Методы погружения произвольных объектов реального мира в линейное пространство для реализации обобщенного линейного подхода к восстановлению зависимостей

Середин Олег Сергеевич
Тульский государственный университет

Моттль Вадим Вячеславович
Вычислительный центр РАН

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 $(c' + c'')\mathbf{x} = c'\mathbf{x} + c''\mathbf{x}$, $c(\mathbf{x}' + \mathbf{x}'') = c\mathbf{x}' + c\mathbf{x}''$.

Обобщенный линейный подход

Линейное пространство восприятия объектов реального мира

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

– $\boldsymbol{x}'+\boldsymbol{x}''=\boldsymbol{x}''+\boldsymbol{x}',\quad (\boldsymbol{x}'+\boldsymbol{x}'')+\boldsymbol{x}'''=\boldsymbol{x}'+(\boldsymbol{x}''+\boldsymbol{x}''')$;

– нулевой элемент $\boldsymbol{x}+\boldsymbol{\phi}=\boldsymbol{x},\ c\boldsymbol{\phi}=\boldsymbol{\phi},\quad (-\boldsymbol{x})+\boldsymbol{x}=\boldsymbol{\phi}$;

– $c'(c''\boldsymbol{x})=(c'c'')\boldsymbol{x},\quad 1\boldsymbol{x}=\boldsymbol{x},\quad 0\boldsymbol{x}=\boldsymbol{\phi}$;

– $(c'+c'')\boldsymbol{x}=c'\boldsymbol{x}+c''\boldsymbol{x},\ c(\boldsymbol{x}'+\boldsymbol{x}'')=c\boldsymbol{x}'+c\boldsymbol{x}''$.

Индефинитное скалярное произведение

$\boldsymbol{x},\boldsymbol{\upsilon}\in\mathbb{X}$ – две произвольные точки в линейном пространстве,

$K(\boldsymbol{x},\boldsymbol{\upsilon}):\mathbb{X}\times\mathbb{X}\to\mathbb{R}$ – двухместая скалярная функция:

(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. Псевдоевклидово линейное пространство. Нормы нет.

Обобщенная линейная модель восстановления зависимости

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

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

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

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

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

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

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

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

Критерий выпуклый, если регуляризующая функция $V(\boldsymbol{v} \,|\, \mu)$ и функция связи $q(y, z)$ выпуклы по $\boldsymbol{v} \in \mathbb{X}$ и $z \in \mathbb{R}$.

Погружение множества объектов в линейное пространство

1. Произвольная функция парного сравнения объектов

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

Симметричная функция парного сравнения $S(\omega', \omega'') = S(\omega'', \omega'):\ \Omega \times \Omega \to \mathbb{R}$

Центральный элемент $\phi \in \Omega$

Симметричная функция общности для центра $\phi \in \Omega$:

$$K_\phi(\omega', \omega'') = \frac{1}{2}\big[S(\omega', \phi) + S(\omega'', \phi) - S(\omega', \omega'')\big]:\ \ \Omega \times \Omega \to \mathbb{R}$$

Пусть множество объектов конечно $|\Omega| = M$ – скорее всего, ***ОЧЕНЬ*** большое число!

Симметрическая матрица общности объектов для центра $\phi \in \Omega$:

$$\mathbf{K}_\phi = \begin{pmatrix} K_\phi(\omega_1, \omega_1) & \cdots & K_\phi(\omega_1, \omega_M) \\ \vdots & \ddots & \vdots \\ K_\phi(\omega_M, \omega_1) & \cdots & K_\phi(\omega_M, \omega_M) \end{pmatrix},$$

собственные числа действительны $\xi_{\phi,i} \in \mathbb{R},\ i = 1, \ldots, M$,

собственные векторы ортонормированы $\mathbf{z}_{\phi,i} \in \mathbb{R}^M,\ \mathbf{z}^T_{\phi,i}\mathbf{z}_{\phi,j} = \begin{cases} 1,\ i = j, \\ 0,\ i \neq j. \end{cases}$

Собственные числа в порядке убывания $\underbrace{\xi_{\phi,1} \geq 0, \ldots, \xi_{\phi,p_\phi} \geq 0}_{\textit{положительные}},\ \underbrace{\xi_{\phi,p_\phi+1} < 0, \ldots, \xi_{\phi,M} < 0}_{\textit{отрицательные}}$

Пара целых чисел $p_\phi + q_\phi = M$ – сигнатура матрицы

Теорема: Сигнатура матрицы $\mathbf{K}_\phi$ не зависит от выбора центра $\phi \in \Omega$.

