An Introduction to Support Vector Machines

By
Ming-Chang Lee

Souce: Documentation for package of R `kernlab' and 'e1071'
Web: http://web.ydu.edu.tw/~alan9956/
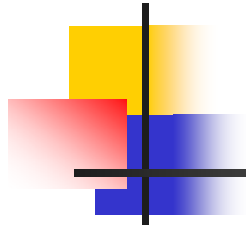Email: alan9956@ms1.hinet.net

Outline

1. Support Vector Machines

2. Binary classification

3. R demo

4. References

1. Support Vector Machines

- History :
  Statistical Learning Theory (Vapnik 1998)

- Development:

  - Binary classification SVM

  - Multi-class SVM

- **Application:**
  - Text categorization
  - Image recognition
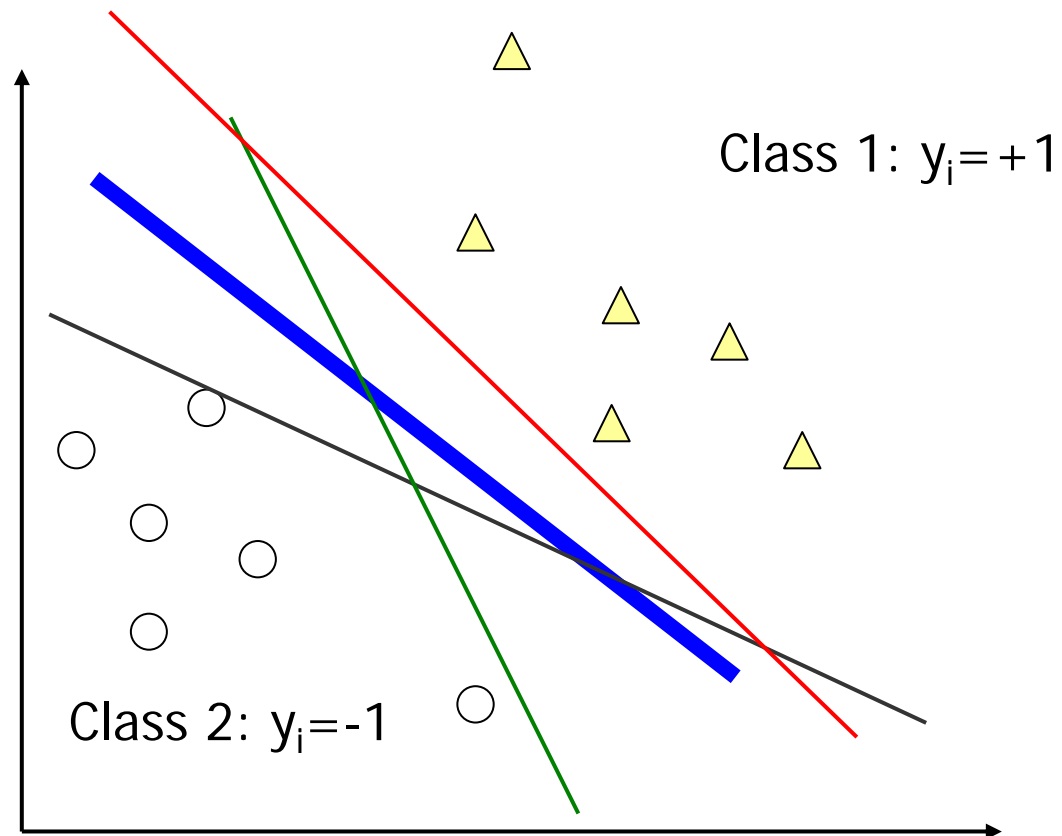 - Hand-written Digit Recognition
  - Bioinformatics

2. Binary classification

- Consider a two-class, linearly separable classification problem :

