### CCIT Report #555

# September 2005

## Smoothing Method of Multipliers for Sum-Max Problems

Michael Zibulevsky

Department of Electrical Engineering Technion - Israel Institute of Technology E-mail: mzib@ee.technion.ac.il

#### Abstract

We study a class of nonsmooth unconstrained optimization problems, which includes the problem of minimizing the sum of pairwise maxima of smooth convex functions. Minimum  $l_1$ -norm approximation is a particular case of this problem. Combining the ideas of Lagrange multipliers and of smooth approximation of max-type function, we obtain an extended notion of nonquadratic augmented Lagrangian. Our approach does not require artificial variables, and preserves sparse structure of Hessian in many practical cases. We present the corresponding method of multipliers, and its convergence analysis for a dual counterpart, resulting in a proximal point maximization algorithm. The practical efficiency of the algorithm is supported by computational results for large-scale problems, arising in structural optimization.

#### 1 Introduction

We consider the non-smooth convex unconstrained optimization problem

$$\min_{x \in \mathbb{R}^n} \left\{ F(x) = f(x) + \sum_{i=1}^m \max\left[\alpha_i h_i(x), \ \beta_i h_i(x)\right] \right\}$$
(1)

where f(x) and  $h_i(x)$ , i = 1, ..., m are smooth convex functions, defined over entire space  $\mathbb{R}^n$ ;  $\alpha_i < \beta_i$  are certain constants. In order to guarantee convexity, we assume, that for any particular index i, the values  $\alpha_i$  and  $\beta_i$  are non-negative, if  $h_i(x)$  is nonlinear<sup>1</sup>. Problem in this setting arises for example in Truss Topology Design [1, 5]. This is a generalization of least  $l_1$  norm approximation (regularization) problem, when the coefficients  $\alpha_i = -1$ ,  $\beta_i = 1$ ,

<sup>&</sup>lt;sup>1</sup>Without loss of generality, one could set  $\alpha_i = 0$  for all *i* by changing *f* to  $f + \sum_i \alpha_i h_i$ and  $\beta_i$  to  $\beta_i - \alpha_i$ . This similification would not change significantly our derivation, nor the computational burden of the method, however it can be useful in various cases.