Обобщенный линейный подход к восстановлению зависимостей по эмпирическим данным

Моттль Вадим Вячеславович
Вычислительный центр РАН

Середин Олег Сергеевич
Тульский государственный университет

Типовая задача восстановления зависимостей в множествах объектов реального мира

Некоторое множество реально существующих объектов $\omega \in \Omega$.

Типовая задача восстановления зависимостей в множествах объектов реального мира

Некоторое множество реально существующих объектов $\omega \in \Omega$.

Некоторое множество значений скрытой характеристики объектов $y \in \mathbb{Y}$.

Типовая задача восстановления зависимостей в множествах объектов реального мира

Некоторое множество реально существующих объектов $\omega \in \Omega$.

Некоторое множество значений скрытой характеристики объектов $y \in \mathbb{Y}$.

Объективно существующая скрытая функция $y(\omega): \Omega \to \mathbb{Y}$.

Типовая задача восстановления зависимостей в множествах объектов реального мира

Некоторое множество реально существующих объектов $\omega \in \Omega$.

Некоторое множество значений скрытой характеристики объектов $y \in \mathbb{Y}$.

Объективно существующая скрытая функция $y(\omega): \Omega \to \mathbb{Y}$.

Желание наблюдателя:

Иметь инструмент оценивания скрытой характеристики для реальных объектов

$\hat{y}(\omega): \Omega \to \mathbb{Y}$; $\quad \hat{y}(\omega) \neq y(\omega)$ – ошибка.

Типовая задача восстановления зависимостей в множествах объектов реального мира

Некоторое множество реально существующих объектов $\omega \in \Omega$.

Некоторое множество значений скрытой характеристики объектов $y \in \mathbb{Y}$.

Объективно существующая скрытая функция $y(\omega): \Omega \to \mathbb{Y}$.

Желание наблюдателя:

Иметь инструмент оценивания скрытой характеристики для реальных объектов

$\hat{y}(\omega): \Omega \to \mathbb{Y}$; $\hat{y}(\omega) \neq y(\omega)$ – ошибка.

Обобщенный линейный подход

Наблюдатель (его компьютер) воспринимает объекты реального мира как точки в некотором линейном пространстве: $\boldsymbol{x}(\omega): \Omega \to \mathbb{X}$ – линейное пространство.

Типовая задача восстановления зависимостей в множествах объектов реального мира

Некоторое множество реально существующих объектов $\omega \in \Omega$.

Некоторое множество значений скрытой характеристики объектов $y \in \mathbb{Y}$.

Объективно существующая скрытая функция $y(\omega): \Omega \to \mathbb{Y}$.

Желание наблюдателя:

Иметь инструмент оценивания скрытой характеристики для реальных объектов $\hat{y}(\omega): \Omega \to \mathbb{Y}$; $\hat{y}(\omega) \neq y(\omega)$ – ошибка.

Обобщенный линейный подход

Наблюдатель (его компьютер) воспринимает объекты реального мира как точки в некотором линейном пространстве: $\mathbf{x}(\omega): \Omega \to \mathbb{X}$ – линейное пространство.

- сложение коммутативно $\mathbf{x}' + \mathbf{x}'' = \mathbf{x}'' + \mathbf{x}'$
  и ассоциативно $(\mathbf{x}' + \mathbf{x}'') + \mathbf{x}''' = \mathbf{x}' + (\mathbf{x}'' + \mathbf{x}''')$;
- существует нулевой элемент $\mathbf{x} + \boldsymbol{\phi} = \mathbf{x}$, $c\boldsymbol{\phi} = \boldsymbol{\phi}$;
- для каждого элемента существует обратный ему элемент $(-\mathbf{x}) + \mathbf{x} = \boldsymbol{\phi}$;
- умножение ассоциативно $c'(c''\mathbf{x}) = (c'c'')\mathbf{x}$;
- умножение на единицу оставляет элемент без изменения $1\mathbf{x} = \mathbf{x}$;
- сложение и умножение дистрибутивны $(c' + c'')\mathbf{x} = c'\mathbf{x} + c''\mathbf{x}$, $c(\mathbf{x}' + \mathbf{x}'') = c\mathbf{x}' + c\mathbf{x}''$.

Типовая задача восстановления зависимостей в множествах объектов реального мира

Некоторое множество реально существующих объектов $\omega \in \Omega$.

Некоторое множество значений скрытой характеристики объектов $y \in \mathbb{Y}$.

Объективно существующая скрытая функция $y(\omega): \Omega \to \mathbb{Y}$.

Желание наблюдателя:

Иметь инструмент оценивания скрытой характеристики для реальных объектов

$\hat{y}(\omega): \Omega \to \mathbb{Y}$; $\hat{y}(\omega) \neq y(\omega)$ – ошибка.

Обобщенный линейный подход

Наблюдатель (его компьютер) воспринимает объекты реального мира как точки в некотором линейном пространстве: $\boldsymbol{x}(\omega): \Omega \to \mathbb{X}$ – линейное пространство.

J. (John) Nelder and R. (Robert) Wedderburn. Generalized Linear Models. *Journal of the Royal Statistical Society, Series A (General)*, 1972, Vol. 135, No. 3.

P. McCullagh and J. Nelder. *Generalized Linear Models*. Chapman & Hall, 1989, 532 p.

Линейное пространство с индефинитным скалярным произведением

$\omega \in \Omega$ – множество объектов реального мира,

$y \in \mathbb{Y}$ – множество значений целевой характеристики,

$\boldsymbol{x} \in \mathbb{X}$ – множество точек линейного пространства восприятия объектов,

Линейное пространство с индефинитным скалярным произведением

$\omega \in \Omega$ – множество объектов реального мира,
$y \in \mathbb{Y}$ – множество значений целевой характеристики,
$\boldsymbol{x} \in \mathbb{X}$ – множество точек линейного пространства восприятия объектов,
$\left\{ \boldsymbol{x}(\omega) \in \mathbb{X} : \omega \in \Omega \right\} \subset \mathbb{X}$ – подмножество образов конкретных объектов.

Линейное пространство с индефинитным скалярным произведением

$\omega \in \Omega$ – множество объектов реального мира,
$y \in \mathbb{Y}$ – множество значений целевой характеристики,
$\boldsymbol{x} \in \mathbb{X}$ – множество точек линейного пространства восприятия объектов,
$\left\{ \boldsymbol{x}(\omega) \in \mathbb{X} : \omega \in \Omega \right\} \subset \mathbb{X}$ – подмножество образов конкретных объектов.

$\boldsymbol{\phi} \in \mathbb{X}$ – нулевая точка линейного пространства,

Линейное пространство с индефинитным скалярным произведением

$\omega \in \Omega$ – множество объектов реального мира,
$y \in \mathbb{Y}$ – множество значений целевой характеристики,
$\boldsymbol{x} \in \mathbb{X}$ – множество точек линейного пространства восприятия объектов,
$\{\boldsymbol{x}(\omega) \in \mathbb{X} : \omega \in \Omega\} \subset \mathbb{X}$ – подмножество образов конкретных объектов.

$\boldsymbol{\phi} \in \mathbb{X}$ – нулевая точка линейного пространства,
$\boldsymbol{x}, \boldsymbol{\upsilon} \in \mathbb{X}$ – две произвольные точки в линейном пространстве,
$K(\boldsymbol{x}, \boldsymbol{\upsilon}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ – скалярная функция, обладающая следующими свойствами:
(1) $K(\boldsymbol{x}, \boldsymbol{\upsilon}) = K(\boldsymbol{\upsilon}, \boldsymbol{x})$ – симметричность,
(2) $K(\boldsymbol{x},\ c'\boldsymbol{\upsilon}' + c''\boldsymbol{\upsilon}'') = c'K(\boldsymbol{x}, \boldsymbol{\upsilon}') + c''K(\boldsymbol{x}, \boldsymbol{\upsilon}'')$ – билинейность.

Линейное пространство с индефинитным скалярным произведением

$\omega \in \Omega$ – множество объектов реального мира,
$y \in \mathbb{Y}$ – множество значений целевой характеристики,
$\boldsymbol{x} \in \mathbb{X}$ – множество точек линейного пространства восприятия объектов,
$\{\boldsymbol{x}(\omega) \in \mathbb{X} : \omega \in \Omega\} \subset \mathbb{X}$ – подмножество образов конкретных объектов.

$\boldsymbol{\phi} \in \mathbb{X}$ – нулевая точка линейного пространства,
$\boldsymbol{x}, \boldsymbol{v} \in \mathbb{X}$ – две произвольные точки в линейном пространстве,
$K(\boldsymbol{x}, \boldsymbol{v}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ – скалярная функция, обладающая следующими свойствами:
(1) $K(\boldsymbol{x}, \boldsymbol{v}) = K(\boldsymbol{v}, \boldsymbol{x})$ – симметричность,
(2) $K(\boldsymbol{x},\ c'\boldsymbol{v}' + c''\boldsymbol{v}'') = c'K(\boldsymbol{x}, \boldsymbol{v}') + c''K(\boldsymbol{x}, \boldsymbol{v}'')$ – билинейность.

Если еще добавить предположение
(3) $K(\boldsymbol{x}, \boldsymbol{x}) \geq 0$, то это будет обычное скалярное произведение, тогда
$\sqrt{K(\boldsymbol{x}, \boldsymbol{x})} = \|\boldsymbol{x}\|$ – норма.

Линейное пространство с индефинитным скалярным произведением

$\omega \in \Omega$ – множество объектов реального мира,
$y \in \mathbb{Y}$ – множество значений целевой характеристики,
$\boldsymbol{x} \in \mathbb{X}$ – множество точек линейного пространства восприятия объектов,
$\{\boldsymbol{x}(\omega) \in \mathbb{X} : \omega \in \Omega\} \subset \mathbb{X}$ – подмножество образов конкретных объектов.

$\boldsymbol{\phi} \in \mathbb{X}$ – нулевая точка линейного пространства,
$\boldsymbol{x}, \boldsymbol{\upsilon} \in \mathbb{X}$ – две произвольные точки в линейном пространстве,
$K(\boldsymbol{x}, \boldsymbol{\upsilon}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ – скалярная функция, обладающая следующими свойствами:
(1) $K(\boldsymbol{x}, \boldsymbol{\upsilon}) = K(\boldsymbol{\upsilon}, \boldsymbol{x})$ – симметричность,
(2) $K(\boldsymbol{x},\ c'\boldsymbol{\upsilon}' + c''\boldsymbol{\upsilon}'') = c'K(\boldsymbol{x}, \boldsymbol{\upsilon}') + c''K(\boldsymbol{x}, \boldsymbol{\upsilon}'')$ – билинейность.

