STAT 7934: Take Home Final

Due by email to gmichail@ufl.edu on Tu Apr 28

No collaboration allowed. Your solutions must be typed. Also include the source file of your code.

(1) Consider the proximal operator

$$\text{prox}_f(x) = \text{argmin}_u\big(f(u) + \frac{1}{2}||u - x||_2^2\big).$$

Give a formula or simple algorithm to calculate the proximal operator for the following functions, where $x \in \mathbb{R}^n$.

(a) ($f(x) = ||x||_1$ with domain of $f = \{x : ||x||_\infty \leq 1\}$.

(b) $f(x) = \max_k\{x_k\}, k = 1, \cdots, n$.

(c) $f(x) = ||Ax - b||_1$,, where $AA' = D$ and D is a diagonal matrix with positive entries.

(d) $f(X) = -\text{logdet}(X)$, where $X \in S^n$ and domain$(f) = S^n_{++}$. For this problem, the proximal operator is defined as

$$\text{prox}_f(x) = \text{argmin}_U\big(f(U) + \frac{1}{2}||U - X||_F^2\big),$$

where $||\cdot||_F$ denotes the Frobenius norm.

(2) Consider the following minimization problem

$$\min(x_1 - x_2),$$

subject to $x \in X = \{x : x_1 \geq 0, x_2 \geq 0\}, x_1 + 1 \leq 0, 1 - x_1 - x_2 \leq 0.$
Derive the dual. What do you observe?

(3) Let $A_1, \cdots, A_J$ be matrices in $\mathbb{R}^{m\times n}$ and let $y_1, \cdots, y_J \in \mathbb{R}$. Consider the trace-norm contrained optimization problem.

$$\min_{X\in\mathbb{R}^{m\times n}} \sum_{j=1}^{J}\big(y_j - \text{trace}(X'A_j)\big)^2,$$

subject to $||X||_* \leq t$, where $||\cdot||_*$ denotes the nuclear norm.

Develop an algorithm to solve this problem. Test your algorithm on an instance of this problem. Provide all the details of how the data were generated and discuss your results.

(4) In a paper in JASA, volume 106, 594-607, 2011 the authors proposed the following minimization problem for the Gaussian graphical model:

$$\min ||\Omega||_1, \text{ subject to } ||S\Omega - I||_\infty \leq t,$$

where S is the empirical covariance matrix obtained from the data. Further, they showed (Lemma 1) that one can instead solve p node-wise problems of the form

$$\min ||\beta||_1, \text{ subject to } ||S\beta - e_i||_\infty \leq t, \beta \in \mathbb{R}^p,$$

where e_i is a standard unit vector in $\mathbb{R}^p$. The latter problem corresponds to the Dantzig selector optimization problem in Candes and Tao, Annals of Statistics 2007 and can be solved by a linear program.

Using the above formulation, consider estimating *jointly* two Gaussian graphical models, under the assumption that they are individually sparse but share connectivity patterns between their nodes, but also exhibit differences.

(a) Formulate the optimization problem: write down the objective function and the constraints.

(b) Devise an algorithm to solve it.

(c) Following the mechanism discussed in Guo et al. (Biometrika, 2011) generate data for networks of size $p = 25, n_i = 150, i = 1, 2$ for a chain, nearest neighbor and scale-free graphs. Apply your algorithm to the data generated and discuss how you tune the regularization parameters. Report false positive, false negative and Frobenius norm estimation error rates over 20 replicates.