IDP-2016

BARCELONA

Methods of embedding real-world entities into a linear space for implementing the generalized linear approach to dependence estimation

Oleg S. Seredin
Tula State University, Tula

Vadim V. Mottl
Computing Center of the Russian Academy of Sciences, Moscow

Laboratory of Data Analysis

The task of dependences estimation on the sets of real-world objects

A set of real-world objects $\omega \in \Omega$.

A set of hidden characteristic values $y \in \mathbb{Y}$.

Real existing hidden function $y(\omega): \Omega \to \mathbb{Y}$.

Observer desire:

To have a tool for estimating the hidden characteristic for the real objects

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

Generalized linear approach

Observer (his/her computer) interpret a real-world objects as points in some linear space:
$\mathbf{x}(\omega): \Omega \to \mathbb{X}$ – linear space.

- addition is commutative $\mathbf{x}' + \mathbf{x}'' = \mathbf{x}'' + \mathbf{x}'$
  and associative $(\mathbf{x}' + \mathbf{x}'') + \mathbf{x}''' = \mathbf{x}' + (\mathbf{x}'' + \mathbf{x}''')$;
- identity element of addition $\mathbf{x} + \boldsymbol{\phi} = \mathbf{x}$, $c\boldsymbol{\phi} = \boldsymbol{\phi}$;
- inverse element of addition $(-\mathbf{x}) + \mathbf{x} = \boldsymbol{\phi}$;
- compatibility of scalar multiplication with field multiplication $c'(c''\mathbf{x}) = (c'c'')\mathbf{x}$;
- identity element of scalar multiplication $1\mathbf{x} = \mathbf{x}$;
- distributivity of multiplication and addition $(c' + c'')\mathbf{x} = c'\mathbf{x} + c''\mathbf{x}$, $c(\mathbf{x}' + \mathbf{x}'') = c\mathbf{x}' + c\mathbf{x}''$.

Generalized Linear Approach

Linear space of real-world object perception

Observer (his/her computer) interpret a real-world objects as points in some linear space:
$\boldsymbol{x}(\omega): \Omega \to \mathbb{X}$ – linear space.

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

– zero element $\boldsymbol{x} + \boldsymbol{\phi} = \boldsymbol{x}$, $c\boldsymbol{\phi} = \boldsymbol{\phi}$, $(-\boldsymbol{x}) + \boldsymbol{x} = \boldsymbol{\phi}$;

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

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

Indefinite scalar product

$\boldsymbol{x}, \boldsymbol{\upsilon} \in \mathbb{X}$ – two arbitrary points in linear space,
$K(\boldsymbol{x}, \boldsymbol{\upsilon}): \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ – scalar function of two arguments,
(1) $K(\boldsymbol{x}, \boldsymbol{\upsilon}) = K(\boldsymbol{\upsilon}, \boldsymbol{x})$ – symmetry,
(2) $K(\boldsymbol{x},\ c'\boldsymbol{\upsilon}' + c''\boldsymbol{\upsilon}'') = c'K(\boldsymbol{x}, \boldsymbol{\upsilon}') + c''K(\boldsymbol{x}, \boldsymbol{\upsilon}'')$ – bilinearity.

If suggest
(3) $K(\boldsymbol{x}, \boldsymbol{x}) \geq 0$, then it will be common inner product, so $\sqrt{K(\boldsymbol{x}, \boldsymbol{x})} = \|\boldsymbol{x}\|$ – is a norm.

For our approach it is enough just (1) and (2).
It is so-called indefinite scalar product in the pseudo-Euclidean space. There is no norm.