Если еще добавить предположение
(3) $K(\boldsymbol{x}, \boldsymbol{x}) \geq 0$, то это будет обычное скалярное произведение, тогда
$\sqrt{K(\boldsymbol{x}, \boldsymbol{x})} = \|\boldsymbol{x}\|$ – норма.

Но нам будет достаточно свойств (1) и (2).
Это индефинитное скалярное произведение.

Линейное пространство с индефинитным скалярным произведением

$\omega \in \Omega$ – множество объектов реального мира,
$y \in \mathbb{Y}$ – множество значений целевой характеристики,
$\boldsymbol{x} \in \mathbb{X}$ – множество точек линейного пространства восприятия объектов,
$\{\boldsymbol{x}(\omega) \in \mathbb{X} : \omega \in \Omega\} \subset \mathbb{X}$ – подмножество образов конкретных объектов.

$\boldsymbol{\phi} \in \mathbb{X}$ – нулевая точка линейного пространства,
$\boldsymbol{x}, \boldsymbol{v} \in \mathbb{X}$ – две произвольные точки в линейном пространстве,
$K(\boldsymbol{x}, \boldsymbol{v}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ – скалярная функция, обладающая следующими свойствами:
(1) $K(\boldsymbol{x}, \boldsymbol{v}) = K(\boldsymbol{v}, \boldsymbol{x})$ – симметричность,
(2) $K(\boldsymbol{x},\ c'\boldsymbol{v}' + c''\boldsymbol{v}'') = c'K(\boldsymbol{x}, \boldsymbol{v}') + c''K(\boldsymbol{x}, \boldsymbol{v}'')$ – билинейность.

Если еще добавить предположение
(3) $K(\boldsymbol{x}, \boldsymbol{x}) \geq 0$, то это будет обычное скалярное произведение, тогда
$\sqrt{K(\boldsymbol{x}, \boldsymbol{x})} = \|\boldsymbol{x}\|$ – норма.

Но нам будет достаточно свойств (1) и (2).
Это индефинитное скалярное произведение.
Псевдоевклидово линейное пространство.

Линейное пространство с индефинитным скалярным произведением

$\omega \in \Omega$ – множество объектов реального мира,
$y \in \mathbb{Y}$ – множество значений целевой характеристики,
$\boldsymbol{x} \in \mathbb{X}$ – множество точек линейного пространства восприятия объектов,
$\{\boldsymbol{x}(\omega) \in \mathbb{X} : \omega \in \Omega\} \subset \mathbb{X}$ – подмножество образов конкретных объектов.

$\boldsymbol{\phi} \in \mathbb{X}$ – нулевая точка линейного пространства,
$\boldsymbol{x}, \boldsymbol{v} \in \mathbb{X}$ – две произвольные точки в линейном пространстве,
$K(\boldsymbol{x}, \boldsymbol{v}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ – скалярная функция, обладающая следующими свойствами:
(1) $K(\boldsymbol{x}, \boldsymbol{v}) = K(\boldsymbol{v}, \boldsymbol{x})$ – симметричность,
(2) $K(\boldsymbol{x},\ c'\boldsymbol{v}' + c''\boldsymbol{v}'') = c'K(\boldsymbol{x}, \boldsymbol{v}') + c''K(\boldsymbol{x}, \boldsymbol{v}'')$ – билинейность.

Если еще добавить предположение
(3) $K(\boldsymbol{x}, \boldsymbol{x}) \geq 0$, то это будет обычное скалярное произведение, тогда
$\sqrt{K(\boldsymbol{x}, \boldsymbol{x})} = \|\boldsymbol{x}\|$ – норма.

Но нам будет достаточно свойств (1) и (2).
Это индефинитное скалярное произведение.
Псевдоевклидово линейное пространство.
Нормы нет.

Линейное пространство с индефинитным скалярным произведением

$\omega \in \Omega$ – множество объектов реального мира,
$y \in \mathbb{Y}$ – множество значений целевой характеристики,
$\boldsymbol{x} \in \mathbb{X}$ – множество точек линейного пространства восприятия объектов,
$\left\{ \boldsymbol{x}(\omega) \in \mathbb{X} : \omega \in \Omega \right\} \subset \mathbb{X}$ – подмножество образов конкретных объектов.

$\boldsymbol{\phi} \in \mathbb{X}$ – нулевая точка линейного пространства,
$\boldsymbol{x}, \boldsymbol{v} \in \mathbb{X}$ – две произвольные точки в линейном пространстве,
$K(\boldsymbol{x}, \boldsymbol{v}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ – скалярная функция, обладающая следующими свойствами:
(1) $K(\boldsymbol{x}, \boldsymbol{v}) = K(\boldsymbol{v}, \boldsymbol{x})$ – симметричность,
(2) $K(\boldsymbol{x},\ c'\boldsymbol{v}' + c''\boldsymbol{v}'') = c'K(\boldsymbol{x}, \boldsymbol{v}') + c''K(\boldsymbol{x}, \boldsymbol{v}'')$ – билинейность.

Если еще добавить предположение
(3) $K(\boldsymbol{x}, \boldsymbol{x}) \geq 0$, то это будет обычное скалярное произведение, тогда
$\sqrt{K(\boldsymbol{x}, \boldsymbol{x})} = \|\boldsymbol{x}\|$ – норма.

Но нам будет достаточно свойств (1) и (2).
Это индефинитное скалярное произведение.
Псевдоевклидово линейное пространство.
Нормы нет.

Марк Григорьевич Крейн (1907-1989), профессор Одесского инженерно-строительного института.

Линейное пространство с индефинитным скалярным произведением

$\omega \in \Omega$ – множество объектов реального мира,
$y \in \mathbb{Y}$ – множество значений целевой характеристики,
$\boldsymbol{x} \in \mathbb{X}$ – множество точек линейного пространства восприятия объектов,
$\left\{ \boldsymbol{x}(\omega) \in \mathbb{X} : \omega \in \Omega \right\} \subset \mathbb{X}$ – подмножество образов конкретных объектов.

$\boldsymbol{\phi} \in \mathbb{X}$ – нулевая точка линейного пространства,
$\boldsymbol{x}, \boldsymbol{v} \in \mathbb{X}$ – две произвольные точки в линейном пространстве,
$K(\boldsymbol{x}, \boldsymbol{v}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ – скалярная функция, обладающая следующими свойствами:
(1) $K(\boldsymbol{x}, \boldsymbol{v}) = K(\boldsymbol{v}, \boldsymbol{x})$ – симметричность,
(2) $K(\boldsymbol{x},\ c'\boldsymbol{v}' + c''\boldsymbol{v}'') = c'K(\boldsymbol{x}, \boldsymbol{v}') + c''K(\boldsymbol{x}, \boldsymbol{v}'')$ – билинейность.

Итак, псевдоевклидово линейное пространство.

Заметим, что если $K_1(\boldsymbol{x}, \boldsymbol{v})$ и $K_2(\boldsymbol{x}, \boldsymbol{v})$ – разные индефинитные скалярные произведения, то их линейная комбинация $K(\boldsymbol{x}, \boldsymbol{v}) = a_1 K_1(\boldsymbol{x}, \boldsymbol{v}) + a_2 K_2(\boldsymbol{x}, \boldsymbol{v})$ с неотрицательными коэффициентами $a_1 \geq 0$ и $a_2 \geq 0$ тоже есть индефинитное скалярное произведение.

Тривиальные примеры погружения множества объектов реального мира в линейное пространство

Множество объектов реального мира $\omega \in \Omega$

Тривиальные примеры погружения множества объектов реального мира в линейное пространство

Множество объектов реального мира $\omega \in \Omega$

Пример 1. Пусть это конечное множество, $|\Omega| = M$ – очень большое число.

Тривиальные примеры погружения множества объектов реального мира в линейное пространство

Множество объектов реального мира $\omega \in \Omega$

Пример 1. Пусть это конечное множество, $|\Omega| = M$ – очень большое число.

С каждым объектом свяжем M-мерный вектор $\boldsymbol{x}_\omega \in \mathbb{R}^M$:

$\boldsymbol{x}_\omega = (x_{\omega,i},\ i \in \Omega),\ x_{\omega,i} = \begin{cases} 1, & i = \omega, \\ 0, & i \neq \omega. \end{cases}$ Это индикаторная функция на Ω, выделяющая данный объект.

Тривиальные примеры погружения множества объектов реального мира в линейное пространство

Множество объектов реального мира $\omega \in \Omega$

Пример 1. Пусть это конечное множество, $|\Omega| = M$ – очень большое число.

С каждым объектом свяжем M-мерный вектор $\boldsymbol{x}_\omega \in \mathbb{R}^M$:

$\boldsymbol{x}_\omega = (x_{\omega,i},\ i \in \Omega),\ x_{\omega,i} = \begin{cases} 1, & i = \omega, \\ 0, & i \neq \omega. \end{cases}$ Это индикаторная функция на Ω, выделяющая данный объект.

Мы погрузили множество объектов в линейное пространство $\mathbb{X} = \mathbb{R}^M$.

Нулевой элемент $\boldsymbol{\phi} = \mathbf{0} \in \mathbb{R}^M$, скалярное произведение $K(\boldsymbol{x}', \boldsymbol{x}'') = \sum_{i \in \Omega} x_i' x_i''$,

норма $\|\boldsymbol{x}\| = \big(K(\boldsymbol{x}, \boldsymbol{x})\big)^{1/2} = \big(\sum_{i \in \Omega} x_i^2\big)^{1/2}$, евклидова метрика $\rho(\boldsymbol{x}', \boldsymbol{x}'') = \big(K(\boldsymbol{x}' - \boldsymbol{x}'', \boldsymbol{x}' - \boldsymbol{x}'')\big)^{1/2}$.

Тривиальные примеры погружения множества объектов реального мира в линейное пространство

Множество объектов реального мира $\omega \in \Omega$

Пример 1. Пусть это конечное множество, $|\Omega| = M$ – очень большое число.

С каждым объектом свяжем M-мерный вектор $\boldsymbol{x}_\omega \in \mathbb{R}^M$:

$\boldsymbol{x}_\omega = (x_{\omega,i},\ i \in \Omega),\ x_{\omega,i} = \begin{cases} 1, & i = \omega, \\ 0, & i \neq \omega. \end{cases}$ Это индикаторная функция на Ω, выделяющая данный объект.

Мы погрузили множество объектов в линейное пространство $\mathbb{X} = \mathbb{R}^M$.