Матрица $\mathbf{K}_\phi = \sum_{i=1}^{p} \xi_{\phi,i}\mathbf{z}_{\phi,i}\mathbf{z}^T_{\phi,i} - \sum_{i=p+1}^{M} \bar{\xi}_{\phi,i}\mathbf{z}_{\phi,i}\mathbf{z}^T_{\phi,i}\ (M \times M)$ – ***ОЧЕНЬ*** большая

Погружение в линейное пространство:

Погружение множества объектов в линейное пространство

1. Произвольная функция парного сравнения объектов

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

Симметричная функция парного сравнения $S(\omega', \omega'') = S(\omega'', \omega'): \ \Omega \times \Omega \to \mathbb{R}$

Центральный элемент $\phi \in \Omega$

Симметричная функция общности для центра $\phi \in \Omega$:

$$K_\phi(\omega', \omega'') = \frac{1}{2}\left[S(\omega', \phi) + S(\omega'', \phi) - S(\omega', \omega'')\right]: \ \ \Omega \times \Omega \to \mathbb{R}$$

Пусть множество объектов конечно $|\Omega| = M$ – скорее всего, ***ОЧЕНЬ*** большое число!

Пара целых чисел $p_\phi + q_\phi = M$ – сигнатура матрицы не зависит от центра $\phi \in \Omega$.

Матрица $\mathbf{K}_\phi = \sum_{i=1}^{p} \xi_{\phi,i} \mathbf{z}_{\phi,i} \mathbf{z}_{\phi,i}^T - \sum_{i=p+1}^{M} \bar{\xi}_{\phi,i} \mathbf{z}_{\phi,i} \mathbf{z}_{\phi,i}^T$ $(M \times M)$ – ***ОЧЕНЬ*** большая

Представим матрицу $\mathbf{K}_\phi$ как совокупность скалярных произведений

$$\mathbf{K}_\phi = \begin{pmatrix} \boldsymbol{x}_{\phi,1}^T \mathbf{J}_p \boldsymbol{x}_{\phi,1} & \cdots & \boldsymbol{x}_{\phi,1}^T \mathbf{J}_p \boldsymbol{x}_{\phi,M} \\ \vdots & \ddots & \vdots \\ \boldsymbol{x}_{\phi,M}^T \mathbf{J}_p \boldsymbol{x}_{\phi,1} & \cdots & \boldsymbol{x}_{\phi,M}^T \mathbf{J}_p \boldsymbol{x}_{\phi,M} \end{pmatrix}, \ \mathbf{J}_p = \begin{pmatrix} \mathbf{I}_{p \times p} & \mathbf{0}_{p \times (M-p)} \\ \mathbf{0}_{(M-p) \times p} & -\mathbf{I}_{(M-p) \times (M-p)} \end{pmatrix}$$

– единичная матрица сигнатуры p

Мы связали элементы произвольного конечного множества $\Omega = \{\omega_1, ..., \omega_M\}$, $\phi \in \Omega$ – центр, с M-мерными векторами действительных признаков объектов $\boldsymbol{x}_{\phi,1} = \boldsymbol{x}_{\phi,\omega_1} \in \mathbb{R}^M, ..., \boldsymbol{x}_{\phi,M} = \boldsymbol{x}_{\phi,\omega_M} \in \mathbb{R}^M$. Центральный объект – вектор $\boldsymbol{x}_\phi \in \mathbb{R}^M$.

Погружение множества объектов в линейное пространство

1. Произвольная функция парного сравнения объектов

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

Симметричная функция парного сравнения $S(\omega', \omega'') = S(\omega'', \omega'): \ \Omega \times \Omega \to \mathbb{R}$

Центральный элемент $\phi \in \Omega$

Симметричная функция общности для центра $\phi \in \Omega$:

$$K_{\phi}(\omega', \omega'') = \frac{1}{2}\big[S(\omega', \phi) + S(\omega'', \phi) - S(\omega', \omega'')\big]: \ \ \Omega \times \Omega \to \mathbb{R}$$

Пусть множество объектов конечно $|\Omega| = M$ – скорее всего, ***ОЧЕНЬ*** большое число!

