ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Chair for Mathematical Information Science

Prof. Dr. H. Bölcskei
Sternwartstrasse 7
CH-8092 Zürich

Mathematics of Information

Spring semester 2022

Problem Set 6

Problem 1 Exam 2020

In this problem you may use—without proof—the following form of the Hölder inequality:

$$|\langle x, y\rangle| \leq \|x\|_1 \|y\|_\infty, \quad \text{for all } x, y \in \mathbb{C}^N,$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $\mathbb{C}^N$.

For a vector $u \in \mathbb{C}^N$, a matrix $B \in \mathbb{C}^{m\times N}$, and a set $S \subset \{1, \ldots, N\}$, we define $u_S \in \mathbb{C}^{|S|}$ to be the vector obtained from u by keeping only the entries indexed by S, and similarly, we define $B_S \in \mathbb{C}^{m\times|S|}$ to be the matrix obtained from B by keeping only the columns indexed by S. We write $\operatorname{im}(B)$ and $\ker(B)$ for the column span and the nullspace of B, respectively. We additionally define the complex sign of u as the vector $\operatorname{sgn}(u) \in \mathbb{C}^N$ given by

$$\left(\operatorname{sgn}(u)\right)_k = \begin{cases} u_k/|u_k|, & \text{if } u_k \neq 0 \\ 0, & \text{else} \end{cases}.$$

In the remainder of the problem we fix an $x \in \mathbb{C}^N \setminus \{0\}$, let $S = \{k \in [N] : x_k \neq 0\}$ be its support, and define $\overline{S} = \{1, \ldots, N\} \setminus S$. We also fix a matrix $A \in \mathbb{C}^{m\times N}$ with columns $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_N \in \mathbb{C}^m$, assume that $A_S^{\mathsf{H}} A_S \in \mathbb{C}^{|S|\times|S|}$ is invertible (where A_S^{H} denotes the Hermitian transpose of A_S), and denote

$$\alpha := \left\|\left(A_S^{\mathsf{H}} A_S\right)^{-1}\right\|_2 \cdot \max_{k\in\overline{S}} \left\|A_S^{\mathsf{H}}\, \boldsymbol{a}_k\right\|_2,$$

where $\|\cdot\|_2$ in the first factor denotes the matrix operator norm with respect to the ℓ^2-norm on $\mathbb{C}^{|S|}$, i.e.,

$$\left\|\left(A_S^{\mathsf{H}} A_S\right)^{-1}\right\|_2 := \max_{\substack{u\in\mathbb{C}^{|S|} \\ \|u\|_2\leq 1}} \left\|\left(A_S^{\mathsf{H}} A_S\right)^{-1} u\right\|_2.$$

a) Let $v \in \mathbb{C}^N$ be arbitrary and write $z = x - v$.

i) Show that $\|x\|_1 = \langle x, \operatorname{sgn}(x)\rangle$.

ii) Establish

$$\|x\|_1 \leq \|z\|_1 - \|v_{\overline{S}}\|_1 + |\langle v, \operatorname{sgn}(x)\rangle|.$$

[Hint: Use $\|\operatorname{sgn}(x_S)\|_\infty = 1$ and $\|v_{\overline{S}}\|_1 = \|z_{\overline{S}}\|_1$.]

b) Let $v \in \ker(A)$ be arbitrary.

i) Show that $A_S v_S + A_{\overline{S}} v_{\overline{S}} = 0$.

ii) Establish $\|v_S\|_2 \leq \alpha\, \|v_{\overline{S}}\|_1$.

