс класса объекта.

Очевидно, что всякая такая функция определяет некоторую разделяющую гиперплоскость в пространстве , задаваемую направляющим вектором  и порогом .

Пусть поступила обучающая совокупность . Представляется естественной эвристическая идея выбрать такую разделяющую гиперплоскость , которая правильно разделяет объекты двух классов:



или, что эквивалентно,  для всех .

Допустим, что для предъявленной обучающей совокупности разделяющая гиперплоскость существует. Но в этом случае существует континуум разделяющих гиперплоскостей. Идея В.Н. Вапника заключается в выборе той из них, которая обеспечивает наибольший «зазор», по-английски, «margin» между гиперплоскостью и ближайшими точками обучающей совокупности как одного, так и другого класса . Вообще говоря, величина зазора  условна, и определяется еще и нормой направляющего вектора, поэтому задача формулируется в виде задачи условной оптимизации:

, , .

Такая концепция обучения названа концепцией оптимальной разделяющей гиперплоскости.

Задачу поиска оптимальной разделяющей гиперплоскости удобно записывать в несколько иной эквивалентной форме. Разделим обе части неравенства  на , тогда надо обеспечить выполнение условий  для  и . При такой замене переменных , поэтому требование  равносильно требованию . Отсюда следует эквивалентная формулировка задачи поиска оптимальной разделяющей гиперплоскости:

 ()

### Выпуклый критерий обучения и его двойственная формулировка

Однако предъявленная обучающая совокупность может оказаться линейно неразделимой, и задача (5) не будет иметь решения. В качестве еще одной эвристики В.Н. Вапник предложил в качестве компромисса «разрешить» некоторым точкам обучающей совокупности располагаться с «неправильной» стороны разделяющей гиперплоскости, потребовав, чтобы такой дефект был минимальным. Наибольшую популярность получил следующий критерий обучения:

 ()

где  – некоторый коэффициент, согласующий два, вообще говоря, взаимно противоречивых требования – обеспечить как можно меньшее значение нормы направляющего вектора и как можно меньшую ошибку классификации в пределах обучающей совокупности. Для выбора коэффициента  в исходном изложении метода нет никаких теоретических суждений, дается лишь эвристическая рекомендация выбирать его значение достаточно большим.

Именно критерий (6) обычно называют критерием обучения по принципу поиска оптимальной разделяющей гиперплоскости.

Задача (6) представляет собой задачу квадратичного программирования с квадратичной целевой функцией  переменных и  линейными ограничениями типа неравенств. Ее особенно удобно решать в двойственной форме в виде задачи квадратичного программирования с  переменными, столькими же ограничениями-неравенствами и одним ограничением-равенством.

Точка минимума критерия (6) удовлетворяет условиям теоремы Куна-Таккера [[[7]](#endnote-7)], которые удобно выразить как условия поиска седловой точки функции Лагранжа относительно множителей Лагранжа:

1.  при ограничениях-неравенствах



1.  при ограничениях-неравенствах



Исходная двойственная задача имеет вид:

 ()

При фиксированных значениях множителей Лагранжа условие минимума функции Лагранжа по направляющему вектору  дает

, ()

условие минимизации по смещению разделяющей гиперплоскости  приводит к равенству

, ()

а из условия минимума по переменным  вытекают равенства

, . ()

Подстановка равенств (8), (9) и (10) в исходную функцию Лагранжа (7) приводит к двойственной задаче обучения относительно только множителей Лагранжа при объектах обучающей совокупности:

 ()

Это задача квадратичного программирования, в которой число переменных равно числу объектов в обучающее совокупности. Ее решение  полностью определяет направляющий вектор  и смещение  оптимальной разделяющей гиперплоскости, получаемые как решение исходной задачи обучения (6).

Значение направляющего вектора оптимальной разделяющей гиперплоскости непосредственно определяется равенством (8), из которого следует, что направляющий вектор есть линейная комбинация векторов признаков объектов, представленных в обучающей совокупности. Очевидно, что векторы признаков только тех объектов участвуют в формировании оптимального направляющего вектора, множители Лагранжа при которых оказались ненулевыми :

. ()

Эти объекты принято называть опорными, поскольку на векторы их признаков как бы «опирается» направляющий вектор оптимальной разделяющей гиперплоскости. Отсюда происходит и название этого метода обучения – метод опорных векторов или, в англоязычной терминологии, Support Vector Machine (SVM).

Значения множителей Лагранжа, получающихся как решение двойственной задачи (11), непосредственно указывают, какие именно объекты обучающей совокупности неправильно классифицируются оптимальной разделяющей гиперплоскостью. Из (10) следует, что если , то , ограничение  не является активным, т.е. , и данный объект классифицирован неправильно. Если же , то в силу (10) , и ограничение  активно, т.е. ошибки классификации данного объекта нет .

Остается найти смещение  оптимальной разделяющей гиперплоскости (6). Рассмотрим опорные объекты в составе обучающей совокупности , причем только те их них, которые правильно классифицируются найденной оптимальной разделяющей гиперплоскостью . Для каждого из этих объектов выполняется равенство . Умножим каждое из этих равенств на  и сложим правые и левые части:



Из сравнения левой и правой части непосредственно следует оптимальное значение , выраженное через направляющий вектор:

. ()

Результат обучения можно представить в виде выражений для направляющего вектора (8) и смещения (13), которые вместе полностью определяют оптимальную разделяющую гиперплоскость (4).

### Решающее правило распознавания. Подмножество опорных объектов.

Некоторые дополнительные особенности этого решения играют чрезвычайно важную роль в так называемой беспризнаковой методологии обучения.

Во-первых, если объекты реального мира из генеральной совокупности  могут быть представлены в компьютере векторами их действительных признаков , то получаемое решающе правило распознавания выражено в терминах не самих векторов признаков, а лишь их попарных скалярных произведений . Линейное дихотомическое решающее правило  (от английского слова *decision*) имеет вид дискриминантной гиперплоскости в пространстве признаков

 , ()

требуя вычисления скалярного произведения вектора признаков нового объекта  с векторами признаков объектов обучающей совокупности 

Во-вторых, как правило, большинство неотрицательных коэффициентов , определяющих направляющий вектор дискриминантной гиперплоскости  в (14), оказываются равными нулю, и сохранять в памяти достаточно лишь небольшое число векторов признаков остальных объектов обучающей совокупности , называемых опорными, которые дали название методу опорных векторов. Таким образом, итоговое решающее правило распознавания класса нового объекта оказывается существенно проще его исходного вида (14):

 , . ()

Это обстоятельство лежит в основе общего подхода к распознаванию образов, в котором предполагается, что объекты представлены лишь некоторой двухместной функцией их парного сравнения  вместо индивидуальных векторов признаков . Такой способ представления объектов особенно адекватен широкому классу приложений, в которых трудно выбрать числовые признаки отдельных объектов, но достаточно легко вычислить некоторую числовую характеристику отношения между любыми двумя объектами. Для того, чтобы процесс обучения, т.е. построения дискриминантной гиперплоскости (15), сохранил все преимущества исходного метода опорных векторов, традиционно принято считать, что функция  должна быть кернелом [3] (в исходной терминологии, потенциальной функцией [[[8]](#endnote-8)]), т.е. быть симметричной  и образовывать неотрицательно определенные матрицы

, , , ()

для любой конечной совокупности объектов, в частности, удовлетворять условию  при . Всякий кернел погружает множество объектов реального мира  в большее гильбертово линейное пространство , в котором играет роль скалярного произведения. Решающее правило распознавания имеет в этом случае вид, аналогичный (15), с тем лишь отличием, что кернел  используется вместо скалярного произведения векторов признаков :

 , . ()

Здесь уже нет векторов признаков (17) уместно называть методом опорных объектов.

## Беспризнаковое распознавание образов

Требование неотрицательной определенности для функции парного сравнения объектов и соответствующих матриц (16) оказывается слишком обременительным для многих прикладных задач анализа данных. Альтернативный подход был предложен Р. Дьюиным и его коллегами [[[9]](#endnote-9),[[10]](#endnote-10)] под названием *реляционного дискриминантного анализа* (*Relational Discriminant Analysis*) и независимо в работах [[[11]](#endnote-11),[[12]](#endnote-12)] (*беспризнаковое распознавание образов*).

Идея заключается в том, чтобы интерпретировать значения произвольной функции парного сравнения между всяким объектом  и всеми объектами обучающей совокупности  как вектор вторичных признаков этого объекта , и применить затем обычный метод опорных векторов в  (14).

Обозначим  множество всех рассматриваемых объектов  и классифицированных на два подмножества  и , ,  некоторой неизвестной индикаторной функцией . Однако, в отличие от классической постановки задачи распознавания образов, не предполагается возможность измерения на объектах каких бы то ни было наблюдаемых признаков , которые позволили бы применять методы обучения, разработанные для векторных признаковых пространств.

Вместо этого будем предполагать, что для любых двух объектов  и  может быть измерена числовая характеристика их сходства 

В частности, химическая, так называемая первичная, структура полимерной молекулы белка, представляющей собой последовательность нескольких десятков или даже сотен остатков 20 существующих в природе аминокислот, полностью выражается цепочкой символов соответствующей длины над 20-буквенным алфавитом. Степень сходства двух аминокислотных последовательностей измеряют, вычисляя показатель парного элайнмента (выравнивания) , который имеет смысл логарифмического отношения правдоподобия двух гипотез – основной гипотезы, что обе последовательности произошли из одной и той же аминокислотной цепочки-прародителя в ходе естественной эволюции, против нулевой гипотезы, что обе последовательности представляют собой случайную комбинацию аминокислот.

По своей природе показатель парного выравнивания может принимать как положительные, так и отрицательные значения, причем его значение, вычисленное для аминокислотной последовательности с самой собой , оказывается положительным и разным для разных последовательностей. Кроме того, матрица значений парного выравнивания, вычисленных для некоторой совокупности аминокислотных цепочек , оказывается положительно определенной или почти положительно определенной.

