Gradient flows for discrete optimization on the hypercube

P.-A. Absil and R. Sepulchre
Institut Montefiore, Université de Liège
Bât. B28 Systèmes, Grande Traverse 10, 4000 Liège, Belgium
absil@montefiore.ulg.ac.be, R.Sepulchre@ulg.ac.be

Keywords: optimization, gradient systems, Lyapunov stability, interior point methods.

Several problems, such as the Traveling Salesman Problem, the Maximum Likelihood Decoding problem or the 0-1 nonlinear programming problem involve the minimization of a *multilinear* function E, i.e. a function E such that $\mathrm{D}^2_{x_i}E \equiv 0$ for all i, over the discrete set $\{0,1\}^n$.

This optimization problem can be solved by a discrete algorithm which searches among adjacent corners whether the cost can be decreased until it reaches a local minimum. Because such a discrete algorithm is usually not efficient, one would prefer to take benefit of the multilinear function E and use an interior point approach for tackling this problem.

This paper is in many points the continuation of an article by Vidyasagar [Vid95] in which Hopfield's continuous-time neural network

$$\dot{u}_i = -\frac{u_i}{\alpha_i} - \frac{\partial E}{\partial x_i}(\boldsymbol{x}) + b_i,\ x_i = g_i(\lambda u_i), \quad i = 1, \dots, n \tag{1}$$

is used in order to find a local minimum of $\hat{E} := E|_{\{0,1\}^n}$. We took the gist out of (1) by considering the gradient system

$$\dot{x}_i = -h(x_i)\frac{\partial E}{\partial x_i}(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^n. \tag{2}$$

where the function h — which we refer to as the *temporisation function* — plays the role of an (inverse) metric and is selected to ensure the invariance of the hypercube $[0,1]^n$. Roughly said, (1) is the flat gradient system in $\mathbb{R}^n$ corresponding to a particular choice of h in (2), in the limiting case $\alpha_i = +\infty$ and $b_i = 0$. In particular, the bias b_i introduced by Vidyasagar is instrumental in the convergence analysis of (1) but we discuss some reasons for not considering a bias term in our analysis of (2).

We show that, under quite general choices of h, the strict local minima of $\hat{E}$ and the asymptotically stable points of (2) coincide, and that the stable points of (2) are either local minima of $\hat{E}$ or lie on a face all the corners of which are local minima of $\hat{E}$.

The geometry of (2) is further analysed by reducing (2) to the flat gradient system $\dot{u}_i = -\partial\bar{E}/\partial u_i$ under a suitable change of variables. Our classification of the temporisation in different categories enlightens connections between (2), (1) and several classical interior point methods for linear programming.

References

[Vid95] M. Vidyasagar, *Minimum-seeking properties of analog neural networks with multilinear objective functions*, IEEE Trans. Autom. Control **40** (1995), 1359–1375.