E9 203: Homework - 3

Assigned on: 26 Feb. 2020, due 09 Mar. 2020

1 Topics

- Uniqueness and uncertainty
- Basic algorithms
- Reweighted algorithms
- Sparse Bayesian learning

2 Problems

1. Given $A \in \mathbb{C}^{m \times N}$ and $\tau > 0$, show that the solution of
$$\min_{\mathbf{z} \in \mathbb{C}^N} \|A\mathbf{z} - \mathbf{y}\|_2^2 + \tau \|\mathbf{z}\|_2^2$$
is given by
$$z^{\#} = \left(A^H A + \tau \mathbf{I}\right)^{-1} A^H \mathbf{y}.$$

2. Given $\mathbf{x} \in \mathbb{R}_+^N$ with $\mathbf{x}_1 \geq \mathbf{x}_2 \geq \cdots \geq \mathbf{x}_N \geq 0$, show that for each $1 \leq s \leq N$ and $r > 1$
$$f(\mathbf{x}) \triangleq \sum_{j=s}^{N} \mathbf{x}_j^r$$
is a convex function of $\mathbf{x}$.

3. (Determinant of a Vandermonde Matrix) The *Vandermonde matrix* associated with $x_0, x_1, \ldots, x_N \in \mathbb{C}$ is defined as
$$\mathbf{V} \triangleq \begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^N \\ 1 & x_1 & x_1^2 & \cdots & x_1^N \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_N & x_N^2 & \cdots & x_N^N \end{bmatrix}.$$
Show that the determinant of $\mathbf{V}$ equals
$$\det \mathbf{V} = \prod_{0 \leq k < l \leq N} (x_l - x_k).$$

4. (Foucart and Rauhut, Thm. 2.13) Given $A \in \mathbb{C}^{m \times N}$, the following statements are equivalent:

   (a) Every s-sparse vector $\mathbf{x} \in \mathbb{C}^N$ is the unique s-sparse solution of $A\mathbf{z} = A\mathbf{x}$, that is, if $A\mathbf{x} = A\mathbf{z}$ and both $\mathbf{x}$ and $\mathbf{z}$ are s-sparse, then $\mathbf{x} = \mathbf{z}$.

(b) The null space $\mathcal{N}(A)$ does not contain any $2s$-sparse vector other than the zero vector, that is,

$$\mathcal{N}(A) \cup \{\mathbf{z} \in \mathcal{C}^N : \|\mathbf{z}\|_0 \leq 2s\} = \{\mathbf{0}\}.$$

(c) For every $S \in [N]$ with $\text{card}(S) \leq 2s$, the submatrix A_S is injective as a map from $\mathbb{C}^S$ to $\mathbb{C}^m$.

(d) Every set of $2s$ columns of A is linearly independent.

5. (Foucart and Rauhut, Ex. 3.1) Let $q > 1$ and let $A \in \mathbb{C}^{m \times N}$ with $m < N$. Prove that there exists a 1-sparse vector that is *not* a minimizer of

$$\min_{\mathbf{z} \in \mathbb{C}^N} \|\mathbf{z}\|_q \text{ subject to } A\mathbf{z} = \mathbf{y}.$$

6. (Foucart and Rauhut, Ex. 3.2) Given $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}$, show that the vector $\mathbf{x} = \left[1, e^{i2\pi/3}, e^{i4\pi/3}\right]^T$ is the unique solution to

$$\min_{\mathbf{z} \in \mathbb{C}^N} \|\mathbf{z}\|_1 \text{ subject to } Az = A\mathbf{x}.$$

This shows that, in the complex setting, a unique ℓ_1 minimizer is not necessarily m-sparse, where m is the number or rows of A.

Notation for the remaining problems: $\mathbf{x} \in \mathbb{R}^N, \mathbf{y} \in \mathbb{R}^m, A \in \mathbb{R}^{m \times N}$. x_i is the i^{th} entry of a vector $\mathbf{x}$.

7. Suppose $g(x)$ is a monotonically increasing, strictly concave function of $x \in \mathbb{R}$ for $x \geq 0$, and that $g(0)$ is bounded. Let $\mathbf{x} = \mathbf{y} - \mathbf{z}$ for some $\mathbf{y} \geq 0, \mathbf{z} \geq 0$, both in $\mathbb{R}^N$. Show that there exists a one-to-one mapping between the local minima of the following two problems, and that the objective functions coincide at the corresponding local minima:

$$\min_{\mathbf{x}} \sum_{i=1}^{N} g(|x_i|) \text{ s.t. } \mathbf{y} = A\mathbf{x}$$

and

$$\min_{\mathbf{y},\mathbf{z}} \sum_{i=1}^{N} \left(g(|y_i|) + g(|z_i|) - g(0)\right) \text{ s.t. } \mathbf{y} = [A \ -A] \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \end{bmatrix}, \mathbf{y} \geq 0, \mathbf{z} \geq 0.$$

8. Show that

$$\lim_{p \to 0} \frac{1}{p} \sum_{i=1}^{N} \left(|x_i|^p - 1\right) = \sum_{i=1}^{N} \log x_i.$$

9. Show that

$$\log(|x_i| + \epsilon) \leq \frac{x_i^2}{\gamma} + \log\left(\frac{(\epsilon^2 + 2\gamma_i)^{\frac{1}{2}} + \epsilon}{2}\right) - \frac{\left[(\epsilon^2 + 2\gamma_i)^{\frac{1}{2}} - \epsilon\right]^2}{4\gamma_i}$$

for all $\epsilon, \gamma_i > 0$, with equality iff $\gamma_i = x_i^2 + \epsilon|x_i|$. Use this to derive an iteratively reweighted ℓ_2 algorithm for recovery of $\mathbf{x}$ from $\mathbf{y} = A\mathbf{x} + \mathbf{n}$.

10. If $\mathbf{y} = A\mathbf{x} + \mathbf{n}$, where $\mathbf{n}$ is Gaussian with zero mean and covariance matrix $\Sigma_n \in \mathbb{R}^{m \times m}$, $\mathbf{x}$ is Gaussian with zero mean and covariance Γ, derive the conditional distribution $p(\mathbf{x}|\mathbf{y})$. What if Σ_n is rank deficient?

11. Show that

$$g_{\text{SBL}}(\mathbf{x}) \triangleq \min_{\gamma \geq 0} \mathbf{x}^T \Gamma^{-1} \mathbf{x} + \log\det\left(\sigma^2 \mathbf{I} + A\Gamma A^T\right)$$

(where $\Gamma = \text{diag}(\gamma)$) is a nondecreasing, concave function of $|\mathbf{x}| = [|x_1|, |x_2|, \ldots, |x_N|]^T$.