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

Solutions to Problem Set 6

Problem 1 Exam 2020 ☕☕☕

(a) (i) We have

$$\|x\|_1 = \sum_{k=1}^{N} |x_k| = \sum_{k\in S} |x_k| = \sum_{k\in S} x_k \cdot \frac{\overline{x_k}}{|x_k|}$$
$$= \sum_{k\in S} x_k \,\overline{\left(\operatorname{sgn}(x)\right)_k} = \sum_{k=1}^{N} x_k \,\overline{\left(\operatorname{sgn}(x)\right)_k} = \langle x, \operatorname{sgn}(x)\rangle,$$

as desired.

(ii) Using the triangle inequality, the Hölder inequality, and $\|z\|_1 = \|z_S\|_1 + \|z_{\overline{S}}\|_1$, we obtain

$$\begin{aligned}
\|x\|_1 &= \langle x, \operatorname{sgn}(x)\rangle = \langle z + v, \operatorname{sgn}(x)\rangle \\
&\le |\langle z, \operatorname{sgn}(x)\rangle| + |\langle v, \operatorname{sgn}(x)\rangle| = |\langle z_S, \operatorname{sgn}(x_S)\rangle| + |\langle v, \operatorname{sgn}(x)\rangle| \\
&\le \|z_S\|_1 \underbrace{\|\operatorname{sgn}(x_S)\|_\infty}_{=1} + |\langle v, \operatorname{sgn}(x)\rangle| = \|z\|_1 - \|z_{\overline{S}}\|_1 + |\langle v, \operatorname{sgn}(x)\rangle| \\
&= \|z\|_1 - \|v_{\overline{S}}\|_1 + |\langle v, \operatorname{sgn}(x)\rangle|.
\end{aligned}$$

(b) (i) We have

$$0 = Av = \sum_{k=1}^{N} \boldsymbol{a}_k v_k = \sum_{k\in S} \boldsymbol{a}_k v_k + \sum_{k\in \overline{S}} \boldsymbol{a}_k v_k = A_S v_S + A_{\overline{S}}\, v_{\overline{S}}.$$

(ii) First note that

$$A_S v_S = -A_{\overline{S}}\, v_{\overline{S}} = -\sum_{k\in\overline{S}} \boldsymbol{a}_k v_k.$$

Using this, the definition of the matrix operator norm, the triangle inequality, and the definition of α, we can argue as follows

$$\begin{aligned}
\|v_S\|_2 &= \|(A_S^{\mathsf{H}}A_S)^{-1}A_S^{\mathsf{H}}A_Sv_S\| \le \|(A_S^{\mathsf{H}}A_S)^{-1}\|_2\|A_S^{\mathsf{H}}A_Sv_S\|_2 \\
&= \|(A_S^{\mathsf{H}}A_S)^{-1}\|_2\Big\|A_S^{\mathsf{H}}\sum_{k\in\overline{S}}\boldsymbol{a}_kv_k\Big\|_2 = \|(A_S^{\mathsf{H}}A_S)^{-1}\|_2\Big\|\sum_{k\in\overline{S}}A_S^{\mathsf{H}}\boldsymbol{a}_kv_k\Big\|_2 \\
&\le \|(A_S^{\mathsf{H}}A_S)^{-1}\|_2\sum_{k\in\overline{S}}\|A_S^{\mathsf{H}}\boldsymbol{a}_k\|_2|v_k| \\
&\le \|(A_S^{\mathsf{H}}A_S)^{-1}\|_2\cdot\max_{k\in\overline{S}}\|A_S^{\mathsf{H}}\boldsymbol{a}_k\|_2\cdot\sum_{k\in\overline{S}}|v_k| \\
&= \alpha\|v_{\overline{S}}\|_1,
\end{aligned}$$

yielding the desired identity.

(c) As $v\in\ker(A)$ and $u\in\operatorname{im}(A^{\mathsf{H}})=(\ker(A))^{\perp}$, we have $\langle v,u\rangle=0$, and so

$$\begin{aligned}
|\langle v,\operatorname{sgn}(x)\rangle| &= |\langle v,\operatorname{sgn}(x)-u\rangle| = |\langle v_S,\operatorname{sgn}(x_S)-u_S\rangle+\langle v_{\overline{S}},\underbrace{\operatorname{sgn}(x_{\overline{S}})}_{=0}-u_{\overline{S}}\rangle| \\
&\le |\langle v_S,\operatorname{sgn}(x_S)-u_S\rangle|+|\langle v_{\overline{S}},-u_{\overline{S}}\rangle| \\
&\le \|u_S-\operatorname{sgn}(x_S)\|_2\|v_S\|_2+\|u_{\overline{S}}\|_\infty\|v_{\overline{S}}\|_1,
\end{aligned}$$

where in the last step we used the Cauchy-Schwarz and the Hölder inequality.