  - training data (x_i, y_i), $i = 1,\ldots,l$. $\{x_i, y_i\}$, $x_i \in R^n$
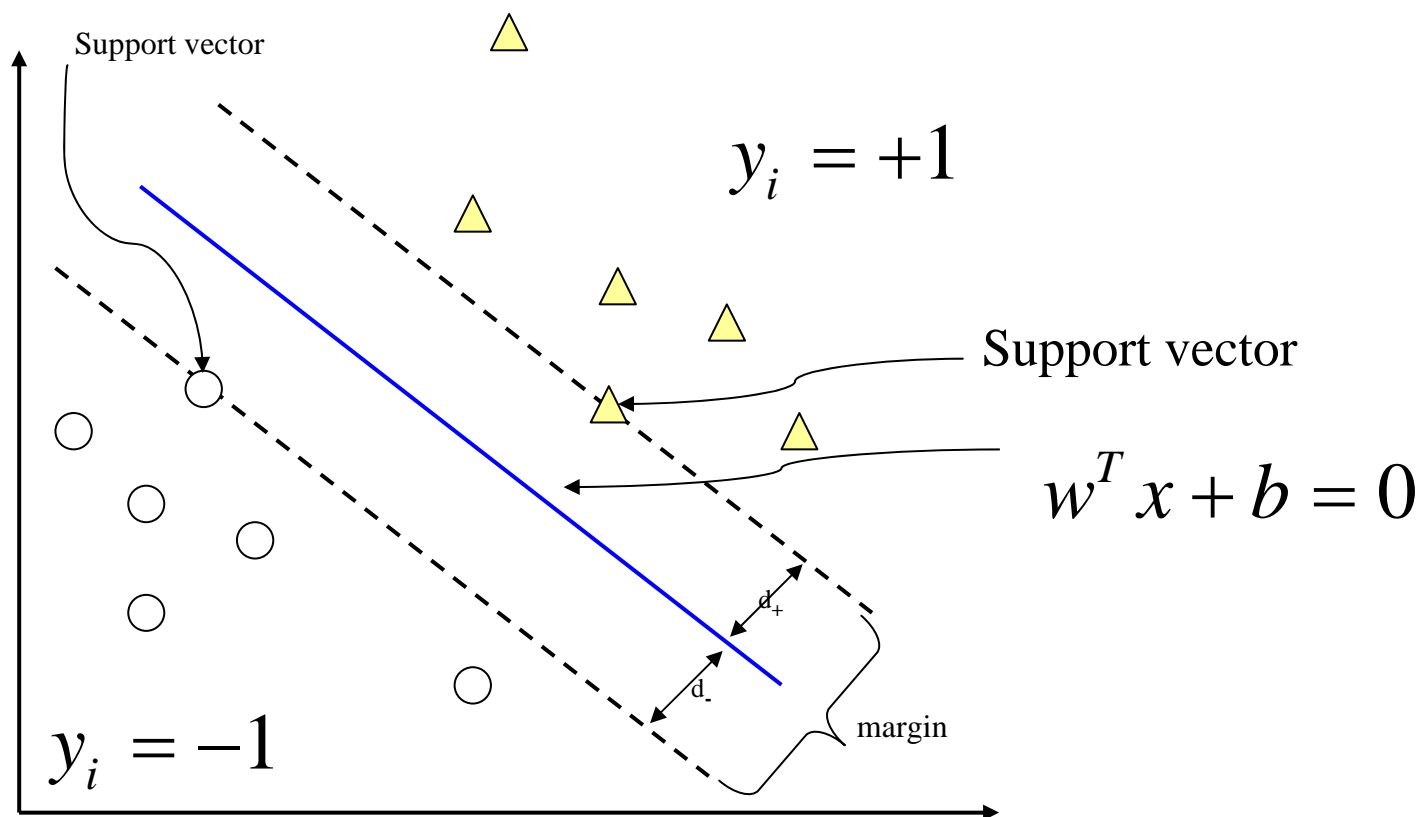
$$y_i = \begin{cases} +1, & x_i \in \text{class 1} \\ -1, & x_i \in \text{class 2} \end{cases}$$

