Assignment1: Optimization for Machine Learning

-by Enric Fita Sanmartin and Shivali Dubey

Exercise 1.1

A graphical model consists of a graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, discrete space of all labellings $\mathcal{Y}_v$ and a corresponding cost vector θ. Here, $\mathcal{V}$ and $\mathcal{V}$ represent the nodes and edges respectively of a given graph $\mathcal{G}$. Our main goal while labelling the nodes for a given graphical model is to minimize the cost, θ_u(s) we pay for assigning label s$\in\mathcal{Y}_u$ to the node u$\in\mathcal{V}u$. Hence, it becomes an optimization problem of labeling y^* with minimal total cost. This is called *energy minimization* or *maximum a posteriori* problem and can be expressed as:

$$y^* = argmin_{y\in\mathcal{Y}_v}\left[E(y;\theta) = \sum_{u\in\mathcal{V}}\theta_u(y_u) + \sum_{uv\in\mathcal{E}}\theta_{uv}(y_u, y_v)\right]$$

.

here, θ_u is called the *unary cost*, which is the preferance for a label y_u. On the other hand, θ_{uv} are the dependencies between labels (y_u and y_v) assigned to different nodes (u and v) which are defined for each edge uv$\in\mathcal{E}$ in the following way:

$$\theta_{uv}(st) = [0, s = t; \alpha, s \neq t]$$

These are called *pairwise costs.* Here, s and t are any pair of labels such that (s,t) $\in \mathcal{Y}_u \times \mathcal{Y}_v$.