Нулевой элемент $\boldsymbol{\phi} = \boldsymbol{0} \in \mathbb{R}^M$, скалярное произведение $K(\boldsymbol{x}', \boldsymbol{x}'') = \sum_{i \in \Omega} x_i' x_i''$,

норма $\|\boldsymbol{x}\| = \big(K(\boldsymbol{x}, \boldsymbol{x})\big)^{1/2} = \Big(\sum_{i \in \Omega} x_i^2\Big)^{1/2}$, евклидова метрика $\rho(\boldsymbol{x}', \boldsymbol{x}'') = \big(K(\boldsymbol{x}' - \boldsymbol{x}'', \boldsymbol{x}' - \boldsymbol{x}'')\big)^{1/2}$.

Пример 2. Пусть $\Omega^0 = \{\omega_1^0, \dots, \omega_n^0\} \subset \Omega$ – выделенное наблюдателем базисное подмножество объектов.

Тривиальные примеры погружения множества объектов реального мира в линейное пространство

Множество объектов реального мира $\omega \in \Omega$

Пример 1. Пусть это конечное множество, $|\Omega| = M$ – очень большое число.

С каждым объектом свяжем M-мерный вектор $\boldsymbol{x}_\omega \in \mathbb{R}^M$:

$\boldsymbol{x}_\omega = (x_{\omega,i},\ i \in \Omega),\ x_{\omega,i} = \begin{cases} 1, & i = \omega, \\ 0, & i \neq \omega. \end{cases}$ Это индикаторная функция на Ω, выделяющая данный объект.

Мы погрузили множество объектов в линейное пространство $\mathbb{X} = \mathbb{R}^M$.

Нулевой элемент $\boldsymbol{\phi} = \boldsymbol{0} \in \mathbb{R}^M$, скалярное произведение $K(\boldsymbol{x}', \boldsymbol{x}'') = \sum_{i \in \Omega} x'_i x''_i$,

норма $\|\boldsymbol{x}\| = \left(K(\boldsymbol{x}, \boldsymbol{x})\right)^{1/2} = \left(\sum_{i \in \Omega} x_i^2\right)^{1/2}$, евклидова метрика $\rho(\boldsymbol{x}', \boldsymbol{x}'') = \left(K(\boldsymbol{x}' - \boldsymbol{x}'', \boldsymbol{x}' - \boldsymbol{x}'')\right)^{1/2}$.

Пример 2. Пусть $\Omega^0 = \{\omega_1^0, \ldots, \omega_n^0\} \subset \Omega$ – выделенное наблюдателем базисное подмножество объектов.

Произвольная функция парного сравнения объектов $S(\omega', \omega''): \Omega \times \Omega \to \mathbb{R}$.

Тривиальные примеры погружения множества объектов реального мира в линейное пространство

Множество объектов реального мира $\omega \in \Omega$

Пример 1. Пусть это конечное множество, $|\Omega| = M$ – очень большое число.

С каждым объектом свяжем M-мерный вектор $\boldsymbol{x}_\omega \in \mathbb{R}^M$:

$$\boldsymbol{x}_\omega = (x_{\omega,i},\ i \in \Omega),\ x_{\omega,i} = \begin{cases} 1, & i = \omega, \\ 0, & i \neq \omega. \end{cases}$$

Это индикаторная функция на Ω, выделяющая данный объект.

Мы погрузили множество объектов в линейное пространство $\mathbb{X} = \mathbb{R}^M$.

Нулевой элемент $\boldsymbol{\phi} = \boldsymbol{0} \in \mathbb{R}^M$, скалярное произведение $K(\boldsymbol{x}', \boldsymbol{x}'') = \sum_{i \in \Omega} x'_i x''_i$,

норма $\|\boldsymbol{x}\| = \big(K(\boldsymbol{x}, \boldsymbol{x})\big)^{1/2} = \Big(\sum_{i \in \Omega} x_i^2\Big)^{1/2}$, евклидова метрика $\rho(\boldsymbol{x}', \boldsymbol{x}'') = \big(K(\boldsymbol{x}' - \boldsymbol{x}'', \boldsymbol{x}' - \boldsymbol{x}'')\big)^{1/2}$.

Пример 2. Пусть $\Omega^0 = \{\omega_1^0, \ldots, \omega_n^0\} \subset \Omega$ – выделенное наблюдателем базисное подмножество объектов.

Произвольная функция парного сравнения объектов $S(\omega', \omega'') : \Omega \times \Omega \to \mathbb{R}$.

С каждым объектом свяжем n-мерный вектор $\boldsymbol{x}_\omega \in \mathbb{R}^n$:

$$\boldsymbol{x}_\omega = \big(S(\omega, \omega_i),\ i \in \Omega^0\big).$$

Это вектор результатов сравнения данного объекта с базисными объектами.

Мы погрузили множество объектов в линейное пространство $\mathbb{X} = \mathbb{R}^n$.

Тривиальные примеры погружения множества объектов реального мира в линейное пространство

Множество объектов реального мира $\omega \in \Omega$

Пример 1. Пусть это конечное множество, $|\Omega| = M$ – очень большое число.

С каждым объектом свяжем M-мерный вектор $\boldsymbol{x}_\omega \in \mathbb{R}^M$:

$\boldsymbol{x}_\omega = (x_{\omega,i},\ i \in \Omega),\ x_{\omega,i} = \begin{cases} 1, & i = \omega, \\ 0, & i \neq \omega. \end{cases}$ Это индикаторная функция на Ω, выделяющая данный объект.

Мы погрузили множество объектов в линейное пространство $\mathbb{X} = \mathbb{R}^M$.

Нулевой элемент $\boldsymbol{\phi} = \boldsymbol{0} \in \mathbb{R}^M$, скалярное произведение $K(\boldsymbol{x}', \boldsymbol{x}'') = \sum_{i \in \Omega} x_i' x_i''$,

норма $\|\boldsymbol{x}\| = \big(K(\boldsymbol{x}, \boldsymbol{x})\big)^{1/2} = \Big(\sum_{i \in \Omega} x_i^2\Big)^{1/2}$, евклидова метрика $\rho(\boldsymbol{x}', \boldsymbol{x}'') = \big(K(\boldsymbol{x}' - \boldsymbol{x}'', \boldsymbol{x}' - \boldsymbol{x}'')\big)^{1/2}$.

Пример 2. Пусть $\Omega^0 = \{\omega_1^0, \ldots, \omega_n^0\} \subset \Omega$ – выделенное наблюдателем базисное подмножество объектов.

Произвольная функция парного сравнения объектов $S(\omega', \omega'') : \Omega \times \Omega \to \mathbb{R}$.

С каждым объектом свяжем n-мерный вектор $\boldsymbol{x}_\omega \in \mathbb{R}^n$:

$\boldsymbol{x}_\omega = \big(S(\omega, \omega_i),\ i \in \Omega^0\big)$. Это вектор результатов сравнения данного объекта с базисными объектами.

Мы погрузили множество объектов в линейное пространство $\mathbb{X} = \mathbb{R}^n$.

Нулевой элемент $\boldsymbol{\phi} = \boldsymbol{0} \in \mathbb{R}^n$, скалярное произведение $K(\boldsymbol{x}', \boldsymbol{x}'') = \sum_{i \in \Omega} x_i' x_i''$,

норма $\|\boldsymbol{x}\| = \big(K(\boldsymbol{x}, \boldsymbol{x})\big)^{1/2} = \Big(\sum_{i \in \Omega} x_i^2\Big)^{1/2}$, евклидова метрика $\rho(\boldsymbol{x}', \boldsymbol{x}'') = \big(K(\boldsymbol{x}' - \boldsymbol{x}'', \boldsymbol{x}' - \boldsymbol{x}'')\big)^{1/2}$.

Обобщенная линейная модель зависимости

Образ объекта в линейном пространстве $\boxed{\boldsymbol{x}(\omega) \in \mathbb{X} \text{ – фантазия наблюдателя}}$

Обобщенная линейная модель зависимости

Образ объекта в линейном пространстве $\boxed{\boldsymbol{x}(\omega) \in \mathbb{X} \text{ – фантазия наблюдателя}}$

Скалярное произведение $\boxed{K(\boldsymbol{x}, \boldsymbol{v}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R} \text{ – фантазия наблюдателя}}$

Обобщенная линейная модель зависимости

Образ объекта в линейном пространстве $\boxed{\boldsymbol{x}(\omega) \in \mathbb{X} \text{ – фантазия наблюдателя}}$

Скалярное произведение $\boxed{K(\boldsymbol{x}, v) : \mathbb{X} \times \mathbb{X} \to \mathbb{R} \text{ – фантазия наблюдателя}}$

Параметры модели зависимости (v, b) $\begin{cases} v \in \mathbb{X} \text{ – направляющая точка } (\text{вектор}) \text{ в том же пространстве} \\ b \in \mathbb{R} \text{ – сдвиг модели} \end{cases}$

Обобщенная линейная модель зависимости

Образ объекта в линейном пространстве $\boxed{\boldsymbol{x}(\omega) \in \mathbb{X} \text{ – фантазия наблюдателя}}$

Скалярное произведение $\boxed{K(\boldsymbol{x}, \boldsymbol{v}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R} \text{ – фантазия наблюдателя}}$

Параметры модели зависимости $(\boldsymbol{v}, b)$ $\begin{cases} \boldsymbol{v} \in \mathbb{X} \text{ – направляющая точка } (\text{вектор}) \text{ в том же пространстве} \\ b \in \mathbb{R} \text{ – сдвиг модели} \end{cases}$

Обобщенный линейный признак объекта $z(\boldsymbol{x}, \boldsymbol{v}, b) = K(\boldsymbol{x}, \boldsymbol{v}) + b$: $\mathbb{X} \xrightarrow{(\boldsymbol{v}, b)} \mathbb{R}$

Обобщенная линейная модель зависимости

Образ объекта в линейном пространстве $\boxed{\boldsymbol{x}(\omega) \in \mathbb{X} \text{ – фантазия наблюдателя}}$

Скалярное произведение $\boxed{K(\boldsymbol{x}, \boldsymbol{v}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R} \text{ – фантазия наблюдателя}}$

Параметры модели зависимости $(\boldsymbol{v}, b)$ $\begin{cases} \boldsymbol{v} \in \mathbb{X} \text{ – направляющая точка } (\text{вектор}) \text{ в том же пространстве} \\ b \in \mathbb{R} \text{ – сдвиг модели} \end{cases}$

Обобщенный линейный признак объекта $z(\boldsymbol{x}, \boldsymbol{v}, b) = K(\boldsymbol{x}, \boldsymbol{v}) + b$: $\mathbb{X} \xrightarrow{(\boldsymbol{v}, b)} \mathbb{R}$

