Tentative syllabus for the course:

Understanding SVM (and associated kernel machines) through the development of a Matlab toolbox

Stéphane Canu
scanu@insa-rouen.fr, asi.insa-rouen.fr\~scanu

17-28 of february 2014, São Paulo

1 Course description

Kernel methods are a class of learning machine that has become an increasingly popular tool for learning tasks such as pattern recognition, classification or novelty detection. This popularity is mainly due to the success of the support vector machines (SVM), probably the most popular kernel method, and to the fact that kernel machines can be used in many applications as they provide a bridge from linearity to non-linearity. This allows the generalization of many well known methods such as PCA or LDA to name a few. Other key points related with kernel machines are convex optimization, duality and related sparcity.

The Objective of this course is to provide an overview of all these issues related with kernels machines. To do so, we will introduce kernel machines and associated mathematical foundations through practical implementation. All lectures will be devoted to the writing of some Matlab functions that, putting all together, will provide a toolbox for learning with kernels.

2 Prerequisites

Students should feel comfortable with the basics of linear algebra and matrix computation, mainly for solving linear systems. Multivariable calculus will be used. An undergraduate level is fine. To summarize, the following background is expected:

- multi variable calculus: matrices and differentials
- linear algebra and associated matrix computation (the basics)
- pre-existing working knowledge of probability and statistics (the basics also)
- programming: basic computer science principles and skills, at a level sufficient to write a reasonably non-trivial computer program.

3 Lectures

Each lecture will be a 3 hours session with 1 hour 30 on theory and the same for practice (programing with Matlab). The following table summarize a tentative course schedule.

[1] http://cs229.stanford.edu/section/cs229-cvxopt2.pdf
[2] www.iip.ist.i.kyoto-u.ac.jp/member/cuturi/Papers/pdk_in_ml.pdf
[3] http://arxiv.org/pdf/math/0701907.pdf
[4] mediamatica.ewi.tudelft.nl/sites/default/files/ML_SVDD_04.pdf
[5] svms.org/parameters/DuanKeerthiPoo2003.pdf
[6] loria.fr/~lauer/MSVMpack
[7] http://mlss.tuebingen.mpg.de/hofmann_slides.pdf
[8] http://www-als.ics.nitech.ac.jp/talk/120409.pdf

	Topics	Readings, useful links	Matlab functions
Lecture 1	Two classes classification problem: Introduction to Support Vector Machines (SVM), Linear SVM in the primal as a quadratic program and LPSVM for variable selection	`asi.insa-rouen.fr/enseignants/~scanu/tutorial_01.pdf`, `CVX`	`SVMClassPrimal`, `SVMValPrimal`
Lecture 2	Karush-Kuhn-Tucker (KKT) optimality conditions, lagrangian duality, Linear SVM in the dual, support vectors	Nocedal & Wright. Numerical Optimization, Springer-Verlag, 1999[1]	`SVMCheckKKT`, `SVMClassDual`, `SVMValDual`
Lecture 3	The non separable case and reformulations: convex hulls and hinge loss, `libsvm` and other QP solver	SVM and Kernel Methods Toolbox, LibSVM	`SVMClass`, `SVMVal`, `SVM_QP`
Lecture 4	Positive definite kernels, associated RKHS, differentiability and other kernels	Cuturi. Positive Definite Kernels in Machine Learning[2]	`SetKernelMatrix`
Lecture 5	SVM in a RKHS, representer theorem and other kernel machines (splines and kernel logistic regression)	Hofmann, Schölkopf, & Smola. Kernel Methods in Machine Learning[3]	`SVMKernel`
Lecture 6	Novelty detection, Support Vector Data Description (SVDD), One Class SVM, minimum enclosing balls and level sets	Tax & Duin, Support Vector Data Description[4]	`SVDDClass`, `SVDDVal`, `OCSVM`
Lecture 7	Machine Learning: tuning the hyperparamters, C, the bandwidth and the KKT gap	Duan & Keerthi, Evaluation of simple performance measures for tuning SVM hyperparameters [5]	`SVMClass_CV`
Lecture 8	MultiClass SVM, estimating the posterior probabilities and structured output	MSVMpack: a Multi-Class Support Vector Machine Package[6] Structure output[7]	`SVMMultiClass`
Lecture 9	How to choose the kernel, multiple kernel learning	SimpleMKL	`SVM_MKL`
Lecture 10	Regularization path for SVM, sub differential, DrSVM (Elastic net SVM) and L_0 SVM	Parametric optimization in machine learning[8]	`LPSVM`, `SVM_RegPath`