General Linear Model of Dependency

The object in linear space $\boxed{\boldsymbol{x}(\omega)\in\mathbb{X} \text{ – just observer imagination}}$

Inner product (indefinite) $\boxed{K(\boldsymbol{x},\boldsymbol{v}):\mathbb{X}\times\mathbb{X}\rightarrow\mathbb{R} \text{ – just observer imagination}}$

Parameters of model of dependency $(\boldsymbol{v},b)$
$$\begin{cases}\boldsymbol{v}\in\mathbb{X} \text{ – directional point (vector) in the same space}\\ b\in\mathbb{R} \text{ – model shift (intercept)}\end{cases}$$

Generalized linear feature of object $z(\boldsymbol{x},\boldsymbol{v},b)=K(\boldsymbol{x},\boldsymbol{v})+b$: $\mathbb{X}\xrightarrow{(\boldsymbol{v},b)}\mathbb{R}$

Goal characteristic of object $y(\omega)\in\mathbb{Y}$ – given by nature.

Link function,
Usually convex on z $\boxed{q(y,z):\mathbb{Y}\times\mathbb{R}\rightarrow\mathbb{R}^{+} \text{ – just observer imagination}}$

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

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

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

Family of convex regularized functions $\boxed{V(\boldsymbol{v}\,|\,\mu):\ \mathbb{X}\xrightarrow{\mu}\mathbb{R}^{+} \text{ – just observer imagination}}$

Training process – find $(\boldsymbol{v}\in\mathbb{X},b\in\mathbb{R})$:
Minimization of regularized empirical risk

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

The criterion is convex if regularized function $V(\boldsymbol{v}\,|\,\mu)$ and link function $q(y,z)$ are convex on $\boldsymbol{v}\in\mathbb{X}$ and $z\in\mathbb{R}$.

Objects Embedding into a Linear Space

1. Arbitrary pairwise function of objects comparing

Set of real-world objects $\omega \in \Omega$

The pairwise comparing function in symmetric $S(\omega', \omega'') = S(\omega'', \omega'):\ \Omega \times \Omega \to \mathbb{R}$

Let us choose an arbitrary element as its "center" $\phi \in \Omega$

Two-argument commonality symmetric function for the center $\phi \in \Omega$:

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

Let us assume, for simplicity sake, that the set of real-world objects finite $|\Omega| = M$ (probably, a VERY huge number!)

The symmetric commonality matrix for some center $\phi \in \Omega$:

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

eigen values are real numbers $\xi_{\phi,i} \in \mathbb{R},\ i = 1, \dots, M$,

eigen vectors are orthonormal $\mathbf{z}_{\phi,i} \in \mathbb{R}^M,\ \mathbf{z}_{\phi,i}^T \mathbf{z}_{\phi,j} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$

Eigen values in decrease order $\underbrace{\xi_{\phi,1} \geq 0, \dots, \xi_{\phi,p_\phi} \geq 0}_{positive},\ \underbrace{\xi_{\phi,p_\phi+1} < 0, \dots, \xi_{\phi,M} < 0}_{negative}$

The pair of integers $p_\phi + q_\phi = M$ – signature of commonality matrix

Theorem: *The signature of matrix* $\mathbf{K}_\phi$ *does not depend on the choice of the center* $\phi \in \Omega$.

Matrix $\mathbf{K}_\phi = \sum_{i=1}^{p} \xi_{\phi,i} \mathbf{z}_{\phi,i} \mathbf{z}_{\phi,i}^T - \sum_{i=p+1}^{M} \bar{\xi}_{\phi,i} \mathbf{z}_{\phi,i} \mathbf{z}_{\phi,i}^T\ (M \times M)$ – VERY huge!

Objects Embedding into a Linear Space

1. Arbitrary pairwise function of objects comparing

Set of real-world objects $\omega \in \Omega$

The pairwise comparing function in symmetric $S(\omega', \omega'') = S(\omega'', \omega')$: $\Omega \times \Omega \to \mathbb{R}$

Let us choose an arbitrary element as its “center” $\phi \in \Omega$

Two-argument commonality symmetric function for the center $\phi \in \Omega$:

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

Let the set of real-world objects finite $|\Omega| = M$ (probably, a VERY huge number!)