В результате оказывается, что такой показатель сходства двух белков обладает свойствами, очень напоминающими свойства скалярного произведения элементов линейного пространства. Такая же особенность характерна для показателей парного сходства объектов во многих других приложениях. Именно такие задачи беспризнакового распознавания образов рассматриваются в настоящей диссертации.

Беспризнаковое распознавание образов основано на эвристической гипотезе, что множество всех потенциальных объектов распознавания можно рассматривать как подмножество линейного пространства  со скалярным произведением (гильбертово пространство), в котором линейные операции определены произвольным неизвестным способом, а роль скалярного произведения играет показатель парного сходства между объектами.

Будем полагать, что в таком расширенном пространстве объектов распознавания определены линейные операции сложения , умножения на действительнозначный коэффициент , , и операция скалярного произведения  в некотором произвольном смысле при обычных ограничениях:

1. операция сложения симметрична и ассоциативна , ;
2. существование специального элемента – нуля  такого что  для любого элемента ;
3. существование обратного элемента  для любого ;
4. умножение на действительнозначный элемент , , ассоциативно  и  для любого ;
5. сложение и умножение дистрибутивны , ;
6. скалярное произведение элементов симметрично  и линейно , ;
7. скалярное произведение элемента с собой обладает свойствами ,  если, и только если  и имеет норму .

Еще раз подчеркнем – совершенно не предполагается, что все элементы гильбертова пространства  реально существуют как объекты распознавания. Мы рассматриваем только реально существующие объекты, и обозначаем это подмножество как , а остальные элементы из  являются лишь продуктами нашего воображения. Именно такое расширение  до  позволяет говорить о «суммировании» реально существующих объектов и об их «умножении» на действительнозначный коэффициент.

Предполагается, что даже если элемент  гильбертова пространства действительно существует , он не может быть представлен иначе как через свое скалярное произведение  с каким либо другим, реально существующим элементом .

Идея погружения множества реально существующих объектов распознавания в гипотетическое линейное пространство со скалярным произведением основана на концепции порождения новых объектов из реально наблюдаемых с помощью подходящих алгебраических операций, выдвинутой в работах Ю.И. Журавлева [[[13]](#endnote-13)]. Для порождения гипотетического гильбертова пространства объектов используется частный вид алгебраических операций, а именно линейные операции. Такое сужение общей концепции Ю.И. Журавлева, позволяющее детально разработать алгоритмы обучения, базируется на интерпретации численной меры сходства реально существующих объектов как скалярного произведения. Это позволяет рассматривать множество реально существующих объектов как подмножество изолированных точек в гипотетическом гильбертовом пространстве.

Если  – некоторый элемент гильбертова пространства, в общем случае воображаемый, то действительнозначная линейная дискриминантная функция , где  – некоторая константа, может быть использована как решающее правило , позволяющая судить о скрытой принадлежности некоторого рассматриваемого объекта , к первому или второму классу, независимо от того существует этот объект в реальности, или нет:

 ()

Здесь элемент  играет роль направляющего элемента соответствующей разделяющей гиперплоскости в гильбертовом пространстве .

Однако, пока у нас нет конструктивного инструмента для выбора направляющего элемента , и, следовательно, построения решающего правила, поскольку любой элемент из  может быть представлен только через свои скалярные произведения с некоторыми другими фиксированными элементами, существующими на самом деле.

Пусть наблюдатель выбрал конечную совокупность реально существующих объектов , называемую базисной совокупностью. В общем случае не предполагается, что элементы базисной совокупности классифицированы, то есть она не рассматривается как обучающая выборка. Базисная совокупность будет играть роль конечного базиса в гильбертовом пространстве, который определяет в нем -мерное подпространство .

Мы ограничимся рассмотрением только тех разделяющих гиперплоскостей, направляющие элементы которых принадлежит множеству , т. е. могут быть выражены в виде линейных комбинаций

, . ()

Соответствующее параметрическое семейство разделяющих гиперплоскостей полностью определяется скалярным произведением их элементов с объектами составляющими базисную совокупность:  .

Отметим, что если , то  для всех . Это означает, что если выбраны направляющие векторы в соответствии с (19) мы ограничиваем наше рассмотрение только теми разделяющими гиперплоскостями, которые ортогональны подпространству, определяемому базисной совокупностью объектов.

Каждому элементу гильбертова пространства  ставится в соответствие целый набор его скалярных произведений, которые мы будем рассматривать как действительнозначный «вектор признаков»

, . ()

Не предполагается, что базис  является полным в гильбертовом пространстве , поэтому произвольный элемент  не может быть представлен линейной комбинацией элементов базисной совокупности, но вектор признаков  полностью определяет проекцию элемента  на это подпространство.

Как результат, все элементы гильбертова пространства, имеющие одни и те же скалярные произведения с базисными элементами , или, другими словами, ту же проекцию на базисное подпространство , линейным решающим правилом (18) с  в форме (19) будут отнесены к одному и тому же классу . Именно поэтому мы будем называть признаки (20) проекционными признаками гильбертова пространства элементов.

Мы пришли к параметрическому семейству решающих правил распознавания образов в гильбертовом пространстве, опирающихся на проекционные признаки объектов:

 , , .

По своей структуре это семейство решающих правил полностью соответствует семейству линейных решающих правил (5). Таким образом, идея проекционных признаков сводит, по крайней мере внешне, задачу беспризнакового распознавания образов в гильбертовом пространстве к классической задаче в обычном линейном пространстве действительнозначных признаков.

Действительно, пусть наблюдателю представлена классифицированная обучающая выборка из объектов , , которая, в общем случае, не совпадает с базисной совокупностью . У наблюдателя нет другого способа увидеть эти объекты, «почувствовать» их, иначе как через скалярные произведения их с объектами базисной совокупности, что как раз эквивалентно вычислению их проекционных признаков

.

Параметры разделяющей гиперплоскости  и  должны выбираться таким образом, чтобы были правильно классифицированы объекты обучающей выборки, и между подвыборками разных классов был положительный зазор :



Если обучающие подвыборки первого и второго классов линейно разделимы существует семейство гиперплоскостей, которые соответствуют этим условиям. Ясно, что зазор  останется положительным после умножения  на положительный коэффициент , , т.е. можно рассматривать направляющие элементы фиксированной нормы . Один из таких элементов, для которого  и выполняются условия задачи (5), будем называть оптимальной разделяющей гиперплоскостью в гильбертовом пространстве.

Поскольку направляющий элемент разделяющей гиперплоскости определен как конечномерный вектор параметров, то такая задача, рассмотренная в базисном подпространстве , полностью совпадает с классической постановкой задачи распознавания образов путем поиска оптимальной разделяющей гиперплоскости. Легко показать, что максимальный зазор обеспечивается выбором направляющего элемента  и порога  из условий

,  .

Однако такой подход оказывается несостоятельным в случае пересекающихся подвыборок классов в базисном подпространстве, и ограничения и, следовательно, становятся несовместными. Для разработки аналогичного критерия для такой обучающей выборки, как и у В.Н. Вапника, вводятся неотрицательные переменные , , и используется компромиссный критерий  с достаточно большим положительным коэффициентом , регулирующим влияние дополнительных переменных . Эта идея заключается в мысленном сдвиге в сторону «своего» класса  тех объектов, которые мешают безошибочному разделению подвыборок, оставляя нетронутыми  объекты, которые не мешают «протиснуть» разделяющую гиперплоскость между подвыборками двух классов.

Таким образом, мы придем к следующей формулировке общей задачи поиска оптимальной разделяющей гиперплоскости в гильбертовом пространстве, которая покрывает как разделимый, так и неразделимый случаи:



Однако, в отличие от классической постановки задачи, норма направляющего элемента искомой гиперплоскости может трактоваться, по крайней мере, двумя способами, либо как норма элемента гильбертова пространства , , либо как норма направляющего вектора в пространстве проекционных признаков , . Эти два варианта нормы приводят к алгоритмам обучения и распознавания, существенно различающимся по своей структуре.

Для случая проекционных признаков решающая функция имеет вид:

 , . (21)

С формальной точки зрения решающее правило (21) остается в классе правил опорных объектов, поскольку каждая сумма  выражает сравнение нового объекта  только с опорными объектами обучающей совокупности . Однако применение правила (21) требует вычисления его вторичных признаков относительно всех объектов обучающей совокупности , , которые все равно надо держать в памяти. Последнее обстоятельство разрушает основное преимущество метода опорных объектов (17) – возможность не запоминать обучающие объекты, не являющиеся опорными.

В работе [[[14]](#endnote-14)] показано, что решающее правило (21) эквивалентно правилу опорных объектов в том и только том случае, когда функция парного сравнения объектов  обладает свойствами кернела.

В работах [[[15]](#endnote-15),[[16]](#endnote-16)] было показано, что если на множестве объектов определен кернел (потенциальная функция), то существует континуум других кернелов, полностью эквивалентных ему в смысле решающего правила распознавания, причем с тем же множеством опорных объектов  для всякой обучающей совокупности, которые отличаются друг от друга только значениями коэффициентов .

## Погружение множества объектов реального мира с пред-евклидовой метрикой в евклидово линейное пространство

Все эти эквивалентные кернелы определяют одну и ту же метрику на множестве объектов , удовлетворяющую, в дополнение к обычному неравенству треугольника

, (22)

еще и требованию условной неотрицательной определенности матриц

, , , (23)

для всех конечных совокупностей объектов. Метрики такого вида уместно называть пред-евклидовыми метриками (proto-Euclidean metric)[[17]](#footnote-1) в отличие от стандартного понятия евклидовых метрик, определяемых в некотором линейном пространстве выбором скалярного произведения. Впрочем, это очень небольшое обобщение, поскольку, как показано в [15,16], всякая пред-евклидова метрика , определенная на заданном множестве  с произвольно выбранным «центральным» элементом , погружает его в некоторое линейное пространство  с нулевым элементом  и скалярным произведением

, (24)

определяющим в этом линейном пространстве евклидову метрику в обычном смысле. Последняя метрика является однозначным продолжением на  исходной метрики  в  относительно произвольного выбора нулевого элемента .