Целевая характеристика объекта $y(\omega) \in \mathbb{Y}$ – задана природой

Обобщенная линейная модель зависимости

Образ объекта в линейном пространстве $\boxed{\boldsymbol{x}(\omega) \in \mathbb{X} \text{ – фантазия наблюдателя}}$

Скалярное произведение $\boxed{K(\boldsymbol{x}, v): \mathbb{X} \times \mathbb{X} \to \mathbb{R} \text{ – фантазия наблюдателя}}$

Параметры модели зависимости (v, b) $\begin{cases} v \in \mathbb{X} \text{ – направляющая точка (вектор) в том же пространстве} \\ b \in \mathbb{R} \text{ – сдвиг модели} \end{cases}$

Обобщенный линейный признак объекта $z(\boldsymbol{x}, v, b) = K(\boldsymbol{x}, v) + b$: $\mathbb{X} \xrightarrow{(v,b)} \mathbb{R}$

Целевая характеристика объекта $y(\omega) \in \mathbb{Y}$ – задана природой

Функция связи (link function), обычно выпуклая по z $\boxed{q(y, z): \mathbb{Y} \times \mathbb{R} \to \mathbb{R}^{+} \text{ – фантазия наблюдателя}}$

Параметрическая функция потерь $q(y, \boldsymbol{x}, v, b) = q\big(y, z(\boldsymbol{x}, v, b)\big): \mathbb{Y} \times \mathbb{X} \xrightarrow{(\boldsymbol{a},b)} \mathbb{R}^{+}$

Обобщенная линейная модель зависимости

Образ объекта в линейном пространстве $\boxed{\boldsymbol{x}(\omega) \in \mathbb{X} \text{ – фантазия наблюдателя}}$

Скалярное произведение $\boxed{K(\boldsymbol{x}, \boldsymbol{v}): \mathbb{X} \times \mathbb{X} \to \mathbb{R} \text{ – фантазия наблюдателя}}$

Параметры модели зависимости $(\boldsymbol{v}, b)$ $\begin{cases} \boldsymbol{v} \in \mathbb{X} \text{ – направляющая точка } \big(\text{вектор}\big) \text{ в том же пространстве} \\ b \in \mathbb{R} \text{ – сдвиг модели} \end{cases}$

Обобщенный линейный признак объекта $z(\boldsymbol{x}, \boldsymbol{v}, b) = K(\boldsymbol{x}, \boldsymbol{v}) + b: \ \mathbb{X} \xrightarrow{(\boldsymbol{v},b)} \mathbb{R}$

Целевая характеристика объекта $y(\omega) \in \mathbb{Y}$ – задана природой

Функция связи (link function), обычно выпуклая по z $\boxed{q(y, z): \mathbb{Y} \times \mathbb{R} \to \mathbb{R}^+ \text{ – фантазия наблюдателя}}$

Параметрическая функция потерь $q(y, \boldsymbol{x}, \boldsymbol{v}, b) = q\big(y, z(\boldsymbol{x}, \boldsymbol{v}, b)\big): \mathbb{Y} \times \mathbb{X} \xrightarrow{(\boldsymbol{a},b)} \mathbb{R}^+$

Решающее правило $\hat{y}(\boldsymbol{x} \mid \boldsymbol{v}, b) = \underset{y \in \mathbb{Y}}{\operatorname{argmin}}\, q(y, \boldsymbol{x}, \boldsymbol{v}, b): \ \mathbb{X} \xrightarrow{(\boldsymbol{v},b)} \mathbb{Y}$

Обобщенная линейная модель зависимости

Образ объекта в линейном пространстве $\boxed{\boldsymbol{x}(\omega)\in\mathbb{X}\ \text{– фантазия наблюдателя}}$

Скалярное произведение $\boxed{K(\boldsymbol{x},\boldsymbol{\upsilon}):\mathbb{X}\times\mathbb{X}\to\mathbb{R}\ \text{– фантазия наблюдателя}}$

Параметры модели зависимости $(\boldsymbol{\upsilon},b)$ $\begin{cases}\boldsymbol{\upsilon}\in\mathbb{X}\ \text{– направляющая точка}\ (\text{вектор})\ \text{в том же пространстве}\\ b\in\mathbb{R}\ \text{– сдвиг модели}\end{cases}$

Обобщенный линейный признак объекта $z(\boldsymbol{x},\boldsymbol{\upsilon},b)=K(\boldsymbol{x},\boldsymbol{\upsilon})+b:\ \mathbb{X}\xrightarrow{(\boldsymbol{\upsilon},b)}\mathbb{R}$

Целевая характеристика объекта $y(\omega)\in\mathbb{Y}$ – задана природой

Функция связи (link function), обычно выпуклая по z $\boxed{q(y,z):\mathbb{Y}\times\mathbb{R}\to\mathbb{R}^{+}\ \text{– фантазия наблюдателя}}$

Параметрическая функция потерь $q(y,\boldsymbol{x},\boldsymbol{\upsilon},b)=q\big(y,z(\boldsymbol{x},\boldsymbol{\upsilon},b)\big):\mathbb{Y}\times\mathbb{X}\xrightarrow{(\boldsymbol{a},b)}\mathbb{R}^{+}$

Решающее правило $\hat{y}(\boldsymbol{x}\,|\,\boldsymbol{\upsilon},b)=\operatorname*{argmin}_{y\in\mathbb{Y}}q(y,\boldsymbol{x},\boldsymbol{\upsilon},b):\ \mathbb{X}\xrightarrow{(\boldsymbol{\upsilon},b)}\mathbb{Y}$

Обучающая совокупность $(\boldsymbol{X},\boldsymbol{Y})=\Big\{\Big(\boldsymbol{x}(\omega_j),y(\omega_j)\Big)=(\boldsymbol{x}_j,y_j),\ j=1,\ldots,N\Big\}$

Обобщенная линейная модель зависимости

Образ объекта в линейном пространстве $\boxed{\boldsymbol{x}(\omega) \in \mathbb{X} \text{ – фантазия наблюдателя}}$

Скалярное произведение $\boxed{K(\boldsymbol{x}, \boldsymbol{v}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R} \text{ – фантазия наблюдателя}}$

Параметры модели зависимости $(\boldsymbol{v}, b)$ $\begin{cases} \boldsymbol{v} \in \mathbb{X} \text{ – направляющая точка } \big(\text{вектор}\big) \text{ в том же пространстве} \\ b \in \mathbb{R} \text{ – сдвиг модели} \end{cases}$

Обобщенный линейный признак объекта $z(\boldsymbol{x}, \boldsymbol{v}, b) = K(\boldsymbol{x}, \boldsymbol{v}) + b$: $\mathbb{X} \xrightarrow{(\boldsymbol{v}, b)} \mathbb{R}$

Целевая характеристика объекта $y(\omega) \in \mathbb{Y}$ – задана природой

Функция связи (link function), обычно выпуклая по z $\boxed{q(y, z) : \mathbb{Y} \times \mathbb{R} \to \mathbb{R}^{+} \text{ – фантазия наблюдателя}}$

Параметрическая функция потерь $q(y, \boldsymbol{x}, \boldsymbol{v}, b) = q\big(y, z(\boldsymbol{x}, \boldsymbol{v}, b)\big) : \mathbb{Y} \times \mathbb{X} \xrightarrow{(\boldsymbol{a}, b)} \mathbb{R}^{+}$

Решающее правило $\hat{y}(\boldsymbol{x} \mid \boldsymbol{v}, b) = \operatorname*{argmin}_{y \in \mathbb{Y}} q(y, \boldsymbol{x}, \boldsymbol{v}, b)$: $\mathbb{X} \xrightarrow{(\boldsymbol{v}, b)} \mathbb{Y}$

Обучающая совокупность $(\boldsymbol{X}, \boldsymbol{Y}) = \Big\{\Big(\boldsymbol{x}(\omega_j), y(\omega_j)\Big) = (\boldsymbol{x}_j, y_j),\ j = 1, \ldots, N\Big\}$

Эмпирический риск $$Q(\boldsymbol{v}, b \mid X, Y) = \frac{1}{N} \sum_{j=1}^{N} q(y_j, \boldsymbol{x}_j, \boldsymbol{v}, b)$$

Обобщенная линейная модель зависимости

Образ объекта в линейном пространстве $\boxed{\boldsymbol{x}(\omega)\in\mathbb{X} \text{ – фантазия наблюдателя}}$

Скалярное произведение $\boxed{K(\boldsymbol{x},\boldsymbol{\upsilon}): \mathbb{X}\times\mathbb{X}\to\mathbb{R} \text{ – фантазия наблюдателя}}$

Параметры модели зависимости $(\boldsymbol{\upsilon}, b)$ $\begin{cases}\boldsymbol{\upsilon}\in\mathbb{X} \text{ – направляющая точка } (\text{вектор}) \text{ в том же пространстве} \\ b\in\mathbb{R} \text{ – сдвиг модели}\end{cases}$

Обобщенный линейный признак объекта $z(\boldsymbol{x},\boldsymbol{\upsilon},b)=K(\boldsymbol{x},\boldsymbol{\upsilon})+b$: $\mathbb{X}\xrightarrow{(\boldsymbol{\upsilon},b)}\mathbb{R}$

Целевая характеристика объекта $y(\omega)\in\mathbb{Y}$ – задана природой

Функция связи (link function), обычно выпуклая по z $\boxed{q(y,z): \mathbb{Y}\times\mathbb{R}\to\mathbb{R}^{+} \text{ – фантазия наблюдателя}}$

Параметрическая функция потерь $q(y,\boldsymbol{x},\boldsymbol{\upsilon},b)=q\big(y,z(\boldsymbol{x},\boldsymbol{\upsilon},b)\big): \mathbb{Y}\times\mathbb{X}\xrightarrow{(\boldsymbol{a},b)}\mathbb{R}^{+}$

Решающее правило $\hat{y}(\boldsymbol{x}\,|\,\boldsymbol{\upsilon},b)=\operatorname*{argmin}_{y\in\mathbb{Y}} q(y,\boldsymbol{x},\boldsymbol{\upsilon},b)$: $\mathbb{X}\xrightarrow{(\boldsymbol{\upsilon},b)}\mathbb{Y}$

Обучающая совокупность $(\boldsymbol{X},\boldsymbol{Y})=\Big\{\big(\boldsymbol{x}(\omega_j),y(\omega_j)\big)=(\boldsymbol{x}_j,y_j),\ j=1,\ldots,N\Big\}$

Эмпирический риск $Q(\boldsymbol{\upsilon},b\,|\,X,Y)=\dfrac{1}{N}\sum_{j=1}^{N}q(y_j,\boldsymbol{x}_j,\boldsymbol{\upsilon},b)$