Мы связали элементы произвольного конечного множества $\Omega = \{\omega_1, ..., \omega_M\}$,

$\phi \in \Omega$ – центр, с M -мерными векторами действительных признаков объектов

$\boldsymbol{x}_{\phi,1} = \boldsymbol{x}_{\phi,\omega_1} \in \mathbb{R}^M, ..., \boldsymbol{x}_{\phi,M} = \boldsymbol{x}_{\phi,\omega_M} \in \mathbb{R}^M$. Центральный объект – вектор $\boldsymbol{x}_{\phi} \in \mathbb{R}^M$.

Погружение в линейное пространство, с двухместной функцией

$K(\boldsymbol{x}', \boldsymbol{x}'') = \boldsymbol{x}'^T \mathbf{J}_p \boldsymbol{x}'': \ \mathbb{R}^M \times \mathbb{R}^M \to \mathbb{R}.$ Ее свойства:

Симметричность $K(\boldsymbol{x}', \boldsymbol{x}'') = K(\boldsymbol{x}'', \boldsymbol{x}')$,

билинейность $K(c'\boldsymbol{x}' + c''\boldsymbol{x}'', \boldsymbol{x}''') = c'K(\boldsymbol{x}', \boldsymbol{x}''') + c''K(\boldsymbol{x}'', \boldsymbol{x}''')$.

Нет свойства $K(\boldsymbol{x}, \boldsymbol{x}) \geq 0$.

Это индефинитное скалярное произведение.

Псевдоевклидово линейное пространство $\mathbb{R}^M$, натянутое на метрическое пространство $\Omega = \{\omega_1, ..., \omega_M\}$. Нормы нет.

Погружение множества объектов в линейное пространство

1. Произвольная функция парного сравнения объектов

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

Симметричная функция парного сравнения $S(\omega', \omega'') = S(\omega'', \omega') : \ \Omega \times \Omega \to \mathbb{R}$

Центральный элемент $\phi \in \Omega$

Симметричная функция общности для центра $\phi \in \Omega$:

$$K_\phi(\omega', \omega'') = \frac{1}{2}\big[S(\omega', \phi) + S(\omega'', \phi) - S(\omega', \omega'')\big] : \ \ \Omega \times \Omega \to \mathbb{R}$$

Пусть множество объектов конечно $|\Omega| = M$ – скорее всего, ***ОЧЕНЬ*** большое число!

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

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

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

Как варьировать направляющий вектор в необозримом пространстве $\boldsymbol{v} \in \mathbb{X}$?

Базисная совокупность объектов $\{\omega_1^0, \dots, \omega_n^0\} \subset \Omega$, их образы $\{\boldsymbol{x}_\phi(\omega_1^0), \dots, \boldsymbol{x}_\phi(\omega_n^0)\}$.

Направляющий вектор – линейная комбинация образов базисных объектов $\quad \boldsymbol{v}(\boldsymbol{a}) = \sum_{i=1}^n a_i \boldsymbol{x}_\phi(\omega_i^0), \ \sum_{i=1}^n a_i = 0.$

Теорема: Обобщенный линейный признак объекта $\quad z(\omega, \boldsymbol{a}, b) = -\frac{1}{2}\sum_{i=1}^n a_i S(\omega, \omega_i^0) + b, \ \sum_{i=1}^n a_i = 0.$

Погружение объектов в линейное пространство является лишь мысленным.

Погружение множества объектов в линейное пространство

2. Двухместная функция расстояния на множестве объектов

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

По-прежнему симметричная функция сравнения $S(\omega', \omega'') = S(\omega'', \omega'):\ \Omega \times \Omega \to \mathbb{R}$.

Дополнительные требования:

нулевое отношение с самим собой $S(\omega, \omega) = 0$, неотрицательность $S(\omega', \omega'') \geq 0$.

Расстояние будем понимать как $d(\omega', \omega'') = \sqrt{S(\omega', \omega'')}$.

Центральный элемент $\phi \in \Omega$.