Making decision boundary

Class 1: $y_i = +1$

Class 2: $y_i = -1$

Large-margin Decision Boundary

margin = (d+) 與(d-) → Goal: Max {margin}

Support vector

$y_i = +1$

Support vector

$w^T x + b = 0$

d_+

d_-

margin

$y_i = -1$

Find maximum margin

$$w^T x_i + b \geq 1 \quad if \ y_i = +1$$
$$w^T x_i + b \leq 1 \quad if \ y_i = +1$$

$$y_i(w^T x_i + b) \geq 1 \quad \forall i = 1, ..., l$$

$$m \arg in = (d_+) + (d_-) = \frac{1}{\|w\|} + \frac{1}{\|w\|} = \frac{2}{\|w\|} = \frac{2}{\sqrt{w^T w}}$$

$$Minimize \quad \cos t \ function, \Phi(w) = \frac{1}{2} w^T w$$

$$Subject \ to \quad y_i(w^T x_i + b) \geq 1 \quad \forall i = 1, ..., l.$$

Lagrange multipliers function

$$Minimize \quad J(w,b,\alpha) = \frac{1}{2}w^T w - \sum_{i=1}^{l} \alpha_i \left[y_i(w^T x_i + b) - 1 \right]$$

$$Subject\ to \quad \alpha_i \geq 0 \quad \forall i = 1,...,l.$$

$$\frac{dJ(w,b,\alpha)}{dw} = 0, \quad w = \sum_{i=1}^{l} \alpha_i y_i x_i$$

$$\frac{dJ(w,b,\alpha)}{db} = 0, \quad \sum_{i=1}^{l} \alpha_i y_i = 0$$

Dual problem

$$J(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{i=1}^{l} \alpha_i y_i w^T x_i - b \sum_{i=1}^{l} \alpha_i y_i + \sum_{i=1}^{l} \alpha_i$$

$$w^T w = \sum_{i=1}^{l} \alpha_i y_i w^T x_i = \sum_{i=1}^{l} \sum_{i=1}^{l} \alpha_i \alpha_j y_i y_j x_i^T x_j$$

Maximize $\quad Q(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{i=1}^{l} \alpha_i \alpha_j y_i y_j x_i^T x_j$

Subject to

$$(1) \sum_{i=1}^{l} \alpha_i y_i = 0$$

$$(2) \alpha_i \geq 0 \ \forall i = 1, 2, ..., l$$

Solution - Support vectors

- Many of the α_i are zero

- Support vectors (SV) x_i :

  - Non-zero α_i

  - The decision boundary is determined only by the SV

  - $$w = \sum_{i=1}^{l} \alpha_i y_i x_i$$

- Testing with a new data x_{new} :

  - Compute $w^T x_{new} + b$

  - If the result is positive , then classify x_{new} as class 1 , class 2 otherwise

3. R demo

- Data set: Glass

- Numbers 214, input variables: 9.

- Training data: 143(2/3)

- Testing data:71(1/3)

- Use SVM to classify

How to get the R?

→ R Binaries

Index of /bin

Name	Last modified	Size	Description
Parent Directory	14-Oct-2005 19:19	-	
linux/	21-Feb-2005 11:27	-	
macos/	19-Apr-2005 09:45	-	
macosx/	19-Oct-2005 15:31	-	
windows/	17-Dec-2004 16:38	-	

Apache/1.3.33 Server at cran.r-project.org Port 80

CRAN
Mirrors
What's new?
Task Views
Search

About R
R Homepage

Software
R Sources
R Binaries
Packages
Other

Documentation
Manuals
FAQs
Contributed
Newsletter

14/19
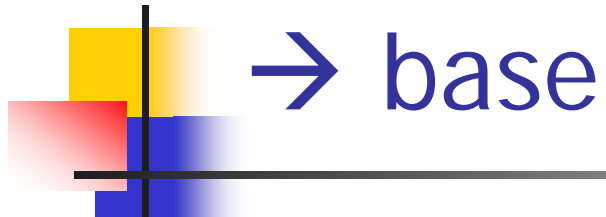
→ base

R for Windows

This directory contains binaries for a base distribution and packages to run on Windows (NT, 95 and later) on Intel and clones (but not NT on Alpha and other platforms).

Note: CRAN does not have Windows systems and cannot check these binaries for viruses. Use the normal precautions with downloaded executables.

Subdirectories:

base   Binaries for base distribution (managed by Duncan Murdoch)
contrib   Binaries of contributed packages (managed by Uwe Ligges)

Please do not submit binaries to CRAN. Package developers might want to contact Duncan Murdoch or Uwe Ligges directly in case of questions / suggestions related to Windows binaries.

You may also want to read the R FAQ and R for Windows FAQ.