Таким образом, согласно [15,16], вместо кернела на исходном множестве объектов  достаточно определить пред-евклидову метрику (23). Тогда соответствующее обобщение метода опорных векторов, сформулированное в [15,16], приведет к решающему правилу распознавания класса нового объекта вида

 , . (25)

Метрический метод опорных объектов (25) является обобщением кернельного метода и существенно более удобен для практического применения, поскольку нет надобности, в дополнение к пред-евклидовой метрике (23), назначать в множестве объектов реального мира  еще и нулевой элемент , выбор которого безразличен. Однако ему присущ тот же недостаток, что и традиционному кернельному методу – если содержательный смысл решаемой задачи анализа данных, как правило, подсказывает естественный выбор метрики на множестве объектов заданной физической природы  (22), то обеспечить наличие у нее свойства пред-евклидовости (23) крайне непросто.

В частности, в задачах классификации биологических полимеров, аминокислотных цепей белков либо нуклеотидных цепей ДНК, общепринятым является способ оптимизационного выравнивания длин двух сравниваемых последовательностей (pair-wise alignment) [[[18]](#endnote-17),[[19]](#endnote-18)]. Аналогичный оптимизационный метод парного сравнения сигналов разной длины, например в задачах компьютерного анализа речи, получил в англоязычной литературе название dynamic time warping [[[20]](#endnote-19),[[21]](#endnote-20)]. Ещё одним примером задачи, в которой естественным образом порождается функция парного отношения между объектами анализа – это распознавание форм бинарных растровых изображений на основе сравнения их скелетных представлений [[[22]](#endnote-21),[[23]](#endnote-22)]. Все эти методы приводят к некоторым метрикам на соответствующем множестве последовательностей либо сигналов (22), но эти метрики принципиально не являются пред-евклидовыми (23). Не обладает свойством пред-евклидовости и метрика на множестве динамических подписей [[[24]](#endnote-23)], задача классификации которых рассматривалась в экспериментальной части нашей статьи [16], где по этой причине наивное применение метрического метода опорных объектов не во всех экспериментах привело к успеху.

Именно эта проблема является предметом систематического изучения в настоящей работе. Мы исходим из предположения, что на множестве объектов реального мира, в котором наблюдателю необходимо решать задачу обучения распознаванию образов, определена произвольная метрика (22), вообще говоря, не являющаяся пред-евклидовой (23). В этом случае любой выбор центрального элемента  приводит к двухместной функции (24), не являющейся кернелом, и метод опорных объектов, а вместе с ним и его метрическое обобщение (25), становятся неприменимыми в их исходном виде. Основная причина заключается в том, что исходное произвольное метрическое пространство удается вложить лишь в псевдоевклидово линейное пространство, в котором метрика определена не для всех пар элементов, и существенно искажается само понятие дискриминантной гиперплоскости.

На первый взгляд, это обстоятельство не позволяет понимать заданную метрику иначе, как один из видов произвольной функции парного сравнения объектов , и наблюдателю остается лишь использовать методологию реляционного дискриминантного анализа (21), разрушающую основное преимущество метода опорных объектов. Тем не менее, в данной статье используется существенная специфика метрики как специального вида функции парного сравнения, удовлетворяющей неравенству треугольника (22). Именно на основе максимальной эксплуатации неравенства треугольника в статье рассматриваются пути «спасения» преимуществ метода опорных объектов, основанные на том факте, что для элементов метрического пространства объектов реального мира, погружаемого в псевдоевклидово линейное пространство как подмножество изолированных точек, сохраняется корректное значение метрики.

## Основные задачи исследования

Для достижения поставленной цели в диссертации сформулированы и решены следующие задачи:

1. Разработка математического аппарата погружения метрического пространства с произвольной метрикой в псевдоевклидово линейное пространство.
2. Разработка параметрического семейства решающих правил различения точек двух классов в псевдоевклидовом линейном пространстве.
3. Разработка критерия обучения распознаванию объектов двух классов в произвольном метрическом пространстве как обобщение классического метода опорных векторов.
4. Разработка численных методов и алгоритмов обучения распознаванию объектов двух классов в произвольном метрическом пространстве.
5. Количественное экспериментальное исследование разработанных методов и алгоритмов.

# Погружение метрического пространства с произвольной метрикой в псевдоевклидово линейное пространство

## Построение псевдо-евклидова линейного пространства

### Общность пар элементов метрического пространства

Пусть  – некоторое множество (генеральную совокупность) объектов реального мира c заданной на нем метрикой  (22). Выберем некоторый элемент  в качестве «центра» метрического пространства и образуем двухместную функцию

. (26)

Будем называть эту функцию общностью пары элементов метрического пространства относительно его центра. Заметим, что общность пары элементов метрического пространства относительно его центра аналогична известному понятию подобности произведения Громова (Gromov product similarity) [[[25]](#endnote-24)].

Из определения общности (26) следует, что общность элемента метрического пространства  с самим собой равна квадрату его расстояния до центра . Следовательно, расстояние между любыми двумя элементами однозначно определяется их общностями друг с другом и с самим собой независимо от выбора центра:

. (27)

**Теорема** **1.** *Если  – общность двух элементов метрического пространства относительно центра , то их общность относительно другого центра  определяется формулой*

. (28)

Доказательство теоремы приведено в Приложении **5.1**.

Предположим для простоты, что множество  конечно , , причем порядок нумерации его элементов в дальнейшем не будет играть существенной роли. Составим симметрическую матрицу значений общности относительно выбранного центра . Эта матрица имеет лишь концептуальное значение в наших рассуждениях. Число строк и столбцов этой матрицы может быть сколь угодно велико, но нам не придется обращаться к полному множеству ее элементов ни в каких вычислительных процедурах.

В силу симметричности матрицы  все её собственные числа и собственные векторы действительны, причем собственные векторы попарно ортогональны. Без ограничения общности можно считать нормы собственных векторов равными единице. Разумеется, собственные числа и векторы этой матрицы зависят от выбора центра метрического пространства :

 (29)

Известно, что матрица  всегда может быть представлена в виде

.

Для произвольной метрики матрица , вообще говоря, не является положительно определенной, поэтому среди ее собственных чисел могут быть как положительные, так и отрицательные числа. Упорядочим собственные числа в порядке убывания, полагая, что , и . Условимся все числа  считать неотрицательными и введем обозначения

, . (30)

Пару целых чисел  принято называть сигнатурой матрицы , которая в нашем случае является матрицей общности заданной метрики [[[26]](#endnote-25), [[27]](#endnote-26)].

**Теорема** **2.** *Сигнатура* *матрицы*  *не зависит от выбора центра метрического пространства* *.*

Доказательство теоремы, полученное Е.О. Черноусовой, приведено в приложении **5.2**.

Значения собственных чисел , вообще говоря, зависят от выбора центра , но после расположения их в порядке убывания (30) число положительных  и отрицательных  чисел остается неизменным в силу теоремы 2, поэтому второй нижний индекс в обозначениях собственных чисел можно отбросить:

, . (31)

Это означает, что сигнатура является характеристикой заданной метрики, а не выбора центра метрического пространства.

В принятых обозначениях собственных чисел выполняются равенства



Здесь все числа  и  неотрицательны и допускают извлечение корней. Введем обозначения:

 (32)

В этих обозначениях получим



Если составить матрицу  из собственных векторов матрицы  как из столбцов , то здесь  – строки этой матрицы с коэффициентами  и . Введем в рассмотрение квадратную матрицу 

, (33)

которую будем называть единичной матрицей сигнатуры . В этих терминах матрица значений общности элементов метрического пространства запишется в виде

 (34)

где векторы  и  являются частями векторов .

Мы связали элементы произвольного конечного метрического пространства , в котором выбран центральный элемент , с -мерными векторами действительных признаков элементов , определяемыми  собственными векторами матрицы значений общности этого конечного множества и  собственными числами ,  (29), причем центральному элементу соответствует нулевой вектор .

### Индефинитное скалярное произведение

Для произвольных пар векторов  линейного пространства  общее выражение для элемента матрицы значений общности (34) определяет двухместную числовую функцию   согласно (33)

. (35)

Здесь нижний индекс сознательно опущен , а не  , поскольку в силу теоремы 2 сигнатура не зависит от выбора центрального элемента .

Тем самым мы погрузили конечное метрическое пространство  с центральным элементом  в -мерное линейное пространство , в котором согласно (34) определена двухместная числовая функция (35).

Будем говорить, что такое линейное пространство  натянуто на конечное метрическое пространство .

Рассмотрим свойства двухместной функции  (35).

1) Симметричность . (36)

2) Билинейность . (37)

Эти два свойства совпадают со свойствами скалярного произведения в линейном пространстве, которое должно обладать еще и третьим свойством:

3) Неотрицательность при совпадающих аргументах . (38)

Однако последним свойством двухместная функция  (35) не обладает, поскольку для некоторых векторов  может иметь место неравенство . Действительно, согласно (17) для произвольного вектора согласно (16)

,

откуда следует, что

)  для векторов , таких что . (39)

Двухместную числовую функцию  (35) принято называть индефинитным скалярным произведением с сигнатурой , а само линейное пространство с индефинитным скалярным произведением – псевдоевклидовым пространством [[[28]](#endnote-27)]. В частности, в случае сигнатуры , , получается обычное евклидово пространство с обычным скалярным произведением, для которого .

Как следствие, в псевдоевклидовом пространстве , натянутом на метрическое пространство , для всех пар векторов   определено значение

, (40)

которое может быть положительным , равным нулю  даже при , либо отрицательным . Метрика в обычном понимании этого термина

 (41)

определена только для тех пар векторов, для которых . Если же , то значение метрики оказывается мнимым как квадратный корень из отрицательного числа.

Таким образом, построенное нами псевдоевклидово пространство  не является, вообще говоря, метрическим, поскольку в нем не для всех пар точек  определено значение метрики , но значение квадрата  (40) определено для всех пар векторов, являясь отрицательным для некоторых из них.