(d) (i) Let $v\in\ker(A)\setminus\{0\}$. We now have

$$\begin{aligned}
|\langle v,\operatorname{sgn}(x)\rangle| &\le \|\operatorname{sgn}(x_S)-u_S\|_2\|v_S\|_2+\|u_{\overline{S}}\|_\infty\|v_{\overline{S}}\|_1 \\
&\le \|\operatorname{sgn}(x_S)-u_S\|_2\cdot\alpha\|v_{\overline{S}}\|_1+\|u_{\overline{S}}\|_\infty\|v_{\overline{S}}\|_1 \\
&= (\alpha\,\|\operatorname{sgn}(x_S)-u_S\|_2+\|u_{\overline{S}}\|_\infty)\cdot\|v_{\overline{S}}\|_1,
\end{aligned}$$

and hence the desired inequality follows by the condition

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

provided we can show that $\|v_{\overline{S}}\|_1>0$ (this is necessary as we have to prove a strict inequality between $|\langle v,\operatorname{sgn}(x)\rangle|$ and $\|v_{\overline{S}}\|_1$).

Indeed, if $\|v_{\overline{S}}\|_1=0$, we would have $\|v_S\|_2\le\alpha\|v_{\overline{S}}\|_1=0$, and so $v_{\overline{S}}=0$ and $v_S=0$. This, however, stands in contradicton to the assumption $v\neq0$, and hence we must have $\|v_{\overline{S}}\|_1>0$, as desired.

(ii) Let $z\in\mathbb{C}^N\setminus\{x\}$ be such that $Az=Ax$, but otherwise arbitrary, and set $v=x-z$. Then $v\neq0$ and $Av=Ax-Az=0$, and thus $v\in\ker(A)\setminus\{0\}$. Hence

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

As z was arbitrary, this establishes that x is the unique minimizer of

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

Problem 2 Winter Exam 2021/2022 ☕☕☕

a) Towards a contradiction, we assume that there are a set $S \subset [n]$, with $|S| = s$, and a vector $v \in \ker(D)\backslash 0$, such that $\|v_{S^c}\|_1 = 0$. This implies that v is supported on S exclusively and hence $Dv_S = 0$ with $v_S \neq 0$. The contradiction is now established by noting that for $s < \text{spark}(D)$ (which is by assumption), $Dv_S = 0$ with $v_S \neq 0$ is not possible as $\text{spark}(D)$ is the smallest number of linearly dependent columns of D.

b) The proof is by the following chain of relations

$$\begin{aligned}
\|(x-z)_{S^c}\|_1 &\leq \|x_{S^c}\|_1 + \|z_{S^c}\|_1 \\
&= \|x_{S^c}\|_1 + \|x\|_1 - \|x\|_1 + \|z_{S^c}\|_1 \\
&= 2\|x_{S^c}\|_1 + \|x_S\|_1 - \|x\|_1 + \|z_{S^c}\|_1 \\
&= 2\|x_{S^c}\|_1 - \|x\|_1 + \|(x-z+z)_S\|_1 + \|z_{S^c}\|_1 \\
&\leq 2\|x_{S^c}\|_1 - \|x\|_1 + \|(x-z)_S\|_1 + \|z\|_1,
\end{aligned} \tag{1}$$

where we used the triangle inequality twice.

c) Let $S \subset [n]$ with $|S| = s$ be the indices of the s largest absolute values of x. First, we note that

$$\sigma_s(x) = \|x_{S^c}\|_1. \tag{2}$$

Second, $\widehat{x}$, by virtue of being a solution of (P1) with $y = Dx$, fulfills $\|\widehat{x}\|_1 \leq \|x\|_1$ and

$$D\widehat{x} = y = Dx \quad \Leftrightarrow \quad D(x - \widehat{x}) = 0.$$

Hence, the vector $v := x - \widehat{x}$ lies in the kernel of D. We have $s < \text{spark}(D)$ by assumption and thus, by subproblem (a), $\|v_{S^c}\|_1 \neq 0$. Further, we have

$$\frac{\|v_S\|_1}{\|v_{S^c}\|_1} \leq \Delta_s(D). \tag{3}$$

Next, we bound

$$\begin{aligned}
\|v_{S^c}\|_1 &\leq 2\|x_{S^c}\|_1 - \|x\|_1 + \|v_S\|_1 + \|\widehat{x}\|_1 \\
&\leq 2\|x_{S^c}\|_1 + \|v_S\|_1 \\
&\leq 2\|x_{S^c}\|_1 + \Delta_s(D)\|v_{S^c}\|_1,
\end{aligned} \tag{4}$$