The pair of integers $p_\phi + q_\phi = M$ – signature (*does not depend on center*)

Matrix $\mathbf{K}_\phi = \sum_{i=1}^{p} \xi_{\phi,i} \mathbf{z}_{\phi,i} \mathbf{z}_{\phi,i}^T - \sum_{i=p+1}^{M} \bar{\xi}_{\phi,i} \mathbf{z}_{\phi,i} \mathbf{z}_{\phi,i}^T$ $(M \times M)$ – is very huge**.**

Let assume $\mathbf{K}_\phi$ as a set of inner products

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

So, we associate the elements of arbitrary finite set $\Omega = \{\omega_1, \ldots, \omega_M\}$,

$\phi \in \Omega$ – center, with M-dimensional vectors of real features of objects

$\boldsymbol{x}_{\phi,1} = \boldsymbol{x}_{\phi,\omega_1} \in \mathbb{R}^M, \ldots, \boldsymbol{x}_{\phi,M} = \boldsymbol{x}_{\phi,\omega_M} \in \mathbb{R}^M$. Zero element – vector of $\boldsymbol{x}_\phi \in \mathbb{R}^M$.

Objects Embedding into a Linear Space

1. Arbitrary pairwise function of objects comparing

Set of real-world objects $\omega \in \Omega$

The pairwise comparing function in symmetric $S(\omega', \omega'') = S(\omega'', \omega')$: $\Omega \times \Omega \to \mathbb{R}$

Let us choose an arbitrary element as its "center" $\phi \in \Omega$

Two-argument commonality symmetric function for the center $\phi \in \Omega$:

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

Let the set of real-world objects finite $|\Omega| = M$ (probably, a VERY huge number!)

So, we associate the elements of arbitrary finite set $\Omega = \{\omega_1, ..., \omega_M\}$,

$\phi \in \Omega$ – center, with M -dimensional vectors of real features of objects

$\boldsymbol{x}_{\phi,1} = \boldsymbol{x}_{\phi,\omega_1} \in \mathbb{R}^M, ..., \boldsymbol{x}_{\phi,M} = \boldsymbol{x}_{\phi,\omega_M} \in \mathbb{R}^M$. Zero element – vector of $\boldsymbol{x}_\phi \in \mathbb{R}^M$.

Embedding in to linear space, with two-argument function

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

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

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

Not holds $K(\boldsymbol{x}, \boldsymbol{x}) \geq 0$.

So it is indefinite inner products.

Pseudo-Euclidean linear space $\mathbb{R}^M$ based on $\Omega = \{\omega_1, ..., \omega_M\}$. There is no norm.

Objects Embedding into a Linear Space

1. Arbitrary pairwise function of objects comparing

Set of real-world objects $\omega \in \Omega$

The pairwise comparing function in symmetric $S(\omega', \omega'') = S(\omega'', \omega'): \ \Omega \times \Omega \to \mathbb{R}$

Let us choose an arbitrary element as its “center” $\phi \in \Omega$

Two-argument commonality symmetric function for the center $\phi \in \Omega$:

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

Let the set of real-world objects finite $|\Omega| = M$ (probably, a VERY huge number!)

Inner product (indefinite) $K(\boldsymbol{x}, \boldsymbol{v}): \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ – just observer imagination

Parameters of model of dependency $(\boldsymbol{v}, b)$ $\begin{cases} \boldsymbol{v} \in \mathbb{X} & \text{– directional point (vector) in the same space} \\ b \in \mathbb{R} & \text{– model shift (intercept)} \end{cases}$

Generalized linear characteristic of object $z(\boldsymbol{x}, \boldsymbol{v}, b) = K(\boldsymbol{x}, \boldsymbol{v}) + b: \ \mathbb{X} \xrightarrow{(\boldsymbol{v}, b)} \mathbb{R}$

How we can search the directional vector in imaginary space $\boldsymbol{v} \in \mathbb{X}$?