### Изометрический образ метрического пространства в псевдоевклидовом линейном пространстве

В псевдоевклидовом линейном пространстве, натянутом на конечное метрическое пространство , индефинитное скалярное произведение любых двух векторов  и  отличается от обычного скалярного произведения  только единичной матрицей сигнатуры  (35). В то же время собственные числа и собственные векторы матрицы значений общности  для конечного метрического пространства с выбранным центральным элементом  (29) определяют координаты конечного множества векторов в линейном пространстве 

,

зависящих от выбора центрального элемента, в которые отображаются элементы самого метрического пространства и которые играют роль базиса в этом линейном пространстве. Условимся обозначать это конечное подмножество в  символом

 (42)

и называть образом метрического пространства в псевдоевклидовом линейном пространстве. Разумеется, образ метрического пространства  (42), существенно зависит от центра через собственные числа и собственные векторы матрицы общности (29).

Как говорилось выше (41), метрика в обычном смысле определена в псевдоевклидовом пространстве  только для тех пар векторов , для которых  . Однако векторы из конечного образа метрического пространства  (42) являются именно такими. Действительно,



т.е. согласно (34) и, далее, (27)

, . (43)

Обратим внимание на тот факт, что хотя линейное пространство , в которое мы погрузили исходное конечное метрическое пространство , не зависит от выбора центрального элемента , погружение (42) существенно зависит от , однако изометричность погружения опять же никак не зависит от этого выбора.

### Частный случай: Погружение метрического пространства с пред-евклидовой метрикой в евклидово линейное пространство

До сих пор мы предполагали, что на конечном множестве объектов реального мира  определена произвольная метрика, т.е. двухместная функция , удовлетворяющая лишь неравенству треугольника (22). Предположим теперь, что исходная метрика удовлетворяет еще и требованию условной неотрицательной определенности матриц  для всякой конечной совокупности элементов (23), т.е. , . Такую метрику будем называть пред-евклидовой.

**Теорема** **3.** *В случае пред-евклидовой метрики матрица  значений общности элементов конечного метрического пространства    (26) неотрицательно определена для любого центра .*

Доказательство приведено в Приложении **5.3**.

В силу утверждения этой теоремы, все собственные числа матрицы  неотрицательны , и сигнатура этой матрицы равна ,  (32). Следовательно, согласно (32)-(34), матрица  может быть представлена как матрица обычных скалярных произведений



векторов , полученных как строки матрицы , составленной из собственных векторов матрицы  как из столбцов, дополнительно умноженные на коэффициенты  (32).

Итак, как и в общем случае, элементы произвольного конечного метрического пространства , в котором выбран центральный элемент , оказались связаны с -мерными векторами действительных признаков элементов , определяемыми  собственными векторами матрицы значений общности этого конечного множества и  собственными числами , причем центральному элементу соответствует нулевой вектор .

Отличие частного случая евклидовой метрики от общего случая произвольной метрики заключается в том, что теперь метрическое пространство погружено в обычное евклидово линейное пространство, а не в псевдоевклидово, как прежде, т.е. в линейное пространство с обычным скалярным произведением, обладающим всеми тремя свойствами (36), (37) и (38). Соответственно, метрика (40)-(41) определена для всех пар точек .

## Аффинные операции в псевдоевклидовом линейном пространстве

### Аффинная комбинация элементов псевдоевклидова пространства

Рассмотрим произвольную конечную неупорядоченную совокупность  элементов псевдоевклидова пространства  (34), т.е. совокупность пар векторов , , , , . Пусть  – вектор коэффициентов при элементах совокупности, в сумме составляющих единицу , где  – вектор, составленный из единиц. Линейная комбинация

, , (44)

называется аффинной комбинацией элементов  коэффициентами . Очевидно, что , где , .

**Теорема** **4.** *Квадрат расстояния любого элемента псевдоевклидова пространства  до аффинной комбинации  (44) согласно (40) определяется равенством*

. (45)

Доказательство теоремы приведено в Приложении **5.4**.

Если совокупность  составлена из произвольных элементов псевдоевклидова пространства, и элемент  также выбран произвольно, то среди квадратов расстояний  и  могут быть, вообще говоря, отрицательные, и квадрат расстояния  может также оказаться отрицательным.

Но если в качестве  и  выступают векторы, соответствующие объектам исходного метрического пространства , , ,

, , (46)

, (47)

то расстояния  и  являются вещественными,  и .

Тем не менее, метрические расстояния  существуют только в том случае, если  для любых . Однако для произвольной метрики на множестве  последнее условие, вообще говоря, не выполняется, и существуют совокупности элементов исходного метрического пространства , такие, что . Более того, для всякой такой совокупности может существовать множество аффинных комбинаций  (44).

Это означает, что вектор **, являющийся аффинной комбинацией векторов  с некоторыми коэффициентами  (44), может не иметь прообраза ни в каком расширении метрического пространства , т.е. может оказаться невозможным даже мысленно добавить соответствующий элемент  в , поскольку для него определены квадраты расстояний до всех других элементов , но не могут быть определены метрические расстояния , если .

В частности, если , то коэффициенты аффинной комбинации векторов  и  определяются одним действительным числом , например,  и . Тогда согласно (44)

. (48)

Такой вектор будем называть соосным упорядоченной паре векторов . Заметим, что согласно (45) для соосного вектора

, . (49)

Действительно, согласно (45)



Если , то определены метрические расстояния

, . (50)

Для любого вектора  квадрат расстояния до соосного вектора (48) согласно (45)

,

т.е.

. (51)

Очевидно, что квадрат расстояния соосного вектора до исходных векторов определяются равенствами

, . (52)

Вообще говоря, квадрат расстояния  может быть отрицательным, как и квадраты расстояний ,  и .

Но если в качестве  и  выступают векторы, соответствующие объектам исходного метрического пространства  и , , то расстояние  является вещественным, , тогда согласно (52)  и . Для векторов , соответствующих другим элементам метрического пространства ,

 (53)

Однако для произвольной метрики на множестве  существуют тройки элементов исходного метрического пространства , такие, что , и невозможно даже мысленно добавить соосный элемент  в , поскольку для него не определены метрические расстояния до других элементов .

### Аффинное псевдоевклидово пространство, натянутое на изометрический образ метрического пространства

Аффинная комбинация (44) определена для любой конечной совокупности  векторов псевдоевклидова пространства , в частности, для образа метрического пространства  (42), т.е. для совокупности всех  векторов, соответствующих элементам исходного метрического пространства  с выбранным центральным элементом  согласно (32)-(34). Всякая совокупность действительных коэффициентов , в сумме составляющих единицу , определяет вектор (46), зависящий от выбора центрального элемента, но находящийся от других векторов  на некоторых расстояниях (47), полностью определяемых коэффициентами . Хотя вектор аффинной комбинации  (46) и зависит от центра, квадрат его расстояния  до всякого вектора , быть может, отрицательный, от центра не зависит.

Множество всех векторов псевдоевклидова пространства (46), являющихся аффинными комбинациями элементов образа  исходного метрического пространства  (42), будем называть аффинным образом метрического пространства и обозначать символом

.

### Частный случай пред-евклидовой метрики: Погружение метрического пространства объектов реального мира в непрерывное метрическое пространство с аффинными операциями

Если исходная метрика является пред-евклидовой в смысле (23), то в силу теоремы 3 сигнатура исходного метрического пространства  равна ,  (32), и линейное пространство , в которое погружается , является обычным евклидовым линейным пространством вместо псевдоевклидового в общем случае.

Теперь всякой совокупности реальных объектов  с выбранным центральным элементом  и коэффициентов  соответствует аффинная комбинация  в обычном евклидовом линейном пространстве (46), для которой определены обычные метрические расстояния  до образов  всех реальных объектов  (47). Этот факт устанавливает следующая теорема.

**Теорема** **5.** *В случае пред-евклидовой метрики  (23) для всякой совокупности объектов ,* , *и коэффициентов  аффинная комбинация  (46)-(47) удовлетворяет неравенству*

.

*Вектор  является единственным в евклидовом линейном пространстве .*

Доказательство теоремы, основанное на непосредственном применении неравенства (23) к совокупности объектов  с коэффициентами , , приведено в приложении **5.5**.

Как следствие, для вектора **, являющегося аффинной комбинацией векторов  с коэффициентами , можно вообразить существование прообраза в некотором расширении метрического пространства , мысленно добавив в него соответствующий элемент , поскольку для него определены метрические расстояния до всех других элементов .

В нашей предыдущей статье [16] такой элемент , не зависящий от выбора центрального элемента , назван аффинной комбинацией элементов ** с коэффициентами , , и для него введено обозначение **.

Если в качестве совокупности объектов рассмотреть все конечное метрическое пространство  и мысленно добавить к нему аффинные комбинации ** со всеми коэффициентами , то в результате мы получим гипотетическое расширенное метрическое пространство , содержащее множество объектов .

В частности, в этом метрическом пространстве для всякой пары элементов  и всякого числа  существует единственный соосный элемент (50)

, . (54)

В отличие от , метрическое пространство  является непрерывным, т.е. содержит вместе с любым элементом  континуум элементов , расположенных к нему не далее сколь угодно малого порога . Но это еще и неограниченное выпуклое метрическое пространство в том смысле, что вместе с любыми двумя элементами , , оно содержит всю определяемую ими ось (54).

В общем случае произвольной метрики  такое погружение невозможно, поскольку не выполняются условия теоремы 5.

# Решающее правило различения объектов двух классов без выбора центрального элемента и критерий обучения по методу опорных векторов

## Диполь в псевдоевклидовом линейном пространстве

### Понятие диполя

Введенный формализм позволяет перейти к рассмотрению задачи обучения распознаванию образов на множестве объектов, представленных только через отношения произвольной метрики. Пусть по-прежнему  есть множество объектов реального мира с заданной на нем метрикой (22), и наблюдателю предоставлена конечная обучающая совокупность объектов вместе с известными индексами их принадлежности к одному из двух классов

 . (55)

Целью наблюдателя является построение решающего правила распознавания классов новых объектов , не представленных в обучающей совокупности, причем единственным свойством каждого нового объекта, доступным наблюдателю, является совокупность его расстояний до объектов обучающей выборки  .