where the first inequality follows from (1) by setting $z = \widehat{x}$, the second is by $\|\widehat{x}\|_1 \leq \|x\|_1$, and the third is due to (3). Using (2) and $1 - \Delta_s(D) > 0$, which is by assumption, we rewrite (4) as

$$\|v_{S^c}\|_1 \leq 2\,\frac{1}{1 - \Delta_s(D)}\,\sigma_s(x).$$

The proof is now concluded upon noting that

$$\begin{aligned}\|x-\widehat{x}\|_1 &= \|v\|_1\\ &= \|v_S\|_1 + \|v_{S^c}\|_1\\ &\leq \Delta_s(D)\|v_{S^c}\|_1 + \|v_{S^c}\|_1\\ &= (1+\Delta_s(D))\|v_{S^c}\|_1\\ &\leq 2\,\frac{1+\Delta_s(D)}{1-\Delta_s(D)}\,\sigma_s(x).\end{aligned}$$

d) Let $\gamma := \Delta_s(D) \in (0,1)$. As $\Delta_s(D)$ is well-defined, there exists a set $S \subset [n]$ with $|S| = s$ and a $v \in \ker(D)\backslash 0$ with $\gamma\|v_{S^c}\|_1 = \|v_S\|_1$. We next note, that for every $b \in \mathbb{R}^n$,

$$0 = D(v_S + v_{S^c} + b - b) \qquad \Leftrightarrow \qquad D(v_S + b) = D(b - v_{S^c}).$$

Next, let $x = v_S + b$ and $z = b - v_{S^c}$ and choose b such that $\|x\|_1 = \|z\|_1$. This can be effected by setting $b = \alpha v_{S^c}$, calculating

$$\begin{aligned}\|x\|_1 &= \|v_S\|_1 + \alpha\|v_{S^c}\|_1 = (\alpha+\gamma)\|v_{S^c}\|_1\\ \|z\|_1 &= (1-\alpha)\|v_{S^c}\|_1,\end{aligned}$$

and finally choosing $\alpha = \frac{1-\gamma}{2}$. Hence,

$$x = v_S + \frac{1-\gamma}{2}v_{S^c} \qquad \text{and} \qquad z = -\frac{1+\gamma}{2}v_{S^c}.$$

Next, we calculate $\sigma_s(x)$. To this end, we first note that $|v_i| \geq |v_j|,\ \forall i \in S, j \in S^c$ as otherwise $\frac{\|v_S\|_1}{\|v_{S^c}\|_1}$ would not be maximal. Since $\frac{1-\gamma}{2} \in (0, 1/2)$ for $\gamma \in (0,1)$, it follows that the indices of the s largest absolute values of x are given by the set S. Hence, we calculate $\sigma_s(x) = \|x_{S^c}\|_1 = \frac{1-\gamma}{2}\|v_{S^c}\|_1$, or equivalently $\frac{2}{1-\gamma}\sigma_s(x) = \|v_{S^c}\|_1$, because $\gamma < 1$ by assumption. Finally, we note that

$$\|x-z\|_1 = \|v_S + v_{S^c}\|_1 = \|v_S\|_1 + \|v_{S^c}\|_1 = (1+\gamma)\|v_{S^c}\|_1 = 2\,\frac{1+\gamma}{1-\gamma}\,\sigma_s(x).$$

Problem 3 Winter Exam 2020/2021

(a) (i) Let us assume that $\|x\|_{2,\infty} = 0$. This implies, that for every $M > 0$,

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

Hence, for every $t > 0$,

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

Consequently, by choosing $M > 0$ but arbitrarily small, we get

$$\operatorname{card}(\{j \in \{1,\dots,N\} \colon |x_j| \geq t\}) = 0, \tag{5}$$

for every $t > 0$. Now, taking t in (5) arbitrarily small, we can conclude that $x = 0$.