Симметричная функция общности для центра $\phi \in \Omega$:

$$K_\phi(\omega', \omega'') = \frac{1}{2}\Big[d^2(\omega', \phi) + d^2(\omega'', \phi) - d^2(\omega', \omega'') \Big]:\ \ \Omega \times \Omega \to \mathbb{R}$$

Пусть множество объектов конечно $|\Omega| = M$ – скорее всего, ***ОЧЕНЬ*** большое число!

В теоретических построениях ничего не изменяется.

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

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

Базисная совокупность объектов $\{\omega_1^0, \ldots, \omega_n^0\} \subset \Omega$, их образы $\{\boldsymbol{x}_\phi(\omega_1^0), \ldots, \boldsymbol{x}_\phi(\omega_n^0)\}$.

Теорема: Обобщенный линейный признак объекта
$$z(\omega, \boldsymbol{a}, b) = -\frac{1}{2}\sum_{i=1}^n a_i d^2(\omega, \omega_i^0) + b,\ \sum_{i=1}^n a_i = 0.$$

Погружение объектов в линейное пространство является лишь мысленным.

Погружение множества объектов в линейное пространство

3. Двухместная функция сравнения является метрикой

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

По-прежнему функция сравнения $S(\omega', \omega'') = S(\omega'', \omega')$: $\Omega \times \Omega \to \mathbb{R}$.

Дополнительные требования:

$S(\omega, \omega) = 0$, неравенство треугольника $S(\omega', \omega'') + S(\omega'', \omega''') \geq S(\omega', \omega''')$

Расстояние будем понимать как $d(\omega', \omega'') = \sqrt{S(\omega', \omega'')}$.

Центральный элемент $\phi \in \Omega$

Симметричная функция общности для центра $\phi \in \Omega$:

$$K_\phi(\omega', \omega'') = \frac{1}{2}\left[d^2(\omega', \phi) + d^2(\omega'', \phi) - d^2(\omega', \omega'')\right]: \quad \Omega \times \Omega \to \mathbb{R}$$

Пусть множество объектов конечно $|\Omega| = M$ – скорее всего, ***ОЧЕНЬ*** большое число!

В теоретических построениях ничего не изменяется.

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

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

Базисная совокупность объектов $\{\omega_1^0, ..., \omega_n^0\} \subset \Omega$, их образы $\{\boldsymbol{x}_\phi(\omega_1^0), ..., \boldsymbol{x}_\phi(\omega_n^0)\}$.

Теорема: Обобщенный линейный признак объекта
$$z(\omega, \boldsymbol{a}, b) = -\frac{1}{2}\sum_{i=1}^{n} a_i d^2(\omega, \omega_i^0) + b, \; \sum_{i=1}^{n} a_i = 0.$$

Погружение объектов в линейное пространство является лишь мысленным.

Критерий обучения: Минимум регуляризованного эмпирического риска

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

Базисная совокупность $\{\omega_i^0, i = 1, \ldots, n\} \subset \Omega$	Нет значений целевой характеристики, есть только матрица парных сравнений $\left[S(\omega_i, \omega_k), \ i, k = 1, \ldots, n\right]$
Обучающая совокупность (часть базисной совокупности) $\{\omega_j, j = 1, \ldots, N\} \subset \Omega$	Известны значения целевой характеристики $y(\omega_j) \in \mathbb{Y}$ и матрица парных сравнений $\left[S(\omega_j, \omega_l), \ \ j, l = 1, \ldots, N\right]$

Мысленное погружение множества объектов в воображаемое линейное пространство со скалярным произведением, вообще говоря, индефинитным:
$\boldsymbol{x}(\omega): \ \Omega \to \mathbb{X}, \ \ K(\boldsymbol{x}', \boldsymbol{x}''): \ \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ – автоматически следуют из $S(\omega', \omega'')$.

Обобщенный линейный признак объекта $z(\boldsymbol{x}_j, \boldsymbol{v}, b) = K(\boldsymbol{x}_j, \boldsymbol{v}) + b$,

$\boldsymbol{v} \in \mathbb{X}$ – искомый направляющий вектор, $b \in \mathbb{R}$ – искомый сдвиг

Функция связи, выбираемая наблюдателем $q(y, z): \ \mathbb{Y} \times \mathbb{R} \to \mathbb{R}^+$, выпуклая по z.