4 Course material

Course web site on Moodle: `http://paca.ime.usp.br/course/view.php?id=776`

- Books
  - V. Vapnik, The Nature of Statistical Learning Theory, Springer, 1995.
  - B. Schölkopf & A. J. Smola, Learning with Kernels, MIT Press, 2002.
  - J. Shawe-Taylor & N. Cristianini, Kernel Methods for Pattern Analysis, Cambridge, 2004.
  - T. Hastie, R. Tibshirani & J. FriedmanThe Elements of Statistical Learning, Springer, 2009. Freely available online at `statweb.stanford.edu/~tibs/ElemStatLearn/`.
  - I. Steinwart & A. Christmann, Support Vector Machines. Springer, New York, 2008.
  - M. Mohri, A. Rostamizadeh & A. Talwalkar, Foundations of Machine Learning, MIT, 2012.
  - J. Nocedal & S. Wright, Numerical Optimization, Springer, 2006.
  - S. Boyd and L. Vandenberghe, Convex optimization, Cambridge University Press, 2004.
- Slides
  - B. Schölkopf tutorial on Kernels at MLSS 2013 at MPI Tuebingen: `mlss.tuebingen.mpg.de/schoelkopf_kernels.pdf`
  - S. Wright tutorial on Optimization (and Learning) at MLSS 2013 `mlss.tuebingen.mpg.de/wright_slides.pdf`
  - Distances and Kernels, M. Cuturi: `www.iip.ist.i.kyoto-u.ac.jp/member/cuturi/Teaching/KAIST/kaist_2013.pdf`
  - Training a support vector machine in the primal O Chapelle `www.kyb.mpg.de/fileadmin/user_upload/files/publications/attachments/neco_%5B0%5D.pdf`
- Courses on Machine learning and related topics (more slides)
  - Machine learning with kernel methods by Jean-Philippe Vert at Mines ParisTech: `cbio.ensmp.fr/~jvert/svn/kernelcourse/course/2014mva/`
  - Machine Learning at Stanford: `cs229.stanford.edu/section/cs229-cvxopt.pdf`
  - Machine learning at Berkley: `www.cs.berkeley.edu/~jordan/courses/294-fall09`
  - Machine Learning at Harvard: `www.seas.harvard.edu/courses/cs281/`
  - Optimization at MIT: `www.mit.edu/~dimitrib/Convex_Slides_2012.pdf`
  - More material on the Machine Learning Summer Schools sites: `www.mlss.cc/`
- Matlab toolbox
  - libsvm: `www.csie.ntu.edu.tw/~cjlin/libsvm/`
  - CVX: Matlab software for disciplined convex programming: `stanford.edu/~boyd/cvx/`
  - SVM KM: `asi.insa-rouen.fr/enseignants/~arakoto/toolbox/`
  - MLOSS: `mloss.org/software/`
- Data set
  - Machine Learning Repository at UC Irvine: `archive.ics.uci.edu/ml/`
  - For objectively comparing ML algorithms: `mlcomp.org`
  - MNIST: `yann.lecun.com/exdb/mnist/`
  - Datasets for deep learning: `deeplearning.net/datasets/`
- Blogs
  - Machine Learning News: `groups.google.com/forum/#!forum/ml-news`
  - John Langford's: `hunch.net`
  - a list of Machine Learning and related area blogs: `www.cs.waikato.ac.nz/~bernhard/good-machine-learning-blogs.html`