Singular Value Decomposition Machine

With various problem formulations

Yongxin (Taylor) Xi
July12, 2007 Group meeting

Core Idea of SVDM

- Suppose there are n data points $x_i \in R^m$ with labels $y_i \in \{0,1\}$

Possible Methods of this learning problem:

- Classification on data in its original space
- Dimension reduction $R^m \to R^p$ first, then construct classifier for the data with reduced dimension p. (noisy, redundant, latent variables)
- Dimension reduction and classification concurrently (SVDM belongs to this group)

Background: SVD & SVM

- SVM
  - Separable Case

$$Min \quad \frac{1}{2} \| q \|^2$$

$$Subject\ to: \ y_i (q^T x_i + b) \geq 1 \quad i = 1, 2, \ldots, n$$

  - Non-separable Case

$$Min \quad \frac{1}{2} \| q \|^2 + C \sum_{i=1}^{n} h_i$$

$$Subject\ to: \ y_i (q^T x_i + b) \geq 1 - h_i, \quad i = 1, 2, \ldots, n$$

$$h_i \geq 0, \forall i$$

Problem #1

Given: data $X = [x_1 \;\; x_2 \;\; \ldots \;\; x_n]$ $x_i \in R^m$ and label $y \in R^n$

Find: 1) $Z_{p \times n} = [z_1 \;\; z_2 \;\; \ldots \;\; z_n]$ $z_i \in R^p$

2) $W_{m \times p} = [w_1 \;\; w_2 \;\; \ldots \;\; w_p]$, with $\{w_i\}$ an ON basis

3) $q \in R^p$ and b

Such that: 1) $Z = W^T X$

2) $$\arg\min_{q,\, b} \; E = \frac{1}{2} \| q \|^2 + C \sum_{i=1}^{n} h_i$$

$$S.t. \qquad h_i \geq 0$$

$$y_i (q^T z_i + b) \geq 1 - h_i, \; i = 1, 2, \ldots n$$

Problem #2 (my formulation)

Given the same data as problem #1

Define: $$E = \frac{1}{2}\|q\|^2 + C\sum_{i=1}^{n} h_i$$

Now find: 1) Z; 2) W; 3) q, b; 4) hi
as in problem #1

Such that: 1) $Z = W^T X$

2) $h_i \geq 0$

3) $y_i(q^T z_i + b) \geq 1 - h_i,\ i = 1, 2, \ldots n$

4) $Min\|X - \hat{X}\|_F^2 + \lambda E,$ where $\hat{X} = W_{m\times p} Z_{p\times n}$

Note

- Prob #1 is the limiting case ($\lambda \to \infty$) of Prob #2

- Prob #2 reduces to finding p principle components of SVD when $\lambda \to 0$

Problem #3

- Two modifications of $\|q\|^2$ (one in objective and the other in constraints) in Prob #1 will lead to Prob #3

$$\arg\min_{q,\,b}\ E = \sum_{i=1}^{n} h_i$$

$$S.t. \qquad h_i \geq 0$$

$$y_i(q^T z_i + b) \geq 1 - h_i,\ i = 1, 2, \ldots n$$

$$\|q\|^2 \leq 1$$

- Prob #3 is the limiting case of prob #1 ($C \to \infty$) if the optimizer q in prob #1 satisfies: $\|q\|^2 \leq 1$

Problem #4 (Fransisco Pereira's)

- Two modifications of $\| q \|^2$ (one in objective and the other in constraints) in Prob #2 will lead to Prob #4
- Change E: (omit $\| q \|^2$) $E = \sum_{i=1}^{n} h_i$
- One more constraint added: $\| q \|^2 \leq 1$

- Prob #4 is a limiting case of prob #2 ($C \to \infty$) if the optimizer q in prob #2 satisfies: $\| q \|^2 \leq 1$

Effect of changing ||q|| term

- We no longer have ||q|| in the objective. Instead of minimizing ||q|| to achieve largest soft margin, we constrain it to be less than 1.

- This constraint regulates the soft margin of the classifier to be greater or equal to 2.

Claim #1

<u>Existence of solution to Prob #1</u>

- Problem #1 has a solution

Proof:

<u>Shown by:</u>

Continuous function on a compact set always achieve its minimum and maximum value on the set.

1) Continuous
2) Bounded
3) Closed (the limit of solution series is feasible)
4) Feasible set non-empty

Claim #2

Uniqueness of solution of problem #1?

- Let (W,q,b,h) be a solution of problem #1. Let R be a pxp ON matrix. Then (WR, $R^T q$, b, h) is also a solution of problem #1.

Proof:

- Intuitive explanations from the subspace perspective

Claim #3

These solutions share the same subspace spanned by the ON basis W

- Subspace spanned by cols of WR is the same as that spanned by cols of W:

  spancols(WR) = spancols(W)

Proof:

Claim #4

- Uniqueness of subspace spanned by cols of W?

1) Non-Unique for separable data (need to show);
   relation between Prob#1 and SVM Problem
2) Non-unique for non-separable data

<u>Counter-examples</u>

There exist
(W1,q1,b1,h1) and (W2,q2,b2,h2)
with spancols(W1)≠spancols(W2)
as two solutions of problem #1.

Questions to answer…

- Does prob #2 have a solution, and is the solution unique? [pr#1 => Yes, ?]
- What about prob #3 and prob #4?