Обучение – выбор $\big(\boldsymbol{\upsilon}\in\mathbb{X}, b\in\mathbb{R}\big)$**:**
Минимизация эмпирического риска

$$\sum_{j=1}^{N}q\big(y_j,z(\boldsymbol{x}_j,\boldsymbol{\upsilon},b)\big)\to\min(\boldsymbol{\upsilon},b),$$
$$z(\boldsymbol{x}_j,\boldsymbol{\upsilon},b)=K(\boldsymbol{x}_j,\boldsymbol{\upsilon})+b$$

Обобщенная линейная модель зависимости

Образ объекта в линейном пространстве $\boxed{\boldsymbol{x}(\omega) \in \mathbb{X} \text{ – фантазия наблюдателя}}$

Скалярное произведение $\boxed{K(\boldsymbol{x}, \boldsymbol{v}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R} \text{ – фантазия наблюдателя}}$

Параметры модели зависимости $(\boldsymbol{v}, b)$ $\begin{cases} \boldsymbol{v} \in \mathbb{X} \text{ – направляющая точка } \big(\text{вектор}\big) \text{ в том же пространстве} \\ b \in \mathbb{R} \text{ – сдвиг модели} \end{cases}$

Обобщенный линейный признак объекта $z(\boldsymbol{x}, \boldsymbol{v}, b) = K(\boldsymbol{x}, \boldsymbol{v}) + b: \ \mathbb{X} \xrightarrow{(\boldsymbol{v}, b)} \mathbb{R}$

Целевая характеристика объекта $y(\omega) \in \mathbb{Y}$ – задана природой

Функция связи (link function), обычно выпуклая по z $\boxed{q(y, z) : \mathbb{Y} \times \mathbb{R} \to \mathbb{R}^{+} \text{ – фантазия наблюдателя}}$

Параметрическая функция потерь $q(y, \boldsymbol{x}, \boldsymbol{v}, b) = q\big(y, z(\boldsymbol{x}, \boldsymbol{v}, b)\big) : \mathbb{Y} \times \mathbb{X} \xrightarrow{(\boldsymbol{a}, b)} \mathbb{R}^{+}$

Обучающая совокупность $(\boldsymbol{X}, \boldsymbol{Y}) = \Big\{ \Big(\boldsymbol{x}(\omega_j), y(\omega_j) \Big) = (\boldsymbol{x}_j, y_j), \ j = 1, \ldots, N \Big\}$

Семейство выпуклых регуляризующих функций $\boxed{V(\boldsymbol{v} \,|\, \mu) : \ \mathbb{X} \xrightarrow{\mu} \mathbb{R}^{+} \text{ – еще одна фантазия наблюдателя}}$

Обобщенная линейная модель зависимости

Образ объекта в линейном пространстве $\boxed{\boldsymbol{x}(\omega) \in \mathbb{X} \text{ – фантазия наблюдателя}}$

Скалярное произведение $\boxed{K(\boldsymbol{x},\boldsymbol{v}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R} \text{ – фантазия наблюдателя}}$

Параметры модели зависимости $(\boldsymbol{v},b)$ $\begin{cases} \boldsymbol{v} \in \mathbb{X} \text{ – направляющая точка } (\text{вектор}) \text{ в том же пространстве} \\ b \in \mathbb{R} \text{ – сдвиг модели} \end{cases}$

Обобщенный линейный признак объекта $z(\boldsymbol{x},\boldsymbol{v},b) = K(\boldsymbol{x},\boldsymbol{v}) + b : \ \mathbb{X} \xrightarrow{(\boldsymbol{v},b)} \mathbb{R}$

Целевая характеристика объекта $y(\omega) \in \mathbb{Y}$ – задана природой

Функция связи (link function), обычно выпуклая по z $\boxed{q(y,z) : \mathbb{Y} \times \mathbb{R} \to \mathbb{R}^+ \text{ – фантазия наблюдателя}}$

Параметрическая функция потерь $q(y,\boldsymbol{x},\boldsymbol{v},b) = q\big(y, z(\boldsymbol{x},\boldsymbol{v},b)\big) : \mathbb{Y} \times \mathbb{X} \xrightarrow{(\boldsymbol{a},b)} \mathbb{R}^+$

Обучающая совокупность $(\boldsymbol{X},\boldsymbol{Y}) = \Big\{\big(\boldsymbol{x}(\omega_j), y(\omega_j)\big) = (\boldsymbol{x}_j, y_j),\ j = 1,\ldots,N\Big\}$

Семейство выпуклых регуляризующих функций $\boxed{V(\boldsymbol{v} \mid \mu) : \ \mathbb{X} \xrightarrow{\mu} \mathbb{R}^+ \text{ – еще одна фантазия наблюдателя}}$

Обучение – выбор $(\boldsymbol{v} \in \mathbb{X}, b \in \mathbb{R})$:
Минимизация регуляризованного эмпирического риска

$$V(\boldsymbol{v} \mid \mu) + c\sum_{j=1}^{N} q\big(y_j, z(\boldsymbol{x}_j,\boldsymbol{v},b)\big) \to \min(\boldsymbol{v},b),$$
$$z(\boldsymbol{x}_j,\boldsymbol{v},b) = K(\boldsymbol{x}_j,\boldsymbol{v}) + b$$

Обобщенная линейная модель зависимости

Образ объекта в линейном пространстве $\boxed{\mathbf{x}(\omega) \in \mathbb{X} \text{ – фантазия наблюдателя}}$

Скалярное произведение $\boxed{K(\mathbf{x}, \boldsymbol{\upsilon}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R} \text{ – фантазия наблюдателя}}$

Параметры модели зависимости $(\boldsymbol{\upsilon}, b)$ $\begin{cases} \boldsymbol{\upsilon} \in \mathbb{X} \text{ – направляющая точка } (\text{вектор}) \text{ в том же пространстве} \\ b \in \mathbb{R} \text{ – сдвиг модели} \end{cases}$

Обобщенный линейный признак объекта $z(\mathbf{x}, \boldsymbol{\upsilon}, b) = K(\mathbf{x}, \boldsymbol{\upsilon}) + b$: $\mathbb{X} \xrightarrow{(\boldsymbol{\upsilon}, b)} \mathbb{R}$

Целевая характеристика объекта $y(\omega) \in \mathbb{Y}$ – задана природой

Функция связи (link function), обычно выпуклая по z $\boxed{q(y, z) : \mathbb{Y} \times \mathbb{R} \to \mathbb{R}^+ \text{ – фантазия наблюдателя}}$

Параметрическая функция потерь $q(y, \mathbf{x}, \boldsymbol{\upsilon}, b) = q\big(y, z(\mathbf{x}, \boldsymbol{\upsilon}, b)\big) : \mathbb{Y} \times \mathbb{X} \xrightarrow{(\mathbf{a}, b)} \mathbb{R}^+$

Обучающая совокупность $(\mathbf{X}, \mathbf{Y}) = \Big\{\big(\mathbf{x}(\omega_j), y(\omega_j)\big) = (\mathbf{x}_j, y_j),\ j = 1, ..., N\Big\}$

Семейство выпуклых регуляризующих функций $\boxed{V(\boldsymbol{\upsilon} \mid \mu) :\ \mathbb{X} \xrightarrow{\mu} \mathbb{R}^+ \text{ – еще одна фантазия наблюдателя}}$

Обучение – выбор $(\boldsymbol{\upsilon} \in \mathbb{X}, b \in \mathbb{R})$:
Минимизация регуляризованного эмпирического риска

$$V(\boldsymbol{\upsilon} \mid \mu) + c \sum_{j=1}^{N} q\big(y_j, z(\mathbf{x}_j, \boldsymbol{\upsilon}, b)\big) \to \min(\boldsymbol{\upsilon}, b),$$
$$z(\mathbf{x}_j, \boldsymbol{\upsilon}, b) = K(\mathbf{x}_j, \boldsymbol{\upsilon}) + b$$

Однако как оптимизировать критерий в необозримом линейном пространстве $\boldsymbol{\upsilon} \in \mathbb{X}$?

Обобщенная линейная модель зависимости

Образ объекта в линейном пространстве $\boxed{\boldsymbol{x}(\omega) \in \mathbb{X} \text{ – фантазия наблюдателя}}$

Скалярное произведение $\boxed{K(\boldsymbol{x}, \boldsymbol{\vartheta}) : \mathbb{X} \times \mathbb{X} \to \mathbb{R} \text{ – фантазия наблюдателя}}$

Параметры модели зависимости $(\boldsymbol{\vartheta}, b)$ $\begin{cases} \boldsymbol{\vartheta} \in \mathbb{X} \text{ – направляющая точка } \left(\text{вектор}\right) \text{ в том же пространстве} \\ b \in \mathbb{R} \text{ – сдвиг модели} \end{cases}$

Обобщенный линейный признак объекта $z(\boldsymbol{x}, \boldsymbol{\vartheta}, b) = K(\boldsymbol{x}, \boldsymbol{\vartheta}) + b$: $\mathbb{X} \xrightarrow{(\boldsymbol{\vartheta}, b)} \mathbb{R}$

Целевая характеристика объекта $y(\omega) \in \mathbb{Y}$ – задана природой

Функция связи (link function), обычно выпуклая по z $\boxed{q(y, z) : \mathbb{Y} \times \mathbb{R} \to \mathbb{R}^+ \text{ – фантазия наблюдателя}}$

Параметрическая функция потерь $q(y, \boldsymbol{x}, \boldsymbol{\vartheta}, b) = q\big(y, z(\boldsymbol{x}, \boldsymbol{\vartheta}, b)\big) : \mathbb{Y} \times \mathbb{X} \xrightarrow{(\boldsymbol{a}, b)} \mathbb{R}^+$

Обучающая совокупность $(\boldsymbol{X}, \boldsymbol{Y}) = \Big\{\big(\boldsymbol{x}(\omega_j), y(\omega_j)\big) = (\boldsymbol{x}_j, y_j),\ j = 1, \ldots, N\Big\}$

Семейство выпуклых регуляризующих функций $\boxed{V(\boldsymbol{\vartheta} \mid \mu) :\ \mathbb{X} \xrightarrow{\mu} \mathbb{R}^+ \text{ – еще одна фантазия наблюдателя}}$

Обучение – выбор $\left(\boldsymbol{\vartheta} \in \mathbb{X}, b \in \mathbb{R}\right)$**:**
Минимизация регуляризованного эмпирического риска

$$V(\boldsymbol{\vartheta} \mid \mu) + c \sum_{j=1}^{N} q\Big(y_j, z(\boldsymbol{x}_j, \boldsymbol{\vartheta}, b)\Big) \to \min(\boldsymbol{\vartheta}, b),$$
$$z(\boldsymbol{x}_j, \boldsymbol{\vartheta}, b) = K(\boldsymbol{x}_j, \boldsymbol{\vartheta}) + b$$

Однако как оптимизировать критерий в необозримом линейном пространстве $\boldsymbol{\vartheta} \in \mathbb{X}$?
Решение: Искать направляющий вектор как линейную комбинацию образов некоторого подмножества базисных объектов: $\left\{\omega_1^0, \ldots, \omega_n^0\right\} \in \Omega$.