Last modified: April 4, 2004, by Friedrich Leisch

CRAN
Mirrors
What's new?
Task Views
Search

About R
R Homepage

Software
R Sources
R Binaries
Packages
Other

Documentation
Manuals
FAQs
Contributed
Newsletter

→ R-2.2.0-win32.exe

R-2.2.0 for Windows

This directory contains a binary distribution of R-2.2.0 to run on Windows 95, 98, ME, NT4.0, 2000 and XP on Intel/clone chips.

Patches to this release are incorporated in the r-patched snapshot build.

CRAN
Mirrors
What's new?
Task Views
Search

A build of the development version (which will eventually become the next major release of R) is available in the r-devel snapshot build.

In this directory:

About R
R Homepage

README.R-2.2.0	Installation and other instructions.
CHANGES	New features of this Windows version.
NEWS	New features of all versions.
R-2.2.0-win32.exe	Setup program (about 25 megabytes). Please download this from a mirror near you. This corresponds to the file named **SetupR.exe** or **rwXXXX.exe** in pre-2.2.0 releases.
old	The previous release.
md5sum.txt	md5sum output for the setup program. A Windows GUI version of md5sum is available at http://www.md5summer.org/; a Windows command line version is available at http://www.etree.org/md5com.html.

Software
R Sources
R Binaries
Packages
Other

Please see the R FAQ for general information about R and the R Windows FAQ for Windows-specific information, including upgrade advice.

Documentation
Manuals
FAQs
Contributed
Newsletter

Note to webmasters: A stable link which will redirect to the current Windows binary release is
<CRAN MIRROR>/bin/windows/base/release.htm.

Last change: 2005-10-06, by Duncan Murdoch

16/19

→ Packages

网址(D) http://cran.r-project.org/

	dse	Bundle of tframe, dse1, dse2
	dyn	Time Series Regression
	dynamicGraph	dynamicGraph
	dynlm	Dynamic Linear Regression
	e1071	Misc Functions of the Department of Statistics (e1071), TU Wien
	eba	Elimination-By-Aspects (EBA) Models
	EbayesThresh	Empirical Bayes thresholding and related methods
	Ecdat	Data sets for econometrics
	eco	R Package for Fitting Bayesian Models of Ecological Inference in
	edci	Edge Detection and Clustering in Images
	effects	Effect Displays for Linear and Generalized Linear Models
	eha	Event History Analysis.
	elasticnet	Elastic Net Regularization and Variable Selection
	ElemStatLearn	Data sets, functions and examples from the book: "The Elements o
	ellipse	Functions for drawing ellipses and ellipse-like confidence region
	elliptic	elliptic functions
	emme2	Read and Write to an EMME/2 databank
	emplik	Empirical likelihood ratio for censored/truncated data
	EMV	Estimation of Missing Values for a Data Matrix
	energy	E-statistics (energy statistics) tests of fit, independence, clus
	ensembleBMA	Probabilistic forecasting using Bayesian Model Averaging
	Epi	A package for statistical analysis in epidemiology.
	epitools	Epidemiology Tools
	epsi	Edge Preserving Smoothing for Images
	equivalence	Provides tests and graphics for assessing tests of equivalence
	evd	Functions for extreme value distributions
	evdbayes	Bayesian Analysis in Extreme Value Theory
	evir	Extreme Values in R
	exactLoglinTest	Monte Carlo Exact Tests for Log-linear models
	exactRankTests	Exact Distributions for Rank and Permutation Tests
	extRemes	Extreme value toolkit.
	Fahrmeir	Data from the book "Multivariate Statistical Modelling Based on
	far	Modelization for Functional AutoRegressive processes
	faraway	Functions and datasets for books by Julian Faraway.

CRAN
Mirrors
What's new?
Task Views
Search

About R
R Homepage

Software
R Sources
R Binaries
Packages
Other

Documentation
Manuals
FAQs
Contributed
Newsletter

4. References

1. Burges, C. (1998) **A tutorial on support vector machines for pattern recognition**, *Data Mining and Knowledge Discovery*, 121-167.

2. Lin C. J., etc. **A Practical Guide to Support Vector Classification,** 2003.
   source: http://www.csie.ntu.edu.tw/~cjlin/

3. Vapnik, V. N. (1998) **Statistical Learning Theory**. Wiley.

THANKS

Q & A