Convex Calibrated Surrogates for Low-Rank Loss Matrices with Applications to Subset Ranking Losses

Harish G. Ramaswamy[1], Shivani Agarwal[1] and Ambuj Tewari[2]

[1]Indian Institute of Science

[2]University of Michigan

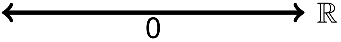
Calibrated Surrogates

Binary Classification

$$\color{blue}{\mathcal{Y}} = \color{red}{\widehat{\mathcal{Y}}} = \{\pm 1\}$$

$$\mathbf{L} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Minimize surrogate loss (e.g. hinge) over $\mathbb{R}$; learn $f : \mathcal{X} \to \mathbb{R}$



Final prediction in $\{\pm 1\}$:

$$\color{red}{h}(x) = \mathrm{sign}(f(x))$$

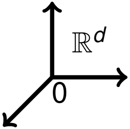
General Multiclass Problem

$$\underset{\text{(classes)}}{\color{blue}{\mathcal{Y}} = \{1, \ldots, n\}}; \underset{\text{(predictions)}}{\color{red}{\widehat{\mathcal{Y}}} = \{1, \ldots, k\}}$$

(predictions)

$$\mathbf{L} = \begin{bmatrix} 0 & 1 & 2 & 1 \\ 1 & 0 & 3 & 2 \\ 4 & 5 & 0 & 1 \end{bmatrix} \text{(classes)}$$

Minimize surrogate loss over $\mathbb{R}^d$; learn $f : \mathcal{X} \to \mathbb{R}^d$



Final prediction in $\{1, \ldots, k\}$:

$$\color{red}{h}(x) = \mathrm{pred}(f(x))$$

Convex Calibrated Surrogates for Low Rank Losses

$$\begin{bmatrix} & \mathbf{L} & \end{bmatrix}_{n \times k} = \begin{bmatrix} \mathbf{A} \end{bmatrix}_{n \times d} \begin{bmatrix} & \mathbf{B} & \end{bmatrix}_{d \times k} + \text{const}$$

Calibrated Convex Surrogate for L

$$\psi_L^*(y, \mathbf{u}) = \sum_{i=1}^{d} (u_i - A_{yi})^2 \qquad \mathrm{pred}_L^*(\mathbf{u}) \in \mathrm{argmin}_{t \in [k]} \sum_{i=1}^{d} u_i B_{it}$$

Application to Subset Ranking

Exponential sized loss matrices with low rank.

Loss matrix	**Rank**	**Efficient predictor**
NDCG	r	✓
Precision@q	r	✓
Expected Rank Utility	r	✓
Mean Average Precision	$\leq r^2$	X
Pairwise Disagreement	$\leq r^2$	X

$r =$ No. of docs. to be ranked

	σ_1	σ_2	$\dots$	$\hat{y}$	$\dots$	$\sigma_{r!}$
$00\dots00$						
$00\dots01$						
$\vdots$						
y						
$\vdots$						
$11\dots11$						

Poster Sat35
Today