Параметрический критерий обучения

Критерий минимизации регуляризованного эмпирического риска:

$$V(v \mid \mu) + c\sum_{j=1}^{N} q\Big(y_j, z(\boldsymbol{x}_j, v, b)\Big) \to \min(v, b), \quad z(\boldsymbol{x}_j, v, b) = K(\boldsymbol{x}_j, v) + b.$$

Параметрический критерий обучения

Критерий минимизации регуляризованного эмпирического риска:

$$V(\boldsymbol{v}\,|\,\mu)+c\sum_{j=1}^{N} q\Big(y_j, z(\mathbf{x}_j,\boldsymbol{v},b)\Big)\to \min(\boldsymbol{v},b), \quad z(\mathbf{x}_j,\boldsymbol{v},b)=K(\mathbf{x}_j,\boldsymbol{v})+b.$$

Представление искомого направляющего вектора как линейной комбинации:

$$\left\{\omega_1^0,\ldots,\omega_n^0\right\}\in\Omega \;\Rightarrow\; \left\{\mathbf{x}_1^0,\ldots,\mathbf{x}_n^0\right\}, \quad \boldsymbol{v}=\sum_{i=1}^{n} a_i\mathbf{x}_i^0, \quad \boldsymbol{a}=(a_1,\ldots,a_N)\in\mathbb{R}^n.$$

Параметрический критерий обучения

Критерий минимизации регуляризованного эмпирического риска:

$$V(\boldsymbol{v} \mid \mu) + c\sum_{j=1}^{N} q\Big(y_j, z(\boldsymbol{x}_j, \boldsymbol{v}, b)\Big) \to \min(\boldsymbol{v}, b), \quad z(\boldsymbol{x}_j, \boldsymbol{v}, b) = K(\boldsymbol{x}_j, \boldsymbol{v}) + b.$$

Представление искомого направляющего вектора как линейной комбинации:

$$\left\{\omega_1^0, \dots, \omega_n^0\right\} \in \Omega \;\Rightarrow\; \left\{\boldsymbol{x}_1^0, \dots, \boldsymbol{x}_n^0\right\}, \quad \boldsymbol{v} = \sum_{i=1}^{n} a_i \boldsymbol{x}_i^0, \quad \boldsymbol{a} = (a_1, \dots, a_N) \in \mathbb{R}^n.$$

В силу линейности индефинитного скалярного произведения:

$$z(\boldsymbol{x}_j, \boldsymbol{a}, b) = K\left(\boldsymbol{x}_j, \sum_{i=1}^{n} a_i \boldsymbol{x}_i^0\right) + b = \sum_{i=1}^{n} a_i K(\boldsymbol{x}_j, \boldsymbol{x}_i^0) + b$$

Параметрический критерий обучения

Критерий минимизации регуляризованного эмпирического риска:

$$V(v\,|\,\mu)+c\sum_{j=1}^{N}q\Big(y_j,z(\boldsymbol{x}_j,v,b)\Big)\to\min(v,b),\quad z(\boldsymbol{x}_j,v,b)=K(\boldsymbol{x}_j,v)+b.$$

Представление искомого направляющего вектора как линейной комбинации:

$$\left\{\omega_1^0,...,\omega_n^0\right\}\in\Omega\;\Rightarrow\;\left\{\boldsymbol{x}_1^0,...,\boldsymbol{x}_n^0\right\},\quad \boldsymbol{v}=\sum_{i=1}^{n}a_i\boldsymbol{x}_i^0,\quad \boldsymbol{a}=(a_1,...,a_N)\in\mathbb{R}^n.$$

В силу линейности индефинитного скалярного произведения:

$$z(\boldsymbol{x}_j,\boldsymbol{a},b)=K\left(\boldsymbol{x}_j,\sum_{i=1}^{n}a_i\boldsymbol{x}_i^0\right)+b=\sum_{i=1}^{n}a_iK(\boldsymbol{x}_j,\boldsymbol{x}_i^0)+b$$

Достаточно рассматривать регуляризующие функции как функции действительного векторного аргумента: $V(\boldsymbol{a}\,|\,\mu):\;\mathbb{R}^n\to\mathbb{R}^+$.

Параметрический критерий обучения

Критерий минимизации регуляризованного эмпирического риска:

$$V(\boldsymbol{v}\,|\,\mu)+c\sum_{j=1}^{N}q\Big(y_j, z(\mathbf{x}_j,\boldsymbol{v},b)\Big)\to \min(\boldsymbol{v},b),\quad z(\mathbf{x}_j,\boldsymbol{v},b)=K(\mathbf{x}_j,\boldsymbol{v})+b.$$

Представление искомого направляющего вектора как линейной комбинации:

$$\left\{\omega_1^0,\ldots,\omega_n^0\right\}\in\Omega\ \Rightarrow\ \left\{\mathbf{x}_1^0,\ldots,\mathbf{x}_n^0\right\},\quad \boldsymbol{v}=\sum_{i=1}^{n}a_i\mathbf{x}_i^0,\quad \boldsymbol{a}=(a_1,\ldots,a_N)\in\mathbb{R}^n.$$

В силу линейности индефинитного скалярного произведения:

$$z(\mathbf{x}_j,\boldsymbol{a},b)=K\left(\mathbf{x}_j,\sum_{i=1}^{n}a_i\mathbf{x}_i^0\right)+b=\sum_{i=1}^{n}a_iK(\mathbf{x}_j,\mathbf{x}_i^0)+b$$

Достаточно рассматривать регуляризующие функции как функции действительного векторного аргумента: $V(\boldsymbol{a}\,|\,\mu):\ \mathbb{R}^n\to\mathbb{R}^+$.

Параметрический критерий минимизации регуляризованного эмпирического риска:

$$\underbrace{V(\boldsymbol{a}\,|\,\mu)}_{\text{выпуклая регуляризующая функция}}+c\sum_{j=1}^{N}\underbrace{q\Big(y_j,z(\mathbf{x}_j,\boldsymbol{a},b)\Big)}_{q(y,z)-\text{ выпуклая функция связи}}\to\min\Big(\boldsymbol{a}\in\mathbb{R}^n, b\in\mathbb{R}\Big),$$

Параметрический критерий обучения

Критерий минимизации регуляризованного эмпирического риска:

$$V(\boldsymbol{\upsilon}\,|\,\mu)+c\sum_{j=1}^{N}q\Big(y_j, z(\mathbf{x}_j,\boldsymbol{\upsilon},b)\Big)\to\min(\boldsymbol{\upsilon},b),\quad z(\mathbf{x}_j,\boldsymbol{\upsilon},b)=K(\mathbf{x}_j,\boldsymbol{\upsilon})+b.$$

Представление искомого направляющего вектора как линейной комбинации:

$$\left\{\omega_1^0,\dots,\omega_n^0\right\}\in\Omega\ \Rightarrow\ \left\{\mathbf{x}_1^0,\dots,\mathbf{x}_n^0\right\},\quad \boldsymbol{\upsilon}=\sum_{i=1}^{n}a_i\mathbf{x}_i^0,\quad \mathbf{a}=(a_1,\dots,a_N)\in\mathbb{R}^n.$$

В силу линейности индефинитного скалярного произведения:

$$z(\mathbf{x}_j,\mathbf{a},b)=K\left(\mathbf{x}_j,\sum_{i=1}^{n}a_i\mathbf{x}_i^0\right)+b=\sum_{i=1}^{n}a_iK(\mathbf{x}_j,\mathbf{x}_i^0)+b$$

Достаточно рассматривать регуляризующие функции как функции действительного векторного аргумента: $V(\mathbf{a}\,|\,\mu):\ \mathbb{R}^n\to\mathbb{R}^+$.

Параметрический критерий минимизации регуляризованного эмпирического риска:

$$\underbrace{V(\mathbf{a}\,|\,\mu)}_{\substack{\text{выпуклая}\\ \text{регуляризующая}\\ \text{функция}}}+c\sum_{j=1}^{N}\underbrace{q\Big(y_j,z(\mathbf{x}_j,\mathbf{a},b)\Big)}_{\substack{q(y,z)-\text{ выпуклая}\\ \text{функция связи}}}\to\min\Big(\mathbf{a}\in\mathbb{R}^n,b\in\mathbb{R}\Big),$$

$$z(\mathbf{x}_j,\mathbf{a},b)=\sum_{i=1}^{n}a_i\underbrace{K(\mathbf{x}_j,\mathbf{x}_i^0)}_{\substack{\text{скалярное}\\ \text{произведение}\\ \text{(индефинитное)}}}+b.$$

Функции связи для некоторых частных видов целевой переменной

Функция связи (link function) $\boxed{q(y,z): \mathbb{Y} \times \mathbb{R} \to \mathbb{R}^{+} \text{ – фантазия наблюдателя}}$

Функции связи для некоторых частных видов целевой переменной

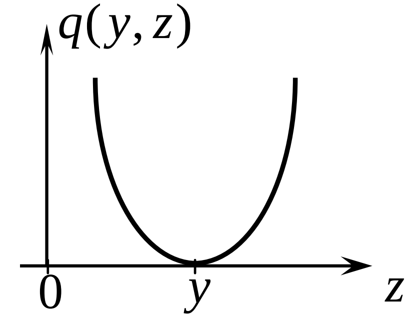
Функция связи (link function) $\boxed{q(y,z): \mathbb{Y}\times\mathbb{R}\to\mathbb{R}^+ \text{ – фантазия наблюдателя}}$

Обобщенный линейный признак объекта $z(\boldsymbol{x},\boldsymbol{a},b)=\sum_{i=1}^{n} a_i K(\boldsymbol{x},\boldsymbol{x}_i^0)+b: \ \boldsymbol{x}\in\mathbb{X}\xrightarrow{(\boldsymbol{a},b)} z\in\mathbb{R}.$

Функции связи для некоторых частных видов целевой переменной

Функция связи (link function) $\boxed{q(y,z): \mathbb{Y}\times\mathbb{R}\to\mathbb{R}^{+} \text{ – фантазия наблюдателя}}$

Обобщенный линейный признак объекта $z(\boldsymbol{x},\boldsymbol{a},b)=\sum_{i=1}^{n} a_i K(\boldsymbol{x},\boldsymbol{x}_i^0)+b:\ \boldsymbol{x}\in\mathbb{X}\xrightarrow{(\boldsymbol{a},b)} z\in\mathbb{R}.$



1) Задача оценивания числовой регрессии.

