

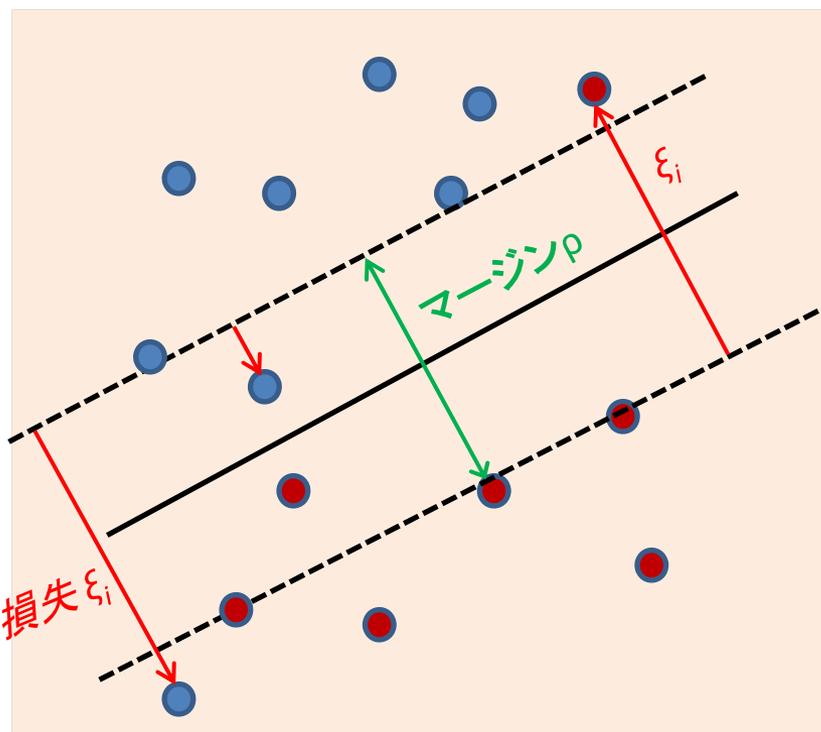
P17 行と列の生成による線形計画ブースティング

畑埜晃平・瀧本英二(九州大学)

1ノルムソフトマージン最適化

2ノルムソフトマージン最適化と並ぶ標準的な分類問題の定式化

「正例と負例をなるべく分離するマージンの大きな」超平面を求める



$$\max \rho - \nu \sum_{i=1}^m \xi_i \quad (\nu \text{は定数}, \nu \geq 1)$$

sub.to.

$$y_i \left(\sum_{j=1}^n \alpha_j h_j(\mathbf{x}_i) \right) \geq \rho - \xi_i, \quad (i = 1, \dots, m, y_i = 1, -1)$$

$$\|\boldsymbol{\alpha}\|_1 = \sum_{i=1}^n |\alpha_i| = 1, \alpha_i \geq 0 \quad \text{1ノルムで正規化}$$

特徴

- ・線形計画問題
- ・最適解が疎になりやすい(特徴選択に有効)

提案手法: Sparse LPBoost

- 与えられた $\varepsilon > 0$ に対して, 1ノルムソフトマージン最適化問題の ε 近似解を出力
- 元の問題を複数の(よりサイズの小さい)線形計画問題に分解して繰り返し解く
- **最適解の疎性を利用し高速化**

実験結果

データセット	例の個数	仮説の個数
Reuters-21578	10,170	30,389
RCV1	20,242	47,237
news20	19,996	1,355,193