Пусть  – псевдоевклидово линейное пространство, натянутое на конечное множество объектов реального мира . Это псевдоевклидово пространство однозначно «привязано» к метрическому пространству , поскольку в силу теоремы 2 его сигнатура фиксирована и определяется только метрикой . Вообще говоря, это пространство может иметь огромную размерность , но нам нигде далее не придется совершать в нем вычислительные операции, оно нам нужно лишь как математическое понятие для дальнейших построений.

Всякий выбор некоторого элемента  в качестве центрального ставит в соответствие каждому элементу  соответствующий ему вектор , т.е. определяет изометрический образ  (42) метрического пространства в , т.е.   (43).

Будем называть дискриминантным диполем упорядоченную пару векторов , а сами векторы  – узлами диполя. Рассмотрим множество всех векторов , соосных паре  в смысле (48) со всеми действительными коэффициентами :

. (56)

Вектор

, , (57)

будем называть центральной точкой диполя.

Условимся рассматривать только такие диполи, квадрат расстояния между узлами которых является положительным согласно (40):

. (58)

Для таких диполей определено метрическое расстояние между узлами .

Пусть  – произвольный вектор в псевдоевклидовом пространстве, например, вектор , соответствующий некоторому объекту реального мира  согласно идее линейного погружения. Тогда формула (51) определяет квадрат расстояния между  и :

. (59)

В силу предположения (58) , и эта функция является квадратичной и строго выпуклой функцией действительного коэффициента , но может принимать, вообще говоря, и отрицательные значения. В частности, отрицательным может быть ее минимальное значение . Тем не менее, пусть  – точка минимума, тогда вектор (56)

 (60)

естественно называть проекцией точки  на луч, образованный диполем .

Поскольку , то , т.е. векторы  (60) и  (57) характеризуются метрическим расстоянием до центра диполя

, (61)

полностью определяемым точкой  и диполем  в псевдоевклидовом пространстве .

Центральная идея методологии обучения распознаванию образов в произвольных метрических пространствах, предлагаемая в данной работе, заключается в использовании расстояния (61) с учетом знака  как параметрического семейства дискриминантных функций, каждая из которых задает некоторое разбиение псевдоевклидова пространства  на три части, определяемое выбором диполя :

 (62)

Следующая теорема придает этой дискриминантной функции конструктивный вид.

**Теорема** **6.** *Точка минимума  функции (59) определяется выражением*

, (63)

*причем значение  не зависит от расстояния между узлами диполя .*

Доказательство теоремы приведено в приложении **5.6**.

Будем называть нейтральное множество, определяемое согласно (62), дискриминантной гиперплоскостью в псевдоевклидовом пространстве :

. (64)

Теорема 6 позволяет записать дискриминантную функцию (62) в метрических терминах:

 (65)

Заметим, что в силу утверждения теоремы 6 здесь безразмерный дробный коэффициент перед длиной диполя  не зависит от самой длины. Именно этот безразмерный коэффициент определяет величину расстояния между проекцией точки  на ось диполя и центром диполя в псевдоевклидовом пространстве  с учетом знака (61)-(62), или, что то же самое, между точкой  и ее проекцией на дискриминантную гиперплоскость (64). Длина диполя является лишь масштабным коэффициентом этой зависимости, никак не влияя на разбиение псевдоевклидова пространства  на «положительную», «нейтральную» и «отрицательную» области .

В частности, при конкретном выборе центрального элемента  каждому реальному объекту  соответствует вектор , поэтому дискриминантная функция в псевдоевклидовом пространстве (65) фактически определяет дискриминантную функцию на множестве объектов реального мира:

 (66)

В следующем разделе мы покажем, что такая дискриминантная функция допускает запись в терминах исходной метрики на множестве , никак не зависящую от выбора в нем центрального элемента , определяющего общность элементов метрического пространства (26) и, далее, его погружение в псевдоевклидово линейное пространство (34).

### Параметрическое семейство дискриминантных функций в псевдоевклидовом линейном пространстве

Для заданной обучающей совокупности (55) задачу обучения естественно понимать как задачу выбора такого диполя , который определял бы разбиение множества обучающих объектов на два класса (65), как можно меньше отличающееся от разбиения, заданного «учителем».

Однако в псевдоевклидовом пространстве существует континуум разных диполей, определяющих одну и ту же дискриминантную функцию вида (65). В частности, достаточно ограничиться диполями фиксированной длины, например, единичной:

 (67)

Расстояние между узлами является не единственной излишней степенью свободы выбора диполя, выражающего желаемую дискриминантную гиперплоскость в , т.е. желаемое решающее правило по отношению к объектам реального мира , можно еще и «перемещать» диполь «параллельно» дискриминантной гиперплоскости. Покажем, что дискриминантную функцию  можно однозначно определить и без строгой фиксации узлов диполя.

Представляется естественным искать дискриминантную функцию, наилучшим образом разделяющую обучающую совокупность (55) в смысле  (67), выражая узлы дискриминантного диполя как неизвестные аффинные комбинации векторов , в которые отображаются объекты самой обучающей совокупности:

 (68)

**Теорема** **7.** *Для всякой точки  и диполя  (68) выполняется равенство*

. (69)

Доказательство теоремы, приведенное в приложении **5.7**, сводится к использованию утверждения теоремы 4.

В равенстве (69) только первая сумма в правой части зависит от предъявленного объекта , являясь линейной комбинацией квадратов его расстояний от объектов обучающей совокупности, причем в качестве коэффициентов выступают разности , сумма которых для любого диполя  должна равняться нулю согласно (68):

. (70)

Следующая теорема показывает, что длина (58) параметрически заданного диполя (68), которая должна быть фиксирована согласно (62), зависит только от коэффициентов .

**Теорема** **8.** *Расстояние между узлами диполя зависит только от исходных расстояний между объектами обучающей совокупности* ****** *(55) и коэффициентов* *******:*

. (71)

Доказательство теоремы приведено в приложении **5.8**.

Значения коэффициентов  (70) определяют ориентацию диполя в псевдоевклидовом пространстве относительно образов объектов обучающей совокупности , оставляя свободными как «параллельный перенос» диполя, так и его «сдвиг» вдоль своей оси. Именно этот «сдвиг» и характеризует вторая двойная сумма в правой части (69), которая является константой по отношению к предъявленному объекту . Обозначим ее символом

. (72)

Подставляя обозначения (70) (71) и (72) в (69) и далее в (67), мы получим эквивалентное выражение для дискримирантной функции, которая, как оказалось, полностью определяется  действительными числами :

 (73)

В частности, если в качестве вектора  выступает образ реального объекта  для некоторого центра , то дискриминантная функция полностью выражается через исходные расстояния этого объекта до объектов обучающей совокупности:

. (74)

В результате обучения должны быть найдены только числа  и  при двух ограничениях типа равенств (73).

В дальнейшем для нас существенное значение будут иметь два факта.

Во-первых, параметрическое семейство решающих правил классификации объекта  (74) определено непосредственно в исходном метрическом пространстве с произвольной метрикой и никак не зависит от выбора в нем центра .

Во-вторых, для произвольной метрики квадратичная форма , выражающая квадрат длины диполя (71), не является условно неотрицательно определенной (23), т.е. может принимать отрицательные значения даже при выполнении равенства .

### Частный случай пред-евклидовой метрики: Дискриминантная гиперплоскость в евклидовом линейном пространстве

В случае пред-евклидовой метрики, когда пространство  сигнатуры ,  (33), является обычным евклидовым линейным пространством с метрикой (40)-(41)

, ,

всякий диполь с несовпадающими узлами , , имеет положительный квадрат длины (58)

.

Тогда правило классификации точек евклидова линейного пространства (67) принимает вид обычной смещенной гиперплоскости с направляющим вектором единичной нормы:

.

После перехода в исходное метрическое пространство (74) специфика евклидовой метрики выражается в том, что в силу условной неотрицательной определенности (23) выполняется неравенство  для всех . Это, казалось бы, небольшое отличие от случая произвольной метрики приведет в разделе 3.2 к фундаментальному упрощению задачи обучения распознаванию образов.

## Метод опорных объектов для обучения распознаванию образов

### Невыпуклая задача обучения по методу опорных объектов: Максимизация зазора между объектами двух классов

Практически буквальным выражением принципа оптимальной дискриминантной гиперплоскости, лежащего в основе метода опорных векторов, является критерий обучения в евклидовом метрическом пространстве, требующий максимизации зазора (марджина в терминологии В.Н. Вапника)

, (75)

между объектами двух классов в обучающей совокупности , в нашем случае, согласно (66) и (69). Мы будем использовать здесь модификацию этого принципа, предложенную в кандидатской диссертации А.И. Татарчука [[[29]](#endnote-28)], заключающуюся в том, что всякое нарушение неравенства (75) «наказывается» функцией потерь

 (76)

в частности,  при , и  при . В качестве критерия обучения естественно искать такую гиперплоскость, которая разделяла бы обучающую выборку на два класса, с одной стороны, с как можно большей величиной зазора , а с другой, с как можно меньшей величиной суммарного штрафа для ошибочно классифицированных объектов обучающей выборки . Баланс таких требований к процессу обучения выражается оптимизационным критерием , где  ‑ структурный параметр, определяющий соотношение требований максимизации зазора и минимизации суммы потерь. Поскольку достаточно искать дискриминантную функцию, образуемую диполем единичной длины (74), то такой критерий обучения относительно искомых узлов диполя естественно записать как следующую задачу оптимизации с ограничениями в исходном метрическом пространстве:

 (77)

Задача (77) есть задача оптимизации в евклидовом линейном пространстве  независимо от того, является ли метрика в множестве объектов  произвольной или евклидовой. Размерность этого пространства определяется числом объектов в обучающей совокупности.

Обратим внимание на важный факт, вытекающий из самой идеи пороговой функции потерь, выражающей главную сущность метода опорных векторов. Очевидно, что задача обучения инвариантна к изменение масштаба приятой метрики , поскольку в этом случае надо пропорционально изменить также порог функции потерь  согласно (74), (75) и (76):

. (78)

Этот факт очевиден в исходной формулировке задачи обучения по методу опорных векторов (77), но потеряет очевидность в ее дальнейших эквивалентных преобразованиях.

