SPARSE RECOVERY AND THE GEOMETRY OF HIGH-DIMENSIONAL RANDOM MATRICES – EXERCISES

B. G. BODMANN
GENE GOLUB SIAM SUMMER SCHOOL 2017

Day 1: Deterministic versus random matrices with the restricted isometry property

Exercise 1. Instead of a direct construction of an $m \times n$ (sensing) matrix A whose *column* vectors are

$$A = [a_1 a_2 \cdots a_n]$$

we may attempt to find a matrix $G = A^* A$, whose entries contain inner products of these column vectors.

(1) State a condition that is equivalent to A having RIP constant δ_s in terms of the spectra of the principal $s \times s$ submatrices of G.
(2) Examine the paper `https://arxiv.org/abs/math/0406134`, where five examples of matrices are presented that possess for $n = 36$ and $m = 15$ optimal bounds for δ_s and smallest values of s. (Note that the matrices need to be normalized by multiplying with an appropriate constant.) If we wish to use the result by Candès requiring $\delta_{2s} \leq \sqrt{2} - 1$, what degree of sparsity in signals can be recovered with sensing matrices obtained by factoring G? You can verify the stated properties with the example of a matrix in this paper, saved in the file `FGE-G36rk15.mat` at `http://www.math.uh.edu/∼bgb/GGSSS2017`.

Exercise 2. Compare numerically obtained RIP constants using sensing matrices with Gaussian i.i.d. entries with the number of measurements required in the paper `DeVoreCompressedSensing.pdf` at `http://www.math.uh.edu/∼bgb/GGSSS2017`.

Exercise 3. Compare numerically obtained RIP constants using sensing matrices with Gaussian i.i.d. entries with the provable performance guarantees in Section 1.6 of the paper `DecodingLP.pdf` at `http://www.math.uh.edu/∼bgb/GGSSS2017`.

Day 2: RIP and Robust width

Exercise 1. It is tempting to consider matrices A whose column vectors can be partitioned into orthonormal bases. A family of vectors $\{a_i^{(r)} : 1 \le r \le b, 1 \le i \le m\}$ is called a family of mutually unbiased bases if

$$\langle a_i^{(r)}, a_{i'}^{(r')} \rangle = \begin{cases} 1/\sqrt{m} & , r \neq r' \\ 0 & , i \neq i', r = r' \\ 1 & , i = i', r = r' \end{cases} .$$

It is known that if $m = 4^p$ for some p, then we can choose $b = m/2$, see the paper `https://arxiv.org/abs/quant-ph/0502024v2`. The Gram matrix corresponding to such an A made up of these column vectors has all off-diagonal entries bounded by $1/\sqrt{m}$. Using the bound for the operator norm in Theorem 5.3 of the paper `FramesGraphsErasures.pdf` considered before, show that the restricted isometry constant is bounded by $\delta_s \le (s-1)/\sqrt{m}$. What is the maximal s for which the Candès proof guarantees robust and stable recovery?

Exercise 2. Prove that if there are constants C_0 and C_1 such that for every $x^\natural \in \ell_2^n$, $\epsilon \ge 0$ and $\eta \in \ell_2^m$ with $\|\eta\| \le \epsilon$, any choice

$$\hat{x} \in \arg\min\{\|x\|_1 : \|Ax - (Ax^\natural + \eta)\| \le \epsilon\}$$

satisfies that for every $a \in \Sigma_s$, $\|\hat{x} - x^\natural\| \le C_0 \|x^\natural - a\|_1 + C_1 \epsilon$, then A satisfies the (ρ, α)-robust width property for ℓ_1^n with constants $\rho = 2C_0$ and $\alpha = 1/(2C_1)$. Hint: Choose $x^\natural \in \mathcal{N}_\alpha(A)$ and let $\epsilon = \alpha\|x^\natural\|$, $\eta = 0$.

Exercise 3. In the paper by Cahill and Mixon, the authors show that if the $m \times n$ matrix A is a scaled co-isometry, so $AA^* = cI$ on ℓ_2^m with some $c > 0$, then $A^*A = cP$ with an orthogonal projection P, and the set $\mathcal{N}_\alpha(A)$ is characterized as a union of subspaces

$$\mathcal{N}_\alpha(A) \cup \{0\} = \cup\{X \subset \ell_2^n : \dim(X) = n - m, \|P_X - P\| < \alpha\},$$

where P_X is the orthogonal projection with range X. Is this set (including 0) convex?