Basic set of the objects $\{\omega_1^0, ..., \omega_n^0\} \subset \Omega$, and their mapping image $\{\boldsymbol{x}_\phi(\omega_1^0), ..., \boldsymbol{x}_\phi(\omega_n^0)\}$.

Directional vector – linear combination of images (fancies) of basic objects

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

Theorem: Generalized linear characteristic of object

$$z(\omega, \boldsymbol{a}, b) = -\frac{1}{2}\sum_{i=1}^{n} a_i S(\omega, \omega_i^0) + b, \ \sum_{i=1}^{n} a_i = 0.$$

So, the embedding into linear space is just our fantasy!

Objects Embedding into a Linear Space

2. Distance pairwise function of objects comparing

Set of real-world objects $\omega \in \Omega$.

Now as before, the comparison function $S(\omega', \omega'') = S(\omega'', \omega'): \ \Omega \times \Omega \to \mathbb{R}$ **symmetric.**

Additional requirements:

non-negativity $S(\omega', \omega'') \geq 0$, zero value for the same argument $S(\omega, \omega) = 0$.

Let us note the distance as $d(\omega', \omega'') = \sqrt{S(\omega', \omega'')}$.

Central element $\phi \in \Omega$.

Two-argument commonality symmetric function for the center $\phi \in \Omega$:

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

Let the set of real-world objects finite $|\Omega| = M$ (probably, a VERY huge number!)

There is no change in theoretical speculations.

Inner product (indefinite) $K(\boldsymbol{x}, \boldsymbol{v}): \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ – observer fantasy

Parameters of model of dependency $(\boldsymbol{v}, b)$ $\begin{cases} \boldsymbol{v} \in \mathbb{X} \text{ – directional point (vector) in the same space} \\ b \in \mathbb{R} \text{ – model shift (intercept)} \end{cases}$

Basic set of the objects $\{\omega_1^0, \ldots, \omega_n^0\} \subset \Omega$, and their mapping image $\{\boldsymbol{x}_\phi(\omega_1^0), \ldots, \boldsymbol{x}_\phi(\omega_n^0)\}$.

Theorem: Generalized linear characteristic of object $z(\omega, \boldsymbol{a}, b) = -\frac{1}{2}\sum_{i=1}^{n} a_i d^2(\omega, \omega_i^0) + b, \ \sum_{i=1}^{n} a_i = 0.$

So, the embedding into linear space is just our fantasy!

Objects Embedding into a Linear Space

3. Pairwise function of objects comparing is a metric

Set of real-world objects $\omega \in \Omega$.

Now as before, the comparison function $S(\omega', \omega'') = S(\omega'', \omega'):\ \Omega \times \Omega \to \mathbb{R}$

Additional requirements:

$S(\omega, \omega) = 0$, triangle inequality $S(\omega', \omega'') + S(\omega'', \omega''') \geq S(\omega', \omega''')$.

Let us note the distance as $d(\omega', \omega'') = \sqrt{S(\omega', \omega'')}$.

Central element $\phi \in \Omega$.

Two-argument commonality symmetric function for the center $\phi \in \Omega$:

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

Let the set of real-world objects finite $|\Omega| = M$ (probably, a VERY huge number!)

There is no change in theoretical speculations.

Inner product (indefinite) $K(\boldsymbol{x}, \boldsymbol{v}): \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ – observer fantasy

Parameters of model of dependency $(\boldsymbol{v}, b)$ $\begin{cases} \boldsymbol{v} \in \mathbb{X} \text{ – directional point (vector) in the same space} \\ b \in \mathbb{R} \text{ – model shift (intercept)} \end{cases}$

Basic set of the objects $\{\omega_1^0, \dots, \omega_n^0\} \subset \Omega$, and their mapping image $\{\boldsymbol{x}_\phi(\omega_1^0), \dots, \boldsymbol{x}_\phi(\omega_n^0)\}$.