В этой задаче переменные  подлежат варьированию на границе, вообще говоря, невыпуклой области . Мы применим идею, лежащую в основе SVM, и заключающуюся в замене исходной задачи на эквивалентную задачу оптимизации внутри этой области , которая, в отличие от классической задачи опорных векторов, является невыпуклой.

Разделим обе части ограничений-неравенств во второй строке (77) на 



и выполним замену переменных:

 (79)

С учетом этой замены ограничение-равенство в последней строке примет вид:

. (80)

Заметим, что требование максимизации зазора  в критерии (77) равносильно требованию минимизации этой квадратичной формы, которая в общем случае произвольной метрики не является условно неотрицательно определенной, и может принимать сколь угодно большие по модулю отрицательные значения.

Мы приходим к следующей задаче обучения, эквивалентной задаче (77):

 (81)

Здесь не учтены ограничения на коэффициенты при объектах обучающей совокупности  (77), однако, как мы увидим ниже, решение задачи (81) будет автоматически удовлетворять этим условиям.

Заметим, что это невыпуклая задача оптимизации, поскольку квадратичная форма в целевой функции и в квадратичном ограничении-неравенстве не является условно неотрицательно определенной при . Задача становится выпуклой, только если заранее известно, что метрика является евклидовой, как это предполагалось в наших предыдущих публикациях [15,16].

### Двойственная форма задачи обучения

Хотя задача обучения (81) получена из существенно более общего предположения произвольной метрики, формальная запись задачи ничем не отличается от ее прежней формулировки для евклидовой метрики [15,16]. Более того, с формальной точки зрения остается полностью справедливой и теорема о двойственной форме задачи обучения, сформулированная в [16].

Задача (81) является задачей минимизации квадратичной целевой функции с двумя совокупностями линейных ограничений типа неравенств, , , и , . В силу невыпуклости квадратичной формы, входящей в состав целевой функции, можно говорить лишь о поиске локального минимума с линейными ограничениями-неравенствами. Такая задача эквивалентна поиску седловой точки функции Лагранжа, аргументами которой, наряду с целевыми переменными , являются также неотрицательные множители Лагранжа, которые обозначим как  для первой группы ограничений и  для второй.

**Теорема** **9.** *Двойственная форма задачи обучения (81) имеет вид:*

 (82)

*Ее решение  полностью определяет параметры решающего правила распознавания (79):*

 (83)

*а также значение максимального зазора в (77)*

. (84)

Доказательство теоремы приведено в приложении **5.9**. Оно в основном повторяет доказательство в статье [16] за исключением того, что рассматриваются лишь необходимые условия минимума функции Лагранжа по переменным .

Из (83) и ограничения-равенства в двойственной задаче (82) с учетом переобозначения (79) вытекает равенство , как мы и обещали выше.

Однако в случае произвольной метрики решение «наивной» двойственной задачи (82) может привести к отрицательному значению квадрата максимального зазора  в (84). Требование максимизации зазора  в исходном критерии (77) равносильно требованию минимизации квадратичной формы (84), которая в общем случае не является условно неотрицательно определенной, и может принимать сколь угодно большие по модулю отрицательные значения. В то же время по своей сути зазор является действительным числом, т.е. , поэтому требование  должно быть дополнено в задаче обучения (81) ограничением . Если метрика является евклидовой, то это ограничение никогда не будет нарушаться и никак не исказит идею обучения.

В силу утверждения  теоремы 9, учет этого ограничения равносилен введению дополнительного ограничения  в двойственную задачу (82). Однако такое ограничение существенно усложнит решение двойственной задачи, являвшейся до сих пор стандартной задачей квадратичного программирования. Вместо строгого ограничения удобно использовать подходящую штрафную функцию, «почти точно» выражающую его требование.

Предлагается следующий вид штрафной функции (рис. 1):

, (85)

где  – достаточно большое число, значение которого само должно ассоциироваться с «очень большим» штрафом, например, .

|  |  |
| --- | --- |
| Рис. 1. Штрафная функция, не допускающая отрицательных значений квадратичной формы. | Действительно, такая функция принимает большое значение  при нулевом значении аргумента (квадратичной формы), быстро стремится к нулю при даже небольших положительных значениях, и уходит в бесконечность при отрицательных значениях аргумента:  (86) |

С учетом штрафной функции (85) предлагается вместо «наивной» двойственной задачи (82) решать задачу

 (87)

Решение двойственной задачи непосредственно определяет решающее правило классификации всякого объекта , в том числе и не участвующего в обучающей совокупности (74). При этом совсем не обязательно возвращаться к исходным значениям параметров решающего правила  и  согласно произведенной нами замене переменных (79), поскольку непосредственное использование значений  и  (83) приведет лишь к изменению масштаба дискриминантной функции без изменения ее знака. Не играет роли и коэффициент  перед ней:

. (88)

Гораздо более существенным является тот факт, что для классификации нового объекта  достаточно сравнить его по расстоянию  только с опорными объектами обучающей совокупности, для которых множители Лагранжа оказались положительными , остальные объекты можно не держать в памяти. Обучение можно считать удачным, если в результате решения двойственной задачи (87) удалось найти решающее правило (88), правильно классифицирующее объекты обучающей совокупности с достаточно большим значением зазора между классами  (84), т.е. с малым значением . Это значение будет гарантированно положительным, поскольку область отрицательных значений «отрезана» штрафной функцией. Как правило, в случае удачного обучения число опорных объектов оказывается небольшим по отношению к общему размеру обучающей совокупности. Отношение числа опорных объектов к общему числу обучающих объектов рассматривается даже как эвристическая оценка вероятности ошибки классификации на генеральной совокупности.

### Различие произвольной и евклидовой метрик

Такая двойственная задача остается правомерной и для евклидовой метрики, поскольку в этом случае значения штрафной функции будут всегда оставаться близкими к нулю согласно (86).

В случае евклидовой метрики исходная двойственная задача (82) является вогнутой задачей квадратичного программирования и легко решается стандартными вычислительными методами. В частности, многократно экспериментально проверена эффективность программных комплексов MOSEK [[[30]](#endnote-29)] и LIBSVM [[[31]](#endnote-30)]. Но эта задача совершенно не адекватна случаю произвольной метрики.

Двойственная задача (87) универсальна, но она не является квадратичной, поскольку функция штрафа не является таковой. Задача относится к общему классу задач выпуклого (вогнутого) программирования, если метрика заведомо является евклидовой, и легко численно решается [[[32]](#endnote-31)].

Но в случае неевклидовой метрики двойственная задача (87) теряет свойство вогнутостости, поскольку квадратичная форма  не является условно неотрицательно определенной. Однако присутствие штрафной функции, отсекающей область отрицательных значений этой квадратичной формы, существенно нивелирует невогнутость целевой функции, и, как показывает опыт, задача хорошо решается обычными методами выпуклого (вогнутого) программирования.

## Класс метрических дискриминантных решающих правил возрастающей сложности

### Преобразование исходной метрики

Итак, принятая наблюдателем метрика  на множестве объектов реального мира , объективно разбитом природой на два класса , определяет параметрическое семейство дискриминантных решающих правил  (74). Для заданной обучающей совокупности (55), результат обучения распознаванию двух классов по методу опорных объектов (87) с фиксированным значением параметра баланса  между требованиями максимизации зазора между классами и степенью его нарушения (77) характеризуется величиной максимально достижимого зазора  – чем больше, тем лучше.

Правда, успех обучения согласно (83), (79) и (74) характеризуется значением зазора  только в пределах обучающей совокупности, в то время как объективное качество полученного решающего правила можно проверить только на генеральной совокупности объектов , и там это качество всегда будет хуже. Мы не будем в данной работе изучать зависимость качества распознавания на всем множестве  от результата обучения по ограниченной выборке, которую принято называть обобщающей способностью метода обучения [5-6].

Нас будет интересовать лишь зависимость разделяющей способности класса решающих правил  (74), выражаемой максимально достижимым зазором  на фиксированной обучающей совокупности, от принятой метрики . Мы рассмотрим семейство преобразований исходной метрики, гарантированно повышающих ее разделяющую способность. Напомним еще раз, что это повышение не обязательно приведет к улучшению качества решающего правила на генеральной совокупности.

Рассмотрим семейство двухместных функций , определяемых преобразованием исходной метрики  (рис. 2):

, . (89)

**Теорема** **10.** *Если  есть метрика на , то  тоже метрика при любом .*

Доказательство теоремы приведено в приложении **5.10**.

|  |  |
| --- | --- |
| Рис. 2. Преобразование метрики. | Нетрудно убедиться, что  ,  .  Очевидно, что , причем  при  и , и что производная убывает, т.е. функция (89) является вогнутой. Кроме того,  , и  при . Таким образом, преобразование (89) является тождественным при . |

Поскольку (89) есть метрика, то все сказанное в предыдущих разделах остается справедливым по отношению к ней. Непосредственная подстановка (89) в решающее правило классификации объектов метрического пространства (74), в исходную «наивную» задачу обучения (77) и далее во все выражения вплоть до двойственной задачи (87) приводит к простой замене  на , причем в силу замечания (78) коэффициент не имеет значения, так что подставлять достаточно . Кроме того, учет условий  и, соответственно,  в (82) делает подстановку еще более простой – достаточно заменить  на :

 (90)

Вместо одного параметрического семейства решающих правил (74) и одной задачи обучения (81), мы получили класс параметрических семейств решающих правил и соответствующих им задач обучения, определяемых выбором еще одного дополнительного параметра . В следующем разделе мы покажем, что с ростом этого параметра увеличивается «свобода» выбора решающего правила классификации в процессе обучения. В силу последнего обстоятельства параметр  уместно назвать структурным параметром критерия обучения. Все сказанное выше есть частный случай при .