Целевая переменная $y\in\mathbb{Y}=\mathbb{R}$ – действительное число.

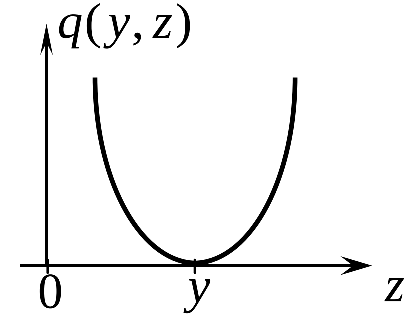
$q(y,z)=(y-z)^2.$

Функции связи для некоторых частных видов целевой переменной

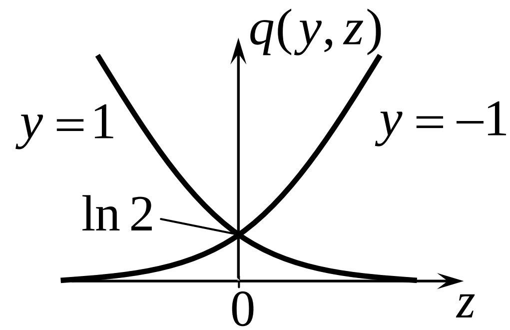
Функция связи (link function) $\boxed{q(y,z): \mathbb{Y}\times\mathbb{R}\to\mathbb{R}^+ \text{ – фантазия наблюдателя}}$

Обобщенный линейный признак объекта $z(\boldsymbol{x},\boldsymbol{a},b)=\sum_{i=1}^{n} a_i K(\boldsymbol{x},\boldsymbol{x}_i^0)+b:\ \boldsymbol{x}\in\mathbb{X}\xrightarrow{(\boldsymbol{a},b)} z\in\mathbb{R}.$

1) Задача оценивания числовой регрессии.
Целевая переменная $y\in\mathbb{Y}=\mathbb{R}$ – действительное число.
$q(y,z)=(y-z)^2.$



2) Задача обучения распознаванию объектов двух классов.
Целевая переменная $y\in\mathbb{Y}=\{-1,1\}$ – индекс класса.
Метод логистической регрессии:
$q(y,z)=\ln\left[1+\exp(-y\,z)\right].$

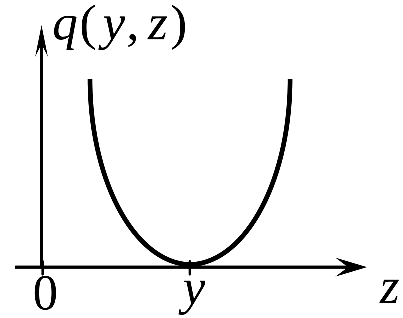


Функции связи для некоторых частных видов целевой переменной

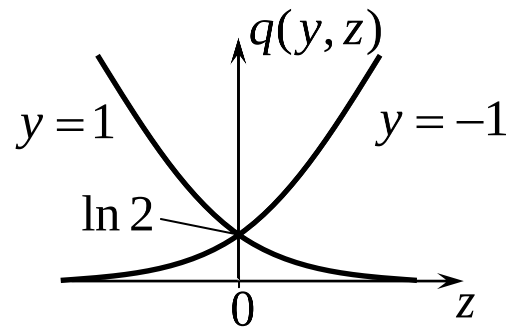
Функция связи (link function) $\boxed{q(y,z): \mathbb{Y}\times\mathbb{R} \to \mathbb{R}^+ \text{ – фантазия наблюдателя}}$

Обобщенный линейный признак объекта $z(\boldsymbol{x},\boldsymbol{a},b) = \sum_{i=1}^{n} a_i K(\boldsymbol{x},\boldsymbol{x}_i^0) + b: \ \boldsymbol{x}\in\mathbb{X} \xrightarrow{(\boldsymbol{a},b)} z\in\mathbb{R}.$

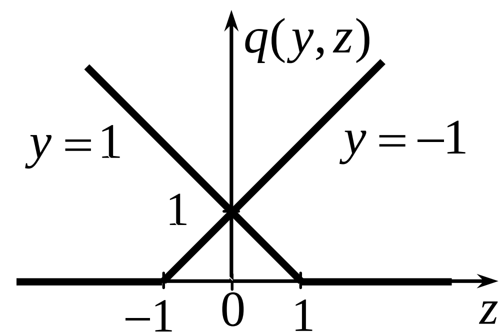
1) Задача оценивания числовой регрессии.
   Целевая переменная $y\in\mathbb{Y}=\mathbb{R}$ – действительное число.
   $q(y,z) = (y-z)^2.$



2) Задача обучения распознаванию объектов двух классов.
   Целевая переменная $y\in\mathbb{Y}=\{-1,1\}$ – индекс класса.
   Метод логистической регрессии:
   $q(y,z) = \ln\left[1+\exp(-y\,z)\right].$



3) Задача обучения распознаванию объектов двух классов.
   Целевая переменная $y\in\mathbb{Y}=\{-1,1\}$ – индекс класса.
   Метод опорных векторов:
   $q(y,z) = \max\left[0,\ 1-yz\right] = \begin{cases} 1-yz, & 1-yz>0, \\ 0, & 1-yz\le 0. \end{cases}$



Некоторые виды регуляризующих функций в конечномерном линейном пространстве

Параметрическое семейство регуляризующих функций

$$\boxed{V(\boldsymbol{a}\,|\,\mu):\ \mathbb{X}=\mathbb{R}^n \xrightarrow{\mu} \mathbb{R}^+ = \text{– фантазия наблюдателя}}$$

Некоторые виды регуляризующих функций в конечномерном линейном пространстве

Параметрическое семейство регуляризующих функций $\boxed{V(\boldsymbol{a}\,|\,\mu):\ \mathbb{X}=\mathbb{R}^n \xrightarrow{\mu} \mathbb{R}^+ = -\text{ фантазия наблюдателя}}$

1) Простейшая квадратичная регуляризация $V(\boldsymbol{a})=\sum_{i=1}^{n} a_i^2$

Некоторые виды регуляризующих функций в конечномерном линейном пространстве

Параметрическое семейство регуляризующих функций $\boxed{V(\boldsymbol{a}\,|\,\mu):\ \mathbb{X}=\mathbb{R}^n \xrightarrow{\mu} \mathbb{R}^+ = -\text{ фантазия наблюдателя}}$

1) Простейшая квадратичная регуляризация $V(\boldsymbol{a})=\sum_{i=1}^{n} a_i^2$

2) Квадратичная регуляризация с разными коэффициентами $V(\boldsymbol{a}\,|\,\boldsymbol{\mu})=\sum_{i=1}^{n} \mu_i a_i^2$

Некоторые виды регуляризующих функций в конечномерном линейном пространстве

Параметрическое семейство регуляризующих функций $\boxed{V(\boldsymbol{a}\,|\,\mu):\ \mathbb{X}=\mathbb{R}^n \xrightarrow{\mu} \mathbb{R}^+ = \text{– фантазия наблюдателя}}$

1) Простейшая квадратичная регуляризация $V(\boldsymbol{a})=\sum_{i=1}^{n} a_i^2$

2) Квадратичная регуляризация с разными коэффициентами $V(\boldsymbol{a}\,|\,\boldsymbol{\mu})=\sum_{i=1}^{n} \mu_i a_i^2$

3) Квадратичная регуляризация для направляющего вектора с упорядоченными компонентами (штраф на «негладкость») $V(\boldsymbol{a})=\sum_{i=1}^{n} (a_i - a_{i-1})^2$

Некоторые виды регуляризующих функций в конечномерном линейном пространстве

Параметрическое семейство регуляризующих функций $\boxed{V(\boldsymbol{a}\,|\,\mu):\ \mathbb{X}=\mathbb{R}^n \xrightarrow{\mu} \mathbb{R}^+ = \text{– фантазия наблюдателя}}$

1) Простейшая квадратичная регуляризация $V(\boldsymbol{a})=\sum_{i=1}^{n} a_i^2$

2) Квадратичная регуляризация с разными коэффициентами $V(\boldsymbol{a}\,|\,\boldsymbol{\mu})=\sum_{i=1}^{n} \mu_i a_i^2$

3) Квадратичная регуляризация для направляющего вектора с упорядоченными компонентами (штраф на «негладкость») $V(\boldsymbol{a})=\sum_{i=1}^{n} (a_i - a_{i-1})^2$

4) Квадратично-модульная регуляризация I (Elastic Net) $V(\boldsymbol{a}\,|\,\mu)=\sum_{i=1}^{n} a_i^2 + \mu\sum_{i=1}^{n} |a_i|$

Некоторые виды регуляризующих функций в конечномерном линейном пространстве

Параметрическое семейство регуляризующих функций $\boxed{V(\boldsymbol{a}\,|\,\mu):\ \mathbb{X}=\mathbb{R}^n \xrightarrow{\ \mu\ } \mathbb{R}^+ = –\ \text{фантазия наблюдателя}}$

1) Простейшая квадратичная регуляризация $V(\boldsymbol{a}) = \sum_{i=1}^{n} a_i^2$

2) Квадратичная регуляризация с разными коэффициентами $V(\boldsymbol{a}\,|\,\boldsymbol{\mu}) = \sum_{i=1}^{n} \mu_i a_i^2$

3) Квадратичная регуляризация для направляющего вектора с упорядоченными компонентами (штраф на «негладкость») $V(\boldsymbol{a}) = \sum_{i=1}^{n} (a_i - a_{i-1})^2$

4) Квадратично-модульная регуляризация I (Elastic Net) $V(\boldsymbol{a}\,|\,\mu) = \sum_{i=1}^{n} a_i^2 + \mu \sum_{i=1}^{n} |a_i|$

5) Квадратично-модульная регуляризация II $V(\boldsymbol{a}\,|\,\mu) = \sum_{i=1}^{n} \begin{pmatrix} 2\mu|\,a_i\,|, & |\,a_i\,| \le \mu, \\ \mu^2 + a_i^2, & |\,a_i\,| > \mu. \end{pmatrix}$

Вероятностная интерпретация минимизации регуляризованного эмпирического риска

Обучение: Минимизация регуляризованного эмпирического риска на обучающей совокупности

$$(X,Y)=\left\{\left(x(\omega_j),y(\omega_j)\right)=(x_j,y_j),\ j=1,...,N\right\}$$

$$V(\boldsymbol{a}\,|\,\mu)+c\sum_{j=1}^{N}q\left(y_j,z(\boldsymbol{x}_j,\boldsymbol{a},b)\right)\to\min(\boldsymbol{a},b),$$

$$z(\boldsymbol{x}_j,\boldsymbol{a},b)=K(\boldsymbol{x}_j,\boldsymbol{a})+b$$

Вероятностная интерпретация минимизации регуляризованного эмпирического риска

Обучение: Минимизация регуляризованного эмпирического риска на обучающей совокупности

$$(X,Y)=\Big\{\Big(x(\omega_j),y(\omega_j)\Big)=(x_j,y_j),\ j=1,...,N\Big\}$$

$$V(\boldsymbol{a}\,|\,\mu)+c\sum_{j=1}^{N}q\Big(y_j,z(\boldsymbol{x}_j,\boldsymbol{a},b)\Big)\to\min(\boldsymbol{a},b),$$

$$z(\boldsymbol{x}_j,\boldsymbol{a},b)=K(\boldsymbol{x}_j,\boldsymbol{a})+b$$

Регуляризующая функция $V(\boldsymbol{a}|\mu)$ и функция связи $q(y,z)$ выпуклы по $\boldsymbol{a}\in\mathbb{R}^n$ и $z\in\mathbb{R}$.