Theorem: Generalized linear characteristic of object $z(\omega, \boldsymbol{a}, b) = -\frac{1}{2}\sum_{i=1}^{n} a_i d^2(\omega, \omega_i^0) + b,\ \sum_{i=1}^{n} a_i = 0.$

So, the embedding into linear space is just our fantasy!

Training Criterion: Minimum of Empirical Regularized Risk

Set of objects $\omega \in \Omega$, pairwise comparison function $S(\omega', \omega''): \Omega \times \Omega \to \mathbb{R}$.

Basic set $\{\omega_i^0, i=1,\dots,n\} \subset \Omega$	There is no information about goal characteristic. Just matrix of pairwise comparisons $\left[S(\omega_i, \omega_k),\ i,k=1,\dots,n\right]$
Training set (a part of basic set) $\{\omega_j, j=1,\dots,N\} \subset \Omega$	The known values of goal characteristic $y(\omega_j) \in \mathbb{Y}$, and matrix of pairwise comparisons $\left[S(\omega_j, \omega_l),\ j,l=1,\dots,N\right]$

Embedding (just mental) of objects into linear imaginary space with inner product, in general with indefinite inner product:

$\boldsymbol{x}(\omega): \Omega \to \mathbb{X}$, $K(\boldsymbol{x}', \boldsymbol{x}''): \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ – directly following from $S(\omega', \omega'')$.

Generalized linear characteristic of object $z(\boldsymbol{x}_j, \boldsymbol{v}, b) = K(\boldsymbol{x}_j, \boldsymbol{v}) + b$,

$\boldsymbol{v} \in \mathbb{X}$ – sought for directional vector, $b \in \mathbb{R}$ – sought for shift.

Link function defines by observer $q(y, z): \mathbb{Y} \times \mathbb{R} \to \mathbb{R}^+$, convex on z.

Regularized function defines by observer $V(\boldsymbol{v} \mid \mu): \mathbb{X} \to \mathbb{R}^+$

Training – finding $(\boldsymbol{v} \in \mathbb{X}, b \in \mathbb{R})$ by the criterion of minimum regularized empirical risk	$V(\boldsymbol{v} \mid \mu) + c\sum_{j=1}^{N} q\left(y_j, z(\boldsymbol{x}_j, \boldsymbol{v}, b)\right) \to \min(\boldsymbol{v}, b)$

Parametric representations of directional vector $\boldsymbol{v}(\boldsymbol{a}) = \sum_{i=1}^{n} a_i \boldsymbol{x}_\phi(\omega_i^0)$, $\sum_{i=1}^{n} a_i = 0$.

Training Criterion: Minimum of Empirical Regularized Risk

Set of objects $\omega \in \Omega$, pairwise comparison function $S(\omega', \omega''): \Omega \times \Omega \to \mathbb{R}$.

Basic set $\{\omega_i^0, i = 1, \ldots, n\} \subset \Omega$	There is no information about goal characteristic. Just matrix of pairwise comparisons $\left[S(\omega_i, \omega_k),\ i, k = 1, \ldots, n\right]$
Training set (a part of basic set) $\{\omega_j, j = 1, \ldots, N\} \subset \Omega$	The known values of goal characteristic $y(\omega_j) \in \mathbb{Y}$, and matrix of pairwise comparisons $\left[S(\omega_j, \omega_l),\ j, l = 1, \ldots, N\right]$

Link function defines by observer $q(y, z): \mathbb{Y} \times \mathbb{R} \to \mathbb{R}^+$, convex on z.