### Обучение во вложенных семействах дискриминантных решающих правил возрастающей сложности

Для анализа эффекта такого обобщения нам достаточно сравнить параметрическое семейство решающих правил распознавания произвольного объекта реального мира на основе класса метрик (89) с учетом (90)

, (91)

и исходное семейство решающих правил (74)

. (92)

Покажем, что при достаточно большом значении структурного параметра  существует решающее правило (91), правильно классифицирующее любую обучающую совокупность.

Рассмотрим произвольный объект реального мира  и найдем ближайший к нему объект в обучающей совокупности , т.е.  для всех :

.

Все слагаемые  умножим и разделим на :



Пусть  – минимальное расстояние между объектами в обучающей совокупности. В составе суммы  в каждом слагаемом



поэтому .

Рассмотрим вектор параметров решающего правила , такой что  при ,  при , и . Тогда

,

и, начиная с некоторого достаточно большого , знак  будет совпадать со знаком , т.е. объект  будет отнесен к тому же классу, что и ближайший к нему объект обучающей совокупности . Очевидно, что все объекты обучающей совокупности  будут правильно классифицированы решающим правилом  с параметрами , если , , если , и . Тем более, все объекты будут правильно классифицированы после того, как параметры решающего правила найдены как результат решения задачи обучения.

Покажем теперь, что для семейства решающих правил на основе исходной произвольной метрики при  (92) найдется обучающая совокупность, которая не может быть правильно классифицирована ни при каких значениях параметров .

Пусть обучающая совокупность состоит из четырех объектов :

. (93)

Попытаемся найти решающе правило вида (92), правильно классифицирующее все объекты:

 (94)

Тогда числа  удовлетворяют неравенствам

 или, что эквивалентно, 

Сложение левых и правых частей первых двух неравенств дает неравенство  , т.е. , а та же операция для вторых двух неравенств дает . Очевидно, что эти две пары неравенств несовместны, т.е. несовместны неравенства (94). Таким образом, не существует решающего правила вида (92), правильно классифицирующего объекты совокупности (93).

### Частный случай исходной пред-евклидовой метрики

Рассмотрим частный случай, когда исходная метрика  является пред-евклидовой (23). Что в этом случае можно утверждать о преобразованной метрике  (89)?

**Теорема** **11.** *Если исходная метрика является пред-евклидовой, то преобразованная метрика также пред-евклидова при любом значении параметра .*

Доказательство теоремы, приведенное в приложении **5.11**, опирается на следующую лемму, утверждение которой имеющую самостоятельное значение.

**Лемма 1.** *В случае пред-евклидовой метрики двухместная функция   является кернелом (потенциальной функцией) на , т.е. удовлетворяет условию (16).*

Доказательство леммы приведено в том же приложении **5.11**.

# Численная реализация двойственной задачи обучения распознаванию образов в множестве объектов с произвольной метрикой и результаты экспериментальных иследований

## Верификация личности по подписи для случая пред-евклидовой метрики

Задача верификации личности по подписи заключается в проверке нулевой гипотезы о том, что рассматриваемая подпись действительно принадлежит заявленному автору (genuine signature), против альтернативной гипотезы, что подпись является сознательной подделкой (skilled forgery).

Рассматриваются динамические подписи из базы данных SVC 2004 [[[33]](#endnote-32)], каждая из которых вводится в компьютер непосредственно в процессе написания (online), и представлена многокомпонентным дискретным сигналом  индивидуальной длины, отражающим ее геометрические и динамические особенности. Степень несходства подписей , играющая роль метрики, вычисляется на основе парного выравнивания соответствующих сигналов разной длины [[[34]](#endnote-33)]. Такая метрика, вообще говоря, не является евклидовой в том смысле, что условие (23) может не выполняться для некоторых совокупностей объектов. Проявление этого факта мы увидим в результатах эксперимента.

Массив содержит динамические подписи 40 лиц  по 20 для каждого из них, причем 10 подписей являются настоящими и еще 10 подделками. Таким образом, обучающая совокупность для каждой персоны  состоит из  подписей, представленных матрицей  попарных расстояний  и снабженных индексом класса  (настоящая) либо  (подделка).

Общий эксперимент заключался в обучении верификации истинности подписи каждого лица  согласно (82), (83) и далее 73, с последующей проверкой обобщающей способности по методу скользящего контроля.

В 3 из 40 частных экспериментов матрица попарных расстояний для совокупности подписей соответствующего лица, как истинных, так и поддельных, оказалась не обладающей свойством условной положительной определенности, что привело к невыпуклости критерия обучения (81) и расходимости процесса его оптимизации. В остальных 37 частных экспериментах ошибка скользящего контроля составила в среднем 3,65% и колебалась от 0% (21 случай) до 15% (2 случая).

Эксперимент показал необходимость обобщения метода на случай произвольной метрики.

## Верификация личности по подписи для случая псевдо-евклидовой метрики

Вообще говоря, в случае произвольной метрики задача (87) со штрафной функцией не является выпуклой. В то же время, опыт показывает, что эта невыпуклость не является «злостной», во всех экспериментах глобальный минимум критерия достигался в результате применения обычных итерационных методов, разработанных для решения задач выпуклого программирования. В данной работе использовался пакет программ, созданный Н.А. Разиным [[[35]](#endnote-34)] и основанный на стандартном методе внутренней точки.

Общая идея метода внутренней точки освещена в литературе [[[36]](#endnote-35)]. Исходная постановка задачи для этого метода такая:



где  – выпуклая функция, а  означает, что каждая компонента  неотрицательна и все  вогнуты.

Штрафы выражаются логарифмическими функциями, а решение задачи безусловной минимизации ищется при помощи метода Ньютона.

Применяемый метод сначала сводит ту задачу, которую он решает, к двойственной, а потом оперирует в терминах этой двойственной задачи. Поэтому мы фактически будем решать задачу, двойственную к двойственной исходной задаче. После этого выписываются условия Каруша-Куна-Такера для новой двойственной задачи и полученная система решается итеративно при помощи метода Ньютона. Поскольку ограничений будет , то и двойственных переменных будет , а активных переменных останется , потому что мы избавились от ограничения-равенства. Итого, переменных в двойственной задаче будет .

# Заключение

Основные результаты работы:

1. Разработан математического аппарата погружение метрического пространства с произвольной метрикой в псевдоевклидово линейное пространство.
2. Разработано параметрического семейства решающих правил различения точек двух классов в псевдоевклидовом линейном пространстве.
3. Разработан критерия обучения распознаванию объектов двух классов в произвольном метрическом пространстве как обобщение классического метода опорных векторов.
4. Разработаны численные методы и алгоритмы обучения распознаванию объектов двух классов в произвольном метрическом пространстве.
5. Проведено количественное экспериментальное исследование разработанных методов и алгоритмов на реальных данных.

# Приложение: доказательство теорем

## Доказательство теоремы 1.

Раскроем правую часть равенства (28) согласно определению общности (26):



Теорема доказана.

## Доказательство теоремы 2.

Пусть имеется совокупность объектов , объект  выбран как центр, и функция общности (26) образует матицу . Пусть теперь другой элемент назначен на роль центра, определяя другую матрицу общности  . Согласно теореме 1 справедливо равенство (28)

.

Введем обозначение   для матриц, в которых все элементы нули, кроме -й строки, состоящей из единиц, и элемента на пересечении -й строки и -го столбца, равного единице:



Нетрудно убедиться, что

.

Заметим, что матрицы  невырождены.

Рассмотрим квадратичные формы  и , . Здесь

.

Как видим, квадратичные формы совпадают при взаимно однозначной подстановке . В силу закона инерции квадратичных форм [25,26] числа положительных, нулевых и отрицательных и собственных чисел матриц  и  совпадают, т.е. совпадают их сигнатуры. Теорема доказана.

## Доказательство теоремы 3.

Пусть  – общность элементов для другого центра . Тогда по теореме 1:



Сумма неотрицательно определенных также неотрицательно определена.

Теорема доказана.

## Доказательство теоремы 4.

Пусть  – произвольный элемент псевдоевклидова пространства. Квадраты его расстояний до элементов совокупности  определяются сигнатурой псевдоевклидова пространства (40). Найдем квадрат расстояния этого элемента  до аффинной комбинации (44), учитывая, что :



Эта двойная сумма может быть представлена в виде



что с учетом (40) дает



откуда следует равенство (45). Теорема доказана.

## Доказательство теоремы 5.

Пусть  – произвольный элемент псевдоевклидова пространства. Квадраты его расстояний до элементов совокупности  определяются сигнатурой псевдоевклидова пространства (40). Найдем квадрат расстояния этого элемента  до аффинной комбинации (44), учитывая, что :



Эта двойная сумма может быть представлена в виде



что с учетом (40) дает



откуда следует равенство (45).

## Доказательство теоремы 6.

По формуле (53)



тогда

.

Следовательно, надо обеспечить равенство

,

т.е.

,

откуда следует



Теорема доказана.

## Доказательство теоремы 7.



Теорема доказана.

## Доказательство теоремы 8.



Здесь , поэтому

Теорема доказана.

## Доказательство теоремы 9.

Функция Лагранжа имеет стандартный вид:



Необходимые условия минимума по целевым переменным имеют обычный вид равенства нулю частных производных:

,

откуда следуют условия , ; (95)

, т.е. ; (96)

, (97)

что вместе с условиями  и  приводит к неравенствам , .

Подстановка полученных условий в функцию Лагранжа приводит к двойственной задаче (82). Равенства  и  (96) дают  в (83).

Если  в полученном решении двойственной задачи, то соответствующее ограничение в исходной задаче (81) неактивно, и , в противном случае  это ограничение активно, т.е. .

Если , то согласно равенству (97) , т.е. ограничение  в (81) активно, и . Напротив, если , то , соответствующее ограничение не активно, и .

Отсюда следует, что согласно (81) в точке решения двойственной задачи выполняются равенства  для всех . Умножение обеих частей этих равенств на  дает равенства , , а их суммирование приводит к равенству

,

из которого следует выражение для  в (83). Наконец, равенство (84) вытекает из (80) и (95). Теорема доказана.