Регуляризующая функция, выбираемая наблюдателем $V(\boldsymbol{v} \,|\, \mu): \ \mathbb{X} \to \mathbb{R}^+$

Обучение – выбор $(\boldsymbol{v} \in \mathbb{X}, b \in \mathbb{R})$ по критерию минимума регуляризованного эмпирического риска	$V(\boldsymbol{v} \,|\, \mu) + c \sum_{j=1}^{N} q\left(y_j, z(\boldsymbol{x}_j, \boldsymbol{v}, b)\right) \to \min(\boldsymbol{v}, b)$

Параметрическое представление направляющего вектора $\boldsymbol{v}(\boldsymbol{a}) = \sum_{i=1}^{n} a_i \boldsymbol{x}_\phi(\omega_i^0), \ \sum_{i=1}^{n} a_i = 0$.

Критерий обучения: Минимум регуляризованного эмпирического риска

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

Базисная совокупность $\{\omega_i^0, i = 1, \ldots, n\} \subset \Omega$	Нет значений целевой характеристики, есть только матрица парных сравнений $\left[S(\omega_i, \omega_k), \ i, k = 1, \ldots, n \right]$
Обучающая совокупность (часть базисной совокупности) $\{\omega_j, \ j = 1, \ldots, N\} \subset \Omega$	Известны значения целевой характеристики $y(\omega_j) \in \mathbb{Y}$ и матрица парных сравнений $\left[S(\omega_j, \omega_l), \ \ j, l = 1, \ldots, N \right]$

Функция связи, выбираемая наблюдателем $q(y, z): \ \mathbb{Y} \times \mathbb{R} \to \mathbb{R}^+$, выпуклая по z.

Регуляризующая функция, выбираемая наблюдателем $V(\boldsymbol{\vartheta} \mid \mu): \ \mathbb{X} \to \mathbb{R}^+$

Обучение: Выбор $\left(\boldsymbol{\vartheta} \in \mathbb{X}, b \in \mathbb{R} \right)$ по критерию минимума регуляризованного эмпирического риска	$V(\boldsymbol{\vartheta} \mid \mu) + c \sum_{j=1}^{N} q\left(y_j, z(\boldsymbol{x}_j, \boldsymbol{\vartheta}, b) \right) \to \min(\boldsymbol{\vartheta}, b)$

Параметрическое представление направляющего вектора $\boldsymbol{\vartheta}(\boldsymbol{a}) = \sum_{i=1}^{n} a_i \boldsymbol{x}_\phi(\omega_i^0), \ \sum_{i=1}^{n} a_i = 0.$

Выпуклая параметрическая регуляризующая функция $V(\boldsymbol{a} \mid \mu): \ \mathbb{R}^n \to \mathbb{R}^+$

Параметрическое обучение: Выбор $\left(\boldsymbol{a} \in \mathbb{R}^n, b \in \mathbb{R} \right)$,	$V(\boldsymbol{a} \mid \mu) + c \sum_{j=1}^{N} q\left(y_j, z(\omega_j, \boldsymbol{a}, b) \right) + b \to \min(\boldsymbol{a}, b),$ $z(\omega_j, \boldsymbol{a}, b) = \sum_{i=1}^{n} \left(S(\omega, \omega_i^0) \right) a_i + b,$ выпуклый критерий.

Отбор подмножества базисных объектов: Квадратично-модульная регуляризация

Регуляризация Elastic Net: $V(\boldsymbol{a}\,|\,\mu) = \sum_{i=1}^{n} a_i^2 + \mu \sum_{i=1}^{n} |a_i|$, $\mu \ge 0$ – параметр селективности

Критерий обучения: $\sum_{i=1}^{n} a_i^2 + \mu \sum_{i=1}^{n} |a_i| + \sum_{j=1}^{N} q\Big(y_j, z(\omega_j, \boldsymbol{a}, b)\Big)$, $z(\omega_j, \boldsymbol{a}, b) = \sum_{i=1}^{n} \Big(S(\omega_j, \omega_i^0)\Big) a_i + b$.