Regularized function defines by observer $V(\boldsymbol{v} \mid \mu): \mathbb{X} \to \mathbb{R}^+$

Training – finding $(\boldsymbol{v} \in \mathbb{X}, b \in \mathbb{R})$ by the criterion of minimum regularized empirical risk	$V(\boldsymbol{v} \mid \mu) + c \sum_{j=1}^{N} q\big(y_j, z(\boldsymbol{x}_j, \boldsymbol{v}, b)\big) \to \min(\boldsymbol{v}, b)$

Parametric representations of directional vector $\boldsymbol{v}(\boldsymbol{a}) = \sum_{i=1}^{n} a_i \boldsymbol{x}_\phi(\omega_i^0),\ \sum_{i=1}^{n} a_i = 0$.

Convex parametric regularized function $V(\boldsymbol{a} \mid \mu): \mathbb{R}^n \to \mathbb{R}^+$

Parametric training: Search $(\boldsymbol{a} \in \mathbb{R}^n, b \in \mathbb{R})$,	$V(\boldsymbol{a} \mid \mu) + c \sum_{j=1}^{N} q\big(y_j, z(\omega_j, \boldsymbol{a}, b)\big) + b \to \min(\boldsymbol{a}, b)$, $z(\omega_j, \boldsymbol{a}, b) = \sum_{i=1}^{n} \big(S(\omega, \omega_i^0)\big) a_i + b$, convex criterion.

Selection of subset of basic objects: L_1 and L_2 regularization combining

Elastic Net regularization: $V(\boldsymbol{a} \mid \mu) = \sum_{i=1}^{n} a_i^2 + \mu \sum_{i=1}^{n} |a_i|,\ \mu \geq 0$ – selectivity parameter.

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

Theorem. Let's values of $(\hat{\lambda}_1, \ldots, \hat{\lambda}_N)$ are decision of dual task of convex programming:

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

Then

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

As larger $\mu \geq 0$, then more the selectivity degree of basic objects.

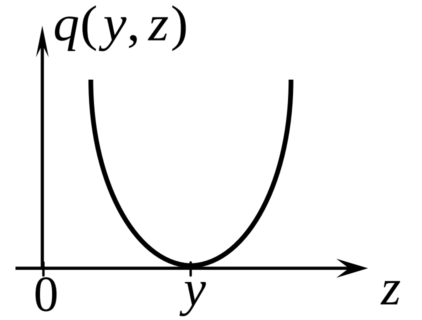
Link Functions for Some Particular Types of Goal Characteristic

Link function $q(y,z): \mathbb{Y}\times\mathbb{R}\to\mathbb{R}^{+}$ – observer fantasy

1) Regression analysis.

Goal characteristic $y\in\mathbb{Y}=\mathbb{R}$ – real number.

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

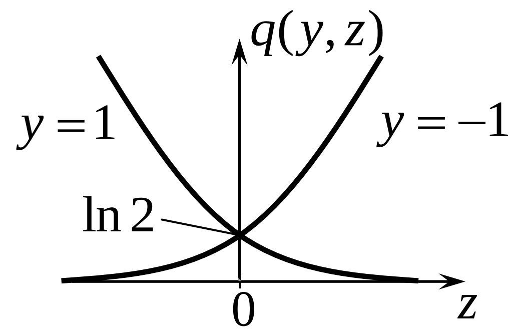


2) Two class pattern recognition task.

Goal characteristic $y\in\mathbb{Y}=\{-1,1\}$ – class index.

Logistic regression:

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

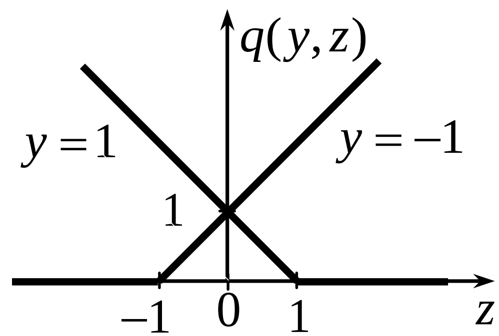


3) Two class pattern recognition task.

Goal characteristic $y\in\mathbb{Y}=\{-1,1\}$ – class index.

Support vector machine:

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



IDP-2016

BARCELONA

Thank you for your attention!