## Доказательство теоремы 10.

Рассмотрим три объекта , , , . Пусть, без ограничения общности, . В силу вогнутости функции  (89) выполняются неравенства

,и ,

следовательно, справедливо неравенство . Исходная метрика по определению удовлетворяет неравенству треугольника (22), т.е. , и в силу неубывания функции получаем . Теорема доказана.

## Доказательство теоремы 11.

*Если исходная метрика является пред-евклидовой, то преобразованная метрика также пред-евклидова при любом значении параметра .*

**Лемма 1.** В случае пред-евклидовой метрики двухместная функция   является кернелом (потенциальной функцией) на , т.е. удовлетворяет условию (16).

**Лемма 2.** Сумма двух условных кернелов – условный кернел.

**Лемма 3.** Произведение двух условных кернелов – условный кернел.

Доказательство теоремы.

Нужно доказать, что  удовлетворяет требованию условной неотрицательной определенности матриц  для всякой конечной совокупности элементов (23), т.е. , . Именно такую метрику мы условились называть пред-евклидовой.

Итак,

.

Далее рассмотрим последовательно части выражения  и покажем, что они удовлетворяют требованиям условного кернела:

а) если  – пред-Евклидова метрика, то  – условный кернел по определению;

б) тогда  – кернел на основании утверждения Леммы 1;

в)  – условный кернел, как сумма кернела и условного кернела (), так как любая константа  – есть условный кернел;

г)  – условный кернел, как произведение пары условных кернелов.

Доказательство Леммы 1.

Доказать, что  – кернел при условии, что  – пред-евклидова метрика. Шаги доказательства:

а)  – есть произведение кернела (неотрицательная константа ) и условного кернела ;

б) разложение функции экспоненты в ряд: показывает, что достаточно опираться на два правила, что произведение условных кернелов и сумма условных кернелов есть условный кернел.

Доказательство Леммы 2.

Сумма двух условных кернелов – условный кернел. Действительно,  и , , тогда .

Доказательство Леммы 3.

Произведение двух условных кернелов – условный кернел.

Докажем более сильное утверждение: произведение двух кернелов – есть кернел. Действительно,  и ,

Представим каждую из матриц  и  как разложение следующего вида:

, , элемент  можно представить как  и

, , элемент  можно представить как ,

тогда 

# Литература

1. . Вапник В.Н. *Восстановление зависимостей по эмпирическим данным*. М., Наука, 1979, 449 с. [↑](#endnote-ref-1)
2. . Браверман Э.М. Опыты по обучению машины распознаванию зрительных образов. *Автоматика и телемеханика*, 1962, т. 23, № 3, с. 349-364. [↑](#endnote-ref-2)
3. . Vapnik V. *Statistical Learning Theory*. John-Wiley & Sons, Inc., 1998, 736 p. [↑](#endnote-ref-3)
4. . Айзерман М.А., Браверман Э.М., Розоноэр Л.И. *Метод потенциальных функций в теории обучения машин*. М.: Наука, 1970, 386 с. [↑](#endnote-ref-4)
5. . Вапник В.Н. *Восстановление зависимостей по эмпирическим данным*. М., Наука, 1979, 449 с. [↑](#endnote-ref-5)
6. . Вапник В.Н., Червоненкис А.Я. *Теория распознавания образов (статистические проблемы обучения)*. М.: Наука, 1974. [↑](#endnote-ref-6)
7. . Kuhn H.W., Tucker A.W. Nonlinear programming // Proceedings of 2nd Berkeley Symposium. Berkeley: University of California Press. – 1951. – pp. 481-492. [↑](#endnote-ref-7)
8. . Айзерман М.А., Браверман Э.М., Розоноэр Л.И. *Метод потенциальных функций в теории обучения машин*. М., Наука, 1970, 386 с. [↑](#endnote-ref-8)
9. . Duin R.P.W, De Ridder D., Tax D.M.J. Featureless classification. *Proceedings of the Workshop on Statistical Pattern Recognition*, Prague, June 1997, pp.37-42. [↑](#endnote-ref-9)
10. . Duin R., Pekalska E., De Ridder D. Relational discriminant analysis. *Pattern Recognition Letters*, Vol. 20, 1999, pp. 1175-1181. [↑](#endnote-ref-10)
11. . Mottl, V., Dvoenko, S., Seredin, O., Kulikowski, C., & Muchnik, I. (2001). Featureless pattern recognition in an imaginary Hilbert space and its application to protein fold classification. In Machine Learning and Data Mining in Pattern Recognition (pp. 322-336). Springer Berlin Heidelberg. [↑](#endnote-ref-11)
12. . Середин О.С. Методы и алгоритмы беспризнакового распознавания образов // Дисс. на соискание звания канд. наук. –М.:, 2001. [↑](#endnote-ref-12)
13. . Журавлев Ю.И. Избранные научные труды. – М.: Издательство Магистр, 1998, 420 с. [↑](#endnote-ref-13)
14. . Середин О.С. Потенциальная функция на множестве объектов распознавания как инструмент их попарного сравнительного представления *Известия ТулГУ, Естественные науки.* Тула: Изд-во ТулГУ, Вып. 1, 2013, С. 177-189. [↑](#endnote-ref-14)
15. . Абрамов В.И., Середин О.С., Сулимова В.В., Моттль В.В. Метод опорных объектов для обучения распознаванию образов в евклидовых метрических пространствах. Доклады 9-й международной конференции «Интеллектуализация обработки информации ИОИ-2012», Будва, Черногория, 16-22 сентября 2012 г. М.: Торус Пресс, 2012, с .5-8. [↑](#endnote-ref-15)
16. . Абрамов В.И., Середин О.С., Моттль В.В. Обучение распознаванию образов в евклидовых метрических пространствах по методу опорных объектов. *Известия ТулГУ, Естественные науки*. Тула: Изд-во ТулГУ, Вып. 2, 2013, с. 119-136. [↑](#endnote-ref-16)
17. Термин предложен К.В. Воронцовым. [↑](#footnote-ref-1)
18. . Needleman S.B., Wunsch, C.D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, Vol. 48, Issue 3, March 1970, pp. 443-453. [↑](#endnote-ref-17)
19. . Smith T.F., Waterman M.S. Identification of Common Molecular Subsequences. *Journal of Molecular Biology*, Vol. 147, 1981, pp. 195–197. [↑](#endnote-ref-18)
20. . Salvador S., Chan Ph. Toward accurate dynamic time warping in linear time and space. *KDD Workshop on Mining Temporal and Sequential Data*, 2004, pp. 70-80. [↑](#endnote-ref-19)
21. . Xiaoyue W., et al. Experimental comparison of representation methods and distance measures for time series data. *Data Mining and Knowledge Discovery*, March 2013, Volume 26, Issue 2, pp. 275-309. [↑](#endnote-ref-20)
22. . Местецкий Л.М. *Непрерывная морфология бинарных изображений. Фигуры, скелеты, циркуляры*. М.: Физматлит, 2009, 288 с. [↑](#endnote-ref-21)
23. . Kushnir O., Seredin O. Parametric Description of Skeleton Radial Function by Legendre Polynomials for Binary Images Comparison // A. Elmoataz et al. (Eds.): ICISP 2014, LNCS 8509, pp. 520–530. Springer International Publishing Switzerland (2014). [↑](#endnote-ref-22)
24. . Mottl V., Lange M., Sulimova V., Ermakov A. Signature verification based on fusion of on-line and off-line kernels. *Proceedings of the 19th International Conference on Pattern Recognition*. Tampa, USA, December 8-11, 2008, pp. 1-4. [↑](#endnote-ref-23)
25. . Деза М.М, Деза Е. *Энциклопедия расстояний*. М.: Наука, 2008.   
    M.M. Deza, E. Deza. *Dictionary of Distances*. Elsevier Science, 2006. [↑](#endnote-ref-24)
26. . Беклемишев Д.В. *Курс аналитической геометрии и линейной алгебры. Учебное пособие для вузов.* М., Высшая школа, 320 с. [↑](#endnote-ref-25)
27. . Умнов А.Е. Аналитическая геометрия и линейная алгебра: учебное пособие. М., МФТИ, 2011, 544 с. [↑](#endnote-ref-26)
28. . Азизов Т.Я., Копачевский Н.Д. *Введение в теорию пространств Крейна. Специальный курс лекций*. Симферополь, ООО Форма, 2010, 112 с. [↑](#endnote-ref-27)
29. . Татарчук А.И. Байесовские методы опорных векторов для обучения распознаванию образов с управляемой селективностью отбора признаков. //Дисс. на соискание звания канд. наук. –М.:, 2014. [↑](#endnote-ref-28)
30. . Mosek Solver Manual. <http://www.gams.com/dd/docs/solvers/mosek.pdf> [↑](#endnote-ref-29)
31. . C.-C. Chang and C.-J. Lin. LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011. [↑](#endnote-ref-30)
32. . Сухарев А.Г., Тимохов А.В., Федоров В.В. Курс методов оптимизации: Учебное пособие. М.: Физматлит, 2005, 386 с. [↑](#endnote-ref-31)
33. . SVC 2004. First International Signature Verification Competition. [Электронный ре-сурс] URL: http://www.cs.ust.hk/svc2004/index.html (дата обращения: 15.05.2013). [↑](#endnote-ref-32)
34. . Mottl V., Lange M., Sulimova V., Ermakov A. Signature verification based on fusion of on-line and off-line kernels. Proceedings of the 19th International Conference on Pattern Recognition. Tampa, USA, December 8-11, 2008, pp. 1-4. [↑](#endnote-ref-33)
35. . Разин Н.А. *Выпуклые критерии и параллелизуемые алгоритмы селективного комбинирования разнородных представлений объектов в задачах восстановления зависимостей по эмпирическим данным*. Дисс. к.ф.-м.н. ВЦ РАН, 2013. [↑](#endnote-ref-34)
36. . *A feasible BFGS interior point algorithm for solving convex minimization problems* / Paul Armand, Jean Charles Gilbert, Sophie Jan Jegou et al. // SIAM Journal on Optimization. — 2000. — V. 11. — P. 200. [↑](#endnote-ref-35)