Вероятностная интерпретация минимизации регуляризованного эмпирического риска

Обучение: Минимизация регуляризованного эмпирического риска на обучающей совокупности

$$(X,Y)=\left\{\Big(x(\omega_j),y(\omega_j)\Big)=(x_j,y_j),\ j=1,\dots,N\right\}$$

$$V(\boldsymbol{a}\,|\,\mu)+c\sum_{j=1}^{N}q\Big(y_j,z(\boldsymbol{x}_j,\boldsymbol{a},b)\Big)\to\min(\boldsymbol{a},b),$$

$$z(\boldsymbol{x}_j,\boldsymbol{a},b)=K(\boldsymbol{x}_j,\boldsymbol{a})+b$$

Регуляризующая функция $V(\boldsymbol{a}|\mu)$ и функция связи $q(y,z)$ выпуклы по $\boldsymbol{a}\in\mathbb{R}^n$ и $z\in\mathbb{R}$.

Вероятностные эвристики наблюдателя

Вероятностная интерпретация минимизации регуляризованного эмпирического риска

Обучение: Минимизация регуляризованного эмпирического риска на обучающей совокупности

$$(X,Y)=\Big\{\big(x(\omega_j),y(\omega_j)\big)=(x_j,y_j),\ j=1,...,N\Big\}$$

$$V(\boldsymbol{a}\,|\,\mu)+c\sum_{j=1}^{N}q\big(y_j,z(\boldsymbol{x}_j,\boldsymbol{a},b)\big)\to\min(\boldsymbol{a},b),$$

$$z(\boldsymbol{x}_j,\boldsymbol{a},b)=K(\boldsymbol{x}_j,\boldsymbol{a})+b$$

Регуляризующая функция $V(\boldsymbol{a}|\mu)$ и функция связи $q(y,z)$ выпуклы по $\boldsymbol{a}\in\mathbb{R}^n$ и $z\in\mathbb{R}$.

Вероятностные эвристики наблюдателя

1) Чем меньше параметрическая функция потерь $q\big(y,z(\boldsymbol{x},\boldsymbol{a},b)\big)$, тем более вероятно появление объекта с характеристиками $(\boldsymbol{x},y)$.

Предполагаемое параметрическое семейство совместных распределений $f(\boldsymbol{x},y\,|\,\boldsymbol{a},b,c)\propto\exp\Big\{-cq\big(y,z(\boldsymbol{x},\boldsymbol{a},b)\big)\Big\}$.

Вероятностная интерпретация минимизации регуляризованного эмпирического риска

Обучение: Минимизация регуляризованного эмпирического риска на обучающей совокупности

$$V(\boldsymbol{a}\,|\,\mu)+c\sum_{j=1}^{N} q\Big(y_j, z(\boldsymbol{x}_j,\boldsymbol{a},b)\Big)\to\min(\boldsymbol{a},b),$$

$$(X,Y)=\Big\{\Big(x(\omega_j),y(\omega_j)\Big)=(x_j,y_j),\ j=1,\dots,N\Big\}$$

$$z(\boldsymbol{x}_j,\boldsymbol{a},b)=K(\boldsymbol{x}_j,\boldsymbol{a})+b$$

Регуляризующая функция $V(\boldsymbol{a}|\mu)$ и функция связи $q(y,z)$ выпуклы по $\boldsymbol{a}\in\mathbb{R}^n$ и $z\in\mathbb{R}$.

Вероятностные эвристики наблюдателя

1) Чем меньше параметрическая функция потерь $q\big(y,z(\boldsymbol{x},\boldsymbol{a},b)\big)$, тем более вероятно появление объекта с характеристиками $(\boldsymbol{x},y)$.

Предполагаемое параметрическое семейство совместных распределений $f(\boldsymbol{x},y\,|\,\boldsymbol{a},b,c)\propto\exp\Big\{-cq\Big(y,z(\boldsymbol{x},\boldsymbol{a},b)\Big)\Big\}$.

2) Чем меньше регуляризующая функция $V(\boldsymbol{a}|\mu)$, тем более априори вероятно значение параметра $\boldsymbol{a}\in\mathbb{R}^n$.

Предполагаемое параметрическое семейство априорных распределений $\psi(\boldsymbol{a}|\mu)\propto\exp\big\{-c'V(\boldsymbol{a}|\mu)\big\}$.

Вероятностная интерпретация минимизации регуляризованного эмпирического риска

Обучение: Минимизация регуляризованного эмпирического риска на обучающей совокупности

$$(X,Y)=\Big\{\big(x(\omega_j),y(\omega_j)\big)=(x_j,y_j),\ j=1,...,N\Big\}$$

$$V(\boldsymbol{a}\,|\,\mu)+c\sum_{j=1}^{N}q\big(y_j,z(\boldsymbol{x}_j,\boldsymbol{a},b)\big)\to\min(\boldsymbol{a},b),$$

$$z(\boldsymbol{x}_j,\boldsymbol{a},b)=K(\boldsymbol{x}_j,\boldsymbol{a})+b$$

Регуляризующая функция $V(\boldsymbol{a}|\mu)$ и функция связи $q(y,z)$ выпуклы по $\boldsymbol{a}\in\mathbb{R}^n$ и $z\in\mathbb{R}$.

Вероятностные эвристики наблюдателя

1) Чем меньше параметрическая функция потерь $q\big(y,z(\boldsymbol{x},\boldsymbol{a},b)\big)$, тем более вероятно появление объекта с характеристиками $(\boldsymbol{x},y)$.

Предполагаемое параметрическое семейство совместных распределений $f(\boldsymbol{x},y\,|\,\boldsymbol{a},b,c)\propto\exp\Big\{-cq\big(y,z(\boldsymbol{x},\boldsymbol{a},b)\big)\Big\}$.

2) Чем меньше регуляризующая функция $V(\boldsymbol{a}|\mu)$, тем более априори вероятно значение параметра $\boldsymbol{a}\in\mathbb{R}^n$.

Предполагаемое параметрическое семейство априорных распределений $\psi(\boldsymbol{a}|\mu)\propto\exp\big\{-c'V(\boldsymbol{a}|\mu)\big\}$.

Байесовский принцип обучения: максимизации апостериорной плотности на обучающей совокупности

$$(\boldsymbol{X},\boldsymbol{Y})=\Big\{\big(\boldsymbol{x}(\omega_j),y(\omega_j)\big)=(\boldsymbol{x}_j,y_j),\ j=1,...,N\Big\}$$

$$P(\boldsymbol{a},b\,|\,\boldsymbol{X},\boldsymbol{Y},\mu,c)\propto$$

$$\psi(\boldsymbol{a}|\mu)\prod_{j=1}^{N}f(\boldsymbol{x}_j,y_j\,|\,\boldsymbol{a},b,c)\to\max(\boldsymbol{a},b).$$

Вероятностная интерпретация минимизации регуляризованного эмпирического риска

Обучение: Минимизация регуляризованного эмпирического риска на обучающей совокупности

$$(X,Y)=\Big\{\big(x(\omega_j),y(\omega_j)\big)=(x_j,y_j),\ j=1,...,N\Big\}$$

$$V(\boldsymbol{a}\,|\,\mu)+c\sum_{j=1}^{N}q\big(y_j,z(\boldsymbol{x}_j,\boldsymbol{a},b)\big)\to\min(\boldsymbol{a},b),$$

$$z(\boldsymbol{x}_j,\boldsymbol{a},b)=K(\boldsymbol{x}_j,\boldsymbol{a})+b$$

Регуляризующая функция $V(\boldsymbol{a}|\mu)$ и функция связи $q(y,z)$ выпуклы по $\boldsymbol{a}\in\mathbb{R}^n$ и $z\in\mathbb{R}$.

Вероятностные эвристики наблюдателя

1) Чем меньше параметрическая функция потерь $q\big(y,z(\boldsymbol{x},\boldsymbol{a},b)\big)$, тем более вероятно появление объекта с характеристиками $(\boldsymbol{x},y)$.

Предполагаемое параметрическое семейство совместных распределений $f(\boldsymbol{x},y\,|\,\boldsymbol{a},b,c)\propto\exp\Big\{-cq\big(y,z(\boldsymbol{x},\boldsymbol{a},b)\big)\Big\}$.

2) Чем меньше регуляризующая функция $V(\boldsymbol{a}|\mu)$, тем более априори вероятно значение параметра $\boldsymbol{a}\in\mathbb{R}^n$.

Предполагаемое параметрическое семейство априорных распределений $\psi(\boldsymbol{a}|\mu)\propto\exp\big\{-c'V(\boldsymbol{a}|\mu)\big\}$.

Выбор значений структурных параметров:
Максимизация полной функции правдоподобия (Evidence Function) на обучающей совокупности

$$F(\boldsymbol{X},\boldsymbol{Y}\,|\,\mu,c)=\int_{\mathbb{R}}\int_{\mathbb{R}^n}\left[\psi(\boldsymbol{a}|\mu)\prod_{j=1}^{N}f(\boldsymbol{x}_j,y_j\,|\,\boldsymbol{a},b,c)\right]d\boldsymbol{a}db\to\max(\mu,c).$$