Теорема. Пусть значения $(\hat{\lambda}_1, ..., \hat{\lambda}_N)$ являются решением двойственной задачи выпуклого программирования:

$$\begin{cases} \dfrac{1}{2\beta} \sum_{i=1}^{n} \left\{ \min\left[\mu/2 + \sum_{j=1}^{N} \lambda_j x_{ij},\ 0,\ \mu/2 - \sum_{j=1}^{N} \lambda_j x_{ij} \right] \right\}^2 - \sum_{j=1}^{N} \underbrace{\min_{z \in \mathbb{R}} \Big(q(y_j, z) + \lambda_j z \Big)}_{\text{вогнута по } \lambda} \to \min(\lambda_1, ..., \lambda_N), \\ \sum_{j=1}^{N} \lambda_j = 0. \end{cases}$$

Тогда

$$\hat{a}_i = \begin{cases} \Big(\sum_{j=1}^{N} \hat{\lambda}_j S(\omega_j, \omega_i^0) + \mu/2 \Big) < 0, & \sum_{j=1}^{N} \hat{\lambda}_j S(\omega_j, \omega_i^0) < -\mu/2, \\ 0, & -\mu/2 \le \sum_{j=1}^{N} \hat{\lambda}_j S(\omega_j, \omega_i^0) \le \mu/2, \\ \Big(\sum_{j=1}^{N} \hat{\lambda}_j S(\omega_j, \omega_i^0) - \mu/2 \Big) > 0, & \sum_{j=1}^{N} \hat{\lambda}_j S(\omega_j, \omega_i^0) > \mu/2, \end{cases} \qquad \begin{cases} \hat{b} = \dfrac{1}{N} \sum_{j=1}^{N} \Big(\hat{z}_j - \sum_{i=1}^{n} \hat{a}_i x_{ij} \Big), \text{ где} \\ \hat{z}_j = \arg\min_{z \in \mathbb{R}} \Big(q(y_j, z) + \hat{\lambda}_j z \Big). \end{cases}$$

Чем больше $\mu \ge 0$, тем выше селективность отбора базисных объектов.

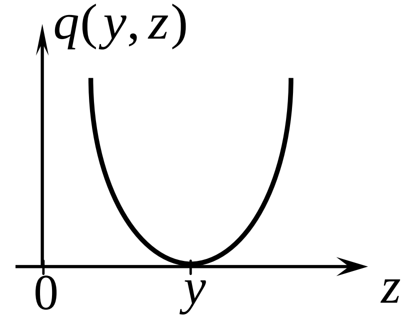
Функции связи для некоторых частных видов целевой переменной

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

1) Задача оценивания числовой регрессии.

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

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

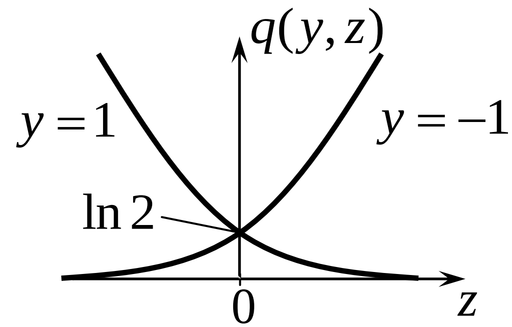


2) Задача обучения распознаванию объектов двух классов.

Целевая переменная $y \in \mathbb{Y} = \{-1, 1\}$ – индекс класса.

Метод логистической регрессии:

$$q(y,z) = \ln\left[1+\exp(-yz)\right].$$

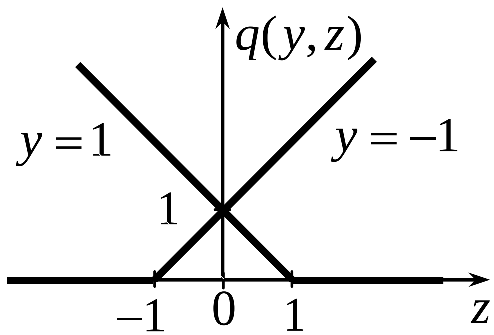


3) Задача обучения распознаванию объектов двух классов.

Целевая переменная $y \in \mathbb{Y} = \{-1, 1\}$ – индекс класса.

Метод опорных векторов:

$$q(y,z) = \max\left[0,\ 1-yz\right] = \begin{cases} 1-yz, & 1-yz>0, \\ 0, & 1-yz \le 0. \end{cases}$$



IDP-2016

BARCELONA

Спасибо за внимание!