Day 3: RIP implies robust width

Exercise 1. Relax or build a raft.

Day 4: Operator monotonicity and operator concavity

Exercise 1. Show that the function $f : t \mapsto t^2$ is not operator monotone on $\mathbb{R}^+$. It is enough to show that the matrices $A = \begin{pmatrix} 3/2 & 0 \\ 0 & 3/4 \end{pmatrix}$ and $B = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$ satisfy $A \geq B$ but not $A^2 \geq B^2$.

Exercise 2. For a random variable X with values in $[0, 1]$, the generating function satisfies the bound

$$\mathbb{E}[e^{rX}] \leq 1 + (e^r - 1)\mathbb{E}[X], \theta \geq 0\,.$$

Is the same true when X is a random positive semidefinite matrix with spectrum in $[0, 1]$? Hint: Use the spectral theorem for each realization of X, take the expectation last.

Exercise 3. Recall Weyl's eigenvalue estimate: For two Hermitian matrices A and B, let $\lambda_i(A)$ denote the i-th largest eigenvalue of A and similarly $\lambda_i(B)$ for B, then if $A \geq B$ (operator inequality), we have $\lambda_i(A) \geq \lambda_i(B)$. Use this to show that if f is monotone on $\mathbb{R}$, then the composition of f with the trace is a real-valued function $X \mapsto \mathrm{tr}[f(X)]$ that is monotone with respect to the partial order on the set of Hermitian matrices.

Exercise 4. Assuming $\mathbb{E}[\sum_{j=1}^m X_j] = I$ as in the lecture, show Talagrand's symmetrization lemma, which yields an upper bound for the expected value of the norm in terms of the symmetrized distribution,

$$\mathbb{E}\|\sum_{j=1}^m X_j - I\| \leq 2\mathbb{E}[\|\sum_j \epsilon_j X_j\|]$$

where each ϵ_j is a zero-mean Rademacher random variable with values in $\{\pm 1\}$ and the family $\{\epsilon_j, X_j : 1 \leq j \leq m\}$ is independent. Hint: First prove the inequality in case the mean is zero, so

$$\mathbb{E}\|\sum_{j=1}^m Y_j\| \leq 2\mathbb{E}[\|\sum_j \epsilon_j Y_j\|]\,.$$

Verify that for any independent (vector-valued) random variables Y, Z, with $\mathbb{E}[Z] = 0$,

$$\mathbb{E}[\|Y\|] \leq \mathbb{E}[\|Y + Z\|]\,.$$

If $\mathbb{E}[Y] = 0$, you may choose Z to be an independent copy of Y.

Day 5: Expected values vs. tail bounds

Exercise 1. Suppose we are interested in measure concentration for X_j as described in the lecture with $\mathbb{E}[\sum_{j=1}^m X_j] = I$ and do not care about logarithmic factors or precise estimates for failure estimates. Explain why it is enough to have

$$\mathbb{E}[\|\sum_{j=1}^{m} X_j - I\|] \leq \delta'$$

for $\delta' < \delta$ in order to obtain

$$\lim_{c\to\infty} \mathbb{P}[\|\sum_{j=1}^{cm} \frac{1}{c} X_j - I\| > \delta] \to 0\,.$$

Exercise 2. Given a non-negative (real-valued) random variable X with expected value E, bound the probability $\mathbb{P}[X \geq rE]$ for a given $r > 1$. Hints:

(1) Argue that among all possible distributions, it suffices to consider X with values in $[0, rE]$, because if $Y = \min\{X, rE\}$, then $\mathbb{P}[X \geq rE] = \mathbb{P}[Y \geq rE]$. Also, $\mathbb{E}[Y] \leq E$.

(2) Next, show that among all random variables with values in $[0, rE]$ and average E, the probability $\mathbb{P}[X = rE]$ is maximized if the random variable takes values in $\{0, rE\}$ only.

(3) Finally, use the normalization of the probability measure and the assumption on the values in $\{0, rE\}$ to express $\mathbb{P}[X = rE]$ in terms of E.