[Hint: Note that $v_S = \left(A_S^{\mathsf{H}} A_S\right)^{-1} A_S^{\mathsf{H}} A_S\, v_S$ and $A_{\overline{S}}\, v_{\overline{S}} = \sum_{k \in \overline{S}} \boldsymbol{a}_k v_k$.]

c) Next, suppose that $v \in \ker(A)$ and $u \in \operatorname{im}(A^{\mathsf{H}})$. Show that

$$|\langle v, \operatorname{sgn}(x)\rangle| \leq \|u_S - \operatorname{sgn}(x_S)\|_2\, \|v_S\|_2 \; + \; \|u_{\overline{S}}\|_\infty\, \|v_{\overline{S}}\|_1.$$

[Hint: Recall that $\operatorname{im}(A^{\mathsf{H}}) = (\ker(A))^\perp$.]

d) Suppose that there exists a $u \in \operatorname{im}(A^{\mathsf{H}})$ satisfying the following condition:

$$\alpha\, \|u_S - \operatorname{sgn}(x_S)\|_2 + \|u_{\overline{S}}\|_\infty < 1.$$

i) Use the results of parts (b) and (c) to show that

$$|\langle v, \operatorname{sgn}(x)\rangle| < \|v_{\overline{S}}\|_1, \quad \text{for all } v \in \ker(A) \setminus \{0\}.$$

ii) Use (a)(ii) and (d)(i) to show that x is the unique solution to the following optimization problem:

$$\min_{z \in \mathbb{C}^N} \|z\|_1 \quad \text{subject to} \quad Az = Ax.$$

Problem 2 Winter Exam 2021/2022 ☕☕☕

Recall the sparse signal recovery procedure

$$\text{(P1)} \qquad \widehat{x} = \arg\min \|\widetilde{x}\|_1 \;\text{ subject to } y = D\widetilde{x},$$

with observation vector $y \in \mathbb{R}^m$ and measurement matrix $D \in \mathbb{R}^{m \times n}$, where $m < n$. In this problem, we are concerned with recovering vectors that are *almost* sparse.

To this end, we define with $s \in \mathbb{N}$, for given $x \in \mathbb{R}^n$,

$$\sigma_s(x) := \inf\{\|x - z\|_1 \mid z \in \mathbb{R}^n,\ \|z\|_0 \leq s\}.$$

Further, for $D \in \mathbb{R}^{m \times n}$ with $m < n$, define for $s < \operatorname{spark}(D)$,

$$\Delta_s(D) = \max_{\substack{S \subset [n] \\ |S| = s}} \max_{v \in \ker(D) \setminus 0} \frac{\|v_S\|_1}{\|v_{S^c}\|_1},$$

where $\ker(D) = \{v \in \mathbb{R}^n \mid Dv = 0\}$, $v_S \in \mathbb{R}^n$ denotes the vector obtained from v according to

$$(v_S)_i = \begin{cases} v_i, & i \in S \\ 0, & i \notin S \end{cases},$$

and S^c stands for the complement of the set S in $[n] = \{1, \ldots, n\}$. You may assume throughout that $\Delta_s(D)$ is well-defined, i.e., that there are S with $|S| < \operatorname{spark}(D)$ and v that achieve the maximum.

a) Prove that if $s < \operatorname{spark}(D)$, then for every set $S \subset [n]$ with $|S| = s$, it holds that

$$\|v_{S^c}\|_1 \neq 0, \qquad \forall v \in \ker(D) \setminus 0.$$

b) Prove the inequality

$$\|(x-z)_{S^c}\|_1 \leq 2\|x_{S^c}\|_1 - \|x\|_1 + \|(x-z)_S\|_1 + \|z\|_1 , \qquad \text{for } x, z \in \mathbb{R}^n,\ S \subset [n].$$

c) Fix $x \in \mathbb{R}^n$, $D \in \mathbb{R}^{m\times n}$, $s < \operatorname{spark}(D)$, and assume that $\Delta_s(D) \in (0,1)$. Prove that every solution $\widehat{x}$ of (P1) with $y = Dx$ approximates x to within error

$$\|x - \widehat{x}\|_1 \leq 2\,\frac{1+\Delta_s(D)}{1-\Delta_s(D)}\,\sigma_s(x).$$

Hint: You may use the results from subproblems (a) and (b).

