ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 9 Statistical Physics for Communication and Computer Science
Homework 9 April 24, 2013

Problem 1 (A generalization of IST and its connection to LASSO.). The Iterative Soft Thresholding algorithm has the form

$$x_i^{t+1} = \eta(x_i^t + (A^T \underline{r}^t)_i; \lambda)$$
$$\underline{r}^t = \underline{y} - A\underline{x}^t$$

starting from the initial condition $x_i^0 = 0$. Consider the following generalization. Let θ_t and b_t be two sequences of scalars (called respectively "thresholds" and "reaction terms") that converge to fixed numbers θ and b. Construct the sequence of estimates according to the iterations

$$x_i^{t+1} = \eta(x_i^t + (A^T \underline{r}^t)_i; \theta_t)$$
$$\underline{r}^t = \underline{y} - A\underline{x}^t + b_t r^{t-1}$$

The goal of the exercise is to prove that: if x^*, r^* is a fixed point of these iterations, then x^* is a stationary point of the LASSO cost function $L(\underline{x}|\underline{y}, A) = \frac{1}{2}||\underline{y} - A\underline{x}||_2^2 + \lambda||x||_1$ for

$$\lambda = \theta(1 - b)$$

Note that this theorem does not say how to specify suitable sequences b_t and θ_t. The point of AMP is that it specifies unambiguously that one should take $b_t = ||x||_0/m$ (for θ_t there is more flexibility). We will see in the next chapter that with this choice *state evolution correctly tracks the average behavior of the iterative algorithm*, which allows to assess its performance.

The proof proceeds in two steps.

(i) Show that the stationarity condition for the LASSO cost function is

$$A^T(\underline{y} - A\underline{x}^*) = \lambda\underline{v}, \qquad v_i = \text{sign}(x_i)$$

where $v_i = \text{sign}(x_i)$ for $x_i \neq 0$ and $v_i \in [-1, +1]$ for $x_i = 0$.

(ii) Show that the fixed point equations corresponding to the iterations above are

$$x_i^* + \theta v_i = x_i^* + (A^T \underline{r}^*)_i$$
$$(1 - b)\underline{r}^* = \underline{y} - A\underline{x}^*$$

Remark that these two equations implies the stationary condition in item (i).

Problem 2 (AMP with automatic adjustment of threshold). In class, starting from the min-sum equations, we derived an AMP algorithm of the form

$$x_i^t = \eta(x_i^{t-1} + (A^T \underline{r})_i; \theta_t)$$
$$\underline{r}^t = \underline{y} - A\underline{x}^t + \frac{||x^t||_0}{m} r^{t-1}$$

We argued that a reasonable choice for $\theta_t = \alpha ||r^t||_2/\sqrt{m}$. There are however other choices that yield good performance. In particular, one of them follows directly from the min-sum equations. The resulting algorithm is slightly more complex and it turns out there is no benefit in performance.

Deduce from the message passing equations obtained after the quadratic approximation, that one can adjust the threshold according to the iterations

$$\frac{\theta_{t+1}}{\lambda} = 1 + \frac{\theta_t}{\lambda} \frac{||x^t||_0}{m}$$

Use the same assumption done in class, namely that $1 + \sum_{j \in \partial a \backslash i} A_{aj}^2 \gamma_{j \to a}^t$ is independent of a and i.