SVM: a very brief introduction

Zhang Hongxin zhx@cad.zju.edu.cn

State Key Lab of CAD&CG, ZJU 2006-03-30

What is Support Vector Machines?



- A most recently developed learning system (in early 90's, by Vapnik and his coworkers), which could be applied in classification and regression;
- It does the following:
 - map original input space to higher dimension feature space, which is implemented implicitly by kennel function, such that "linear decision boundaries constructed in the high dimensional feature space correspond to nonlinear decision boundaries in the input space" [1];
 - in feature space, an optimal separating hyperplane is constructed, which could be determined by solving an optimization problem;
 - Lagrange multipliers and dual theory can then be applied to convert this optimization problem to a convex quadratic program subject to linear constraints.
- Advantages:
 - Better generalization
 - Global optimum

SVM applications



- SVMs were originally proposed by Boser, Guyon and Vapnik in 1992 and gained increasing popularity in late 1990s.
- SVMs are currently among the best performers for a number of classification tasks ranging from text to genomic data.
- SVMs can be applied to complex data types beyond feature vectors (e.g. graphs, sequences, relational data) by designing kernel functions for such data.
- SVM techniques have been extended to a number of tasks such as regression [Vapnik et al. '97], principal component analysis [Schölkopf et al. '99], etc.
- Most popular optimization algorithms for SVMs use *decomposition* to hill-climb over a subset of a_i 's at a time, e.g. SMO [Platt '99] and [Joachims '99]
- Tuning SVMs remains a black art: selecting a specific kernel and parameters is usually done in a try-and-see manner.

Useful tools



• SVM-Light

http://svmlight.joachims.org/

LibSVM

http://www.csie.ntu.edu.tw/~cjlin/libsvm/

Tutorial



 http://www1.cs.columbia.edu/~rkuang/candid acy/Burges98.pdf