d) Fix $D \in \mathbb{R}^{m\times n}$, $s < \operatorname{spark}(D)$, and assume that $\Delta_s(D) \in (0,1)$. Show that one can find $x \in \mathbb{R}^n$ and $z \in \mathbb{R}^n$, with $\|x\|_1 = \|z\|_1$ and $Dx = Dz$, such that

$$\|x - z\|_1 = 2\,\frac{1+\Delta_s(D)}{1-\Delta_s(D)}\,\sigma_s(x).$$

Problem 3 Winter Exam 2020/2021 ☕☕☕

In this problem, for a finite set A, we denote by $\operatorname{card}(A) \in \mathbb{N}_0$ the cardinality of A. One way to define *compressibility* of a vector $x \in \mathbb{C}^N$ is to say that x is *compressible* if the number

$$\operatorname{card}\left(\{j \in \{1,\dots,N\} \colon |x_j| \geq t\}\right)$$

of its significant components is small. This leads to the idea of quantifying compressibility through the following quasinorm on $\mathbb{C}^N$

$$\|x\|_{2,\infty} = \inf\left\{ M \geq 0\colon\ \operatorname{card}\left(\{j \in \{1,\dots,N\} \colon |x_j| \geq t\}\right) \leq \frac{M^2}{t^2},\ \text{for all } t > 0 \right\}.$$

In this problem, we will verify that $\|x\|_{2,\infty}$, indeed, constitutes a quasinorm, i.e., it satisfies the norm axioms, except for the triangle inequality which is replaced by

$$\|x+y\|_{2,\infty} \leq K(\|x\|_{2,\infty} + \|y\|_{2,\infty}) \tag{1}$$

for some $K > 0$. We will also compare the properties of $\|x\|_{2,\infty}$ with those of the usual ℓ_2-norm.

a) Show that for every $x \in \mathbb{C}^N$ and $\lambda \in \mathbb{C}$,
   i) $\|x\|_{2,\infty} = 0 \implies x = 0$,
   ii) $\|\lambda x\|_{2,\infty} = |\lambda|\|x\|_{2,\infty}$.

b) Next, we show that the quasinorm $\|x\|_{2,\infty}$ does not satisfy the triangle inequality. To this end, let us fix $x = (1, 2^{-1/2})$ and $y = (2^{-1/2}, 1)$.
   i) Calculate $\|x\|_{2,\infty}$ and $\|y\|_{2,\infty}$.
   ii) Calculate $\|x+y\|_{2,\infty}$ and use (b)(i) to conclude that in fact

$$\|x+y\|_{2,\infty} > \|x\|_{2,\infty} + \|y\|_{2,\infty}.$$

Hint: Use (a)(ii) to simplify your calculations.

c) Next, we establish (1) by proving a more general result. To this end, let us fix $x^1 = (x_1^1, \dots, x_N^1), x^2 = (x_1^2, \dots, x_N^2), \dots, x^k = (x_1^k, \dots, x_N^k) \in \mathbb{C}^N$ and $t > 0$.

i) Show that

$$\{j \in \{1, \dots, N\} \colon |x_j^1 + \dots + x_j^k| \geq t\} \subset \bigcup_{i \in \{1, \dots, k\}} \{j \in \{1, \dots, N\} \colon |x_j^i| \geq t/k\}.$$

ii) Use (c)(i) to prove that

$$\|x^1 + \dots + x^k\|_{2,\infty} \leq k\big(\|x^1\|_{2,\infty}^2 + \dots + \|x^k\|_{2,\infty}^2\big)^{1/2},$$

where $\|x^i\|_{2,\infty}^2$ stands for $(\|x^i\|_{2,\infty})^2$.

iii) Now employ (c)(ii) to show that

$$\|x^1 + \dots + x^k\|_{2,\infty} \leq k\big(\|x^1\|_{2,\infty} + \dots + \|x^k\|_{2,\infty}\big).$$

d) Next, we compare the quasinorm $\|x\|_{2,\infty}$ with the usual ℓ_2-norm. To this end, let us fix $x = (x_1, \dots, x_N) \in \mathbb{C}^N$ and assume that

$$\|x\|_{2,\infty} = \max_{k \in \{1, \dots, N\}} k^{1/2} x_k^*, \tag{2}$$

where $x^* \in \mathbb{R}_+^N$ denotes the nonincreasing rearrangement of $(|x_1|, \dots, |x_N|) \in \mathbb{R}_+^N$. Use (2) to prove that

$$\|x\|_{2,\infty} \leq \|x\|_2.$$