(ii) The statement is obvious for $\lambda = 0$. Hence, we can assume that $\lambda \neq 0$. Now, observe that

$$\{j \in \{1, \dots, N\} : |\lambda x_j| \geq t\} = \{j \in \{1, \dots, N\} : |x_j| \geq t/|\lambda|\}.$$

Therefore,

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

(b) (i) Note that

$$\operatorname{card}\big(\{j \in \{1,2\} : |x_j| \geq t\}\big) = \begin{cases} 2, & \text{if } \ t \leq 2^{-1/2}, \\ 1, & \text{if } \ 2^{-1/2} < t \leq 1, \\ 0, & \text{if } \ t > 1. \end{cases}$$

Therefore, for all $t > 0$,

$$\operatorname{card}\big(\{j \in \{1,2\} : |x_j| \geq t\}\big) \leq \frac{1}{t^2}.$$

On the other hand,

$$\operatorname{card}\big(\{j \in \{1,2\} : |x_j| \geq 1\}\big) = 1,$$

and hence, we get

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

Next, using

$$\operatorname{card}\big(\{j \in \{1,2\} : |x_j| \geq t\}\big) = \operatorname{card}\big(\{j \in \{1,2\} : |y_j| \geq t\}\big), \text{ for all } t > 0,$$

yields

$$\|y\|_{2,\infty} = \|x\|_{2,\infty} = 1.$$

(ii) Thanks to (a)(ii), we have

$$\|x + y\|_{2,\infty} = \|(1 + 2^{-1/2}, 1 + 2^{-1/2})\|_{2,\infty} = (1 + 2^{-1/2})\|(1,1)\|_{2,\infty}.$$

We are therefore left with having to calculate $\|(1,1)\|_{2,\infty}$. Let us fix $z := (1,1)$. Then

$$\operatorname{card}\big(\{j \in \{1,2\} \colon |z_j| \geq t\}\big) = \begin{cases} 2, & \text{if} \quad t \leq 1, \\ 0, & \text{if} \quad t > 1, \end{cases}$$

and hence

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

Consequently, we get

$$\|x+y\|_{2,\infty} = 2^{1/2}(1 + 2^{-1/2}) = 2^{1/2} + 1 > 2 = \|x\|_{2,\infty} + \|y\|_{2,\infty}.$$

(c) (i) If $|x_j^1 + \cdots + x_j^k| \geq t$ for some $j \in \{1, \ldots, N\}$, then we have that $|x_j^i| \geq t/k$ for this j and some $i \in \{1, \ldots, k\}$. This allows us to conclude that

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

(ii) From (c)(i) we get that

$$\begin{aligned} &\operatorname{card}\big(\{j \in \{1,\ldots,N\} \colon |x_j^1 + \cdots + x_j^k| \geq t\}\big) \\ &\leq \sum_{i \in \{1,\ldots,k\}} \operatorname{card}\big(\{j \in \{1,\ldots,N\} \colon |x_j^i| \geq t/k\}\big) \\ &\leq \sum_{i \in \{1,\ldots,k\}} \frac{\|x^i\|_{2,\infty}^2}{(t/k)^2} \\ &= \frac{k^2\big(\|x^1\|_{2,\infty}^2 + \cdots + \|x^k\|_{2,\infty}^2\big)}{t^2}. \end{aligned}$$

We therefore obtain

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

(iii) We have

$$\begin{aligned} &\frac{\big(\|x^1\|_{2,\infty}^2 + \cdots + \|x^k\|_{2,\infty}^2\big)}{\big(\|x^1\|_{2,\infty} + \cdots + \|x^k\|_{2,\infty}\big)^2} \\ &= \left(\frac{\|x^1\|_{2,\infty}}{\big(\|x^1\|_{2,\infty} + \cdots + \|x^k\|_{2,\infty}\big)}\right)^2 + \cdots + \left(\frac{\|x^k\|_{2,\infty}}{\big(\|x^1\|_{2,\infty} + \cdots + \|x^k\|_{2,\infty}\big)}\right)^2 \\ &\leq \frac{\|x^1\|_{2,\infty}}{\big(\|x^1\|_{2,\infty} + \cdots + \|x^k\|_{2,\infty}\big)} + \cdots + \frac{\|x^k\|_{2,\infty}}{\big(\|x^1\|_{2,\infty} + \cdots + \|x^k\|_{2,\infty}\big)} = 1, \end{aligned}$$

and hence

$$\big(\|x^1\|_{2,\infty}^2 + \cdots + \|x^k\|_{2,\infty}^2\big)^{1/2} \leq \|x^1\|_{2,\infty} + \cdots + \|x^k\|_{2,\infty}. \tag{6}$$

Combining this with (c)(ii), we obtain

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

(d) For every $k \in \{1, \ldots, N\}$, we can write

$$\|x\|_2^2 = \sum_{j=1}^{N} (x_j^*)^2 \geq \sum_{j=1}^{k} (x_j^*)^2 \geq k(x_k^*)^2.$$

Raising to the power $1/2$ and taking the maximum over k yields the desired result.