THE WEST COAST OPTIMIZATION MEETING

Depts. of Mathematics and Applied Mathematics, University of Washington

All talks will be in Guggenheim 317

FRIDAY, NOVEMBER 19, 1999

6:30–9:30+ Party at Terry Rockafellar's home, 4531 NE 93rd Street, 206–527–9637
The cost per person will be $10/"student" and $15/"others."

SATURDAY, NOVEMBER 20, 1999

8:30–9:00 ————Refreshments in Guggenheim 408, the Applied Math Lounge

9:00–9:10 "Graphical properties of set-valued partial differential operators," Teemu Pennanen, University of Washington

9:10–9:50 "Error bounds for conic optimization problems," Madhu Nayakkankuppam, Simon Fraser University

9:50–10:15 ————Refreshments in Guggenheim 408, the Applied Math Lounge

10:15–10:25 "Dualizing the properties of saddle functions," Rafal Goebel, University of Washington

10:25–11:05 "Soft margin boosting using column generation," Kristin Bennett, Microsoft and Rensselaer Polytechnic Institute

11:05–11:30 ————Refreshments in Guggenheim 408, the Applied Math Lounge

11:30–12:10 "Statistical estimation: a variational analysis viewpoint," Roger Wets, University of California-Davis

12:10–14:00 ————Lunch expedition to University Avenue

14:00–14:40 "Convex optimization with Bregman distances," Patrick Combettes, City University of New York

14:40–15:20 "A non-interior continuation method for semidefinite programming," Paul Tseng, University of Washington

---

The **West Coast Optimization Meeting** occurs twice each year. Contact:

Prof. J. M. Borwein at the Dept. of Mathematics and Statistics, Simon Fraser University, Vancouver: (604) 291-3070, e-mail jborwein@cecm.sfu.ca

Prof. R. T. Rockafellar at the Dept. of Mathematics, University of Washington, Seattle: (206) 543-1916, e-mail rtr@math.washington.edu

TALK ABSTRACTS for WCOM

Seattle, November 20, 1999

Teemu Pennanen, *Graphical Properties of Set-Valued Partial Differential Operators*

This talk is concerned with partial differential operators in the generalized divergence form D^*AD, where D is a linear derivative operator and A is a set-valued monotone mapping. I will give new conditions for maximal monotonicity, graphical convergence and measurability for mappings in this form. These results are based on more general results on composite mappings that I have derived using an abstract duality framework developed in my thesis.

Madhu Nayakkankuppam, *Error Bounds for Conic Optimization Problems*

We consider the problem of minimizing a linear function over the intersection of an affine space and a symmetric cone. Under assumptions which are generically satisfied, we derive a quantitative error bound for the solution via a simple application of the Kantorovich theorem.

Rafal Goebel, *Dualizing the Properties of Saddle Functions*

Several properties of a convex function f are reflected in properties of f^*, the function conjugate to f. For example, f is differentiable if and only if f^* is strictly convex. The talk will briefly present corresponding results in the setting of saddle functions. The motivation for investigating these issues comes from optimal control theory and differential games.

Kristin Bennett, *Soft Margin Boosting Using Column Generation* (based on joint work with Ayhan Demiriz and John Shawe-Taylor)

Support Vector Machines (SVM) and Boosting are two major new research topics to emerge in machine learning in the last ten years. Using the classification problem, we will examine how mathematical programming plays a key role in both of these methods and then show how they can be combined to create a powerful new approach. The underlying problem is to construct a function that discriminates between the members of two classes. In SVM the objective is to choose a function that maximizes the separation or margin between the two classes while minimizing the misclassification error. The SVM can be formulated as a linear or quadratic programming depending on the choice of metrics.

The idea in boosting is to not use a single discriminant function, but rather to use a linear combination of many discriminant functions (called weak learners) in order to improve future prediction. These popular methods have been shown to be gradient descent algorithms that maximize the margin of separation much the same as in SVM. The catch is that these boosting approaches can be sensitive to outliers, slow to converge, and prone to local minima. In this work we address these problems by adapting the soft margin cost function of support vector machines to boosting. We prove that minimizing the soft margin error function directly optimizes a generalization error bound.

We formulate the boosting problem as if all possible weak learners had already been generated. The class labels produced by the weak learners become the new feature space of the problem. The boosting task becomes to construct a learning function in the label space that minimizes misclassification accuracy and maximizes the soft margin. The resulting linear program can be efficiently solved using column generation techniques developed for large-scale optimization problems. The columns of the linear program, each corresponding to one weak learner, are generated as needed. In essence the learning machine becomes an "oracle" that generates columns needed by a simplex algorithm. The dual variables of the linear program provide the misclassification costs used by the learning machine. The resulting "SoftBoost" algorithm has many attractive properties. The algorithm has well defined convergence criteria. It converges in a finite number of iterations to a globally optimal solution. In computational experiments, the algorithm performs very well both in terms of computational time and generalization testing. The algorithm requires very few iterations. Thus few weak learners are actually generated and even fewer appear in the optimal learning ensemble. Extensions of this approach to more general boosting problems are also discussed.

Roger Wets, *Statistical Estimation: A Variational Analysis Viewpoint*

Statistics and Optimization have been closely linked from the very outset. The search for a 'best' estimator (least squares, maximum likelihood, etc.) certainly relies on optimization tools. On the other hand, Statistics has often provided the motivation for the development of algorithmic procedures for certain classes of optimization problems. However, it's only relatively recently, more specifically in connection with the development of an approximation and sampling theory for variational problems, that the full connection has come to light. This in turn suggests a more comprehensive approach to the formulation of statistical estimation questions.

Patrick Combettes, *Convex Optimization with Bregman Distances* (based on joint work with H. H. Bauschke and J. M Borwein)

In convex optimization, many fundamental notions such as projection, proximation, Fejer monotonicity, or nonexpansiveness are intrinsically tied to the metric of the underlying Banach space. Replacing the standard metric by a so-called Bregman distance (roughly speaking, a suitable nonnegative function of two variables induced by a convex function) leads to new mathematical objects that may be more useful than their standard counterparts in certain problems. This talk will consist of three parts. In the first part, several classes of operators associated with Bregman distances will be presented and analyzed (projectors, proximators, resolvents, firm operators, etc.). The concept of Legendreness, first introduced by Rockafellar in finite-dimensional spaces, is re-examined and shown to play a crucial role in our infinite-dimensional analysis. The second part will be devoted to the study of algorithms based on Bregman distances. The notion of Bregman-monotonicity will be seen to provide a convenient framework to investigate the convergence of such algorithms. Applications (including convex feasibility and best Bregman approximation problems) will be discussed in the third part.

Paul Tseng, *A Non-Interior Continuation Method for Semidefinite programming*

There recently has been much interest in non-interior continuation/smoothing methods for solving linear/nonlinear complementarity problems. We describe extensions of such methods to complementarity problems defined over the cone of block-diagonal symmetric positive semidefinite real matrices. These extensions involve the Chen-Mangasarian class of smoothing functions and the smoothed Fischer-Burmeister function. Issues such as existence of Newton directions, boundedness of iterates, global convergence, and local superlinear convergence will be studied. Preliminary numerical experience on semidefinite linear programs will be reported.