Tarea

Fecha de entrega: 10 de noviembre de 2016

1. Sea $t > 0$ y $h(x) = \|x\|_1$. Calcule $\operatorname{prox}_{th}(x)$.

2. Considere la función $f(\beta) = \sum_{i=1}^{n} \left(-y_i x_i^T \beta + \log(1 + \exp(x_i^T \beta))\right)$ con $(x_i, y_i) \in \mathbb{R}^p \times \{0, 1\}$. Usando los datos del archivo Datos.csv, donde la última columna corresponde a los valores de y_i, minimice $h(\beta) = f(\beta) + \lambda\|\beta\|_1$ con diferentes valores de $\lambda > 0$. Para todos lo métodos inicie las iteraciones en $\beta_0 = 0$ y grafique $\log(h(\beta_k) - h^*)$ en cada iteración. Grafique además $\|\beta^*\|_1$ para cada valor de λ. ¿Puede encontrar algún método que sea mejor para valores pequeños de λ? ¿Para grandes?

   a) Use el método del subgradiente. Explique claramente la forma de escoger el subgradiente y el tamaño de los pasos. Sugerencia: Consulte el ejemplo 3.1.5.5 (página 133) del libro de Nesterov.

   b) Use el método del subgradiente estocástico. Explique cómo lo hace.

   c) Use el método proximal con *backtracking*.

   d) Use el método proximal acelerado.

3. Considere el problema de minimizar $\sum_{i=1}^{n} f_i(\eta_i) + \lambda\|\beta\|_1$ sujeto a $\eta_i - \beta = 0$ para $i = 1, \ldots, m$, con $f_i(\eta) = -y_i x_i^T \eta + \log(1 + \exp(x_i^T \eta))$. Note que este problema es equivalente al anterior. Implemente el siguiente algoritmo ADMM con consenso:

$$\eta_i^{k+1} := \operatorname*{argmin}_{\eta} \left(f_i(\eta) + \frac{\alpha_k}{2}\|\eta - \beta^k + \mu_i^k\|^2 \right)$$

$$\beta^{k+1} := \operatorname*{argmin}_{\beta} \left(\lambda\|\beta\|_1 + \frac{n\alpha_k}{2}\|\beta - \bar{\eta}^{k+1} - \bar{\mu}^k\|^2 \right)$$

$$\mu_i^{k+1} := \mu_i^k + \eta_i^{k+1} - \beta^{k+1},$$

   donde $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

Mauricio Junca