Laboratorium 2

Aleksander Milach

02/04/2020

Exercise 1

Let $A \in \mathbb{R}^{3\times 4}$ and $b \in \mathbb{R}^3$ as follows

$$A := \begin{pmatrix} 1 & 0 & 0 & 1/\sqrt{3} \\ 0 & 1 & 0 & 1/\sqrt{3} \\ 0 & 0 & 1 & 1/\sqrt{3} \end{pmatrix} \quad and \quad b := \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

A vector $x^* = (1, 1, -1, 0)$ is a solution of $Ax = b$. One may notice, that since $supp(x) = \{1, 2, 3\}$, $A_I = A_I^{-1} = (A_I^T A_I)^{-1} = I$, thus $A_{\bar{I}}^T A_I (A_I^T A_I)^{-1} sign(x_I^*) = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})(1, 1, -1)^T = 1/\sqrt{3} < 1$, hence x^* satisfies irrepresentable condition.

However $\tilde{x} = (0, 0, -2, \sqrt{3})$ is a different solution of Ax = b, while it is more sparse than x^*, therefore x^* is not one of the sparsest solutions.

Exercise 2

Let us set a patricular seed and generate a 200×500 Gaussian 'random' matrix B with i.i.d normal entries, and let us define matrix A as the matrix B with normalized columns.

1)

Using the definition of mutual coherence of a matrix, we calculate that $M(A) = 0.3021009$. According to one of the propositions, a solution of a linear system is both the unique sparsest solution and the unique BP minimizer, if its l_0 norm is strictly smaller than $\frac{1+1/M(A)}{2} = 2.1550764$. Thus if a solution has two or less non - null components, then it is both the unique sparsest solution and a unique BP minimizer. Hence the set $K_0 = \{1, 2\}$.

2)

To check for which $k \in \{1, ..., 200\}$ the irrepresentable condition holds, we can write a loop, that calculates a value on the left hand side of that condition and after that check for which values of k it is smaller than 1 (for example by using a which function), after that we obtain a set K_1.

```
##  [1]  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
```

3)

One can notice, that $sign(\tilde{x}) = sign(x^*)$ and $supp(\tilde{x}) = supp(x^*)$. Moreover, from characterization of BP minimizer, a solution x is a unique BP minimizer if, and only if for each non - zero vector h from $ker(A)$ $|\sum\limits_{i \in supp(x)} sign(x_i)h_i| < \sum\limits_{i \notin supp(x)} |h_i|$, thus a solution of a linear system is a unique BP minimizer by depending on its sign vector and support set. Hence $\tilde{x}$ is a unique BP minimizer if, and only if x^* is a unique BP minimizer.

4)

Based on the corrolary from task 3, we can calculate a set of sparsities $k \in K_2$ for which $\tilde{x}$ is a BP minimizer of $Ax = \tilde{b}$, since the it is the same set as the set of sparsities for which x^* is a BP minimizer of $Ax = b$. In this case we check in loop for which $k \in \{1, ..., 200\}$ a numerically obtained BP minimizer is equal to $(1, ..., 1, 0, ..., 0)$, where k is a number of ones in this vector.

```
##  [1]  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
## [24] 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46
## [47] 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63
```

We see that $K_0 \subset K_1 \subset K_2$, since mutual coherence condition implies irrepresentable condidition, and irreperesnetable condition implies a solution being a BP minimizer.

5)

From the propositions concerning spark of a matrix, we know that if for a solution x_1, $||x_1||_0 <$ spark(A) / 2, then x_1 is a unique sparsest solution. Additionally, spark of a matrix is always smaller or equal $m + 1$, where $m = min(n, p)$, where n and p are number of rows and columns of the matrix, thus in our case its $m = 200$. Moreover we know that the set of real matrices with 200 rows and a particular number of columns (eg. 200), which spark is smaller or equal 200 is negligable with respect to Lebesgue measure. Thus we can assume that spark(A) = 201, thus x_1 is the unique sparsest solution if $||x_1||_0 \leq 100$.

From previous point, we know that a solution x_2 is a unique BP mimizer if $||x_2||_0 \leq 63$, so in this case it is a unique sparsest solution because automatically $||x_2||_0 \leq 100$ too, thus our set $K_3 = \{1, ..., 63\}$.

A set K_0 is a lot smaller, then a set K_3, while both sets are sets of sparsities for which a solution of a linear system is a unique sparsest solution and a unique BP minimizer. Thus we notice, that the mutual coherence condition is a very strong condition for checking, whether a solution is both a unique sparsest solution and a unique BP minimizer.

Exercise 6

Let $A := (A_1|...|A_p) \in \mathbb{R}^{n\times p}$ where $||A_1||_2 = ... = ||A_p||_2 = 1$, $b \in col(A)$ and x^* be a solution of the linear system $Ax = b$. Let $I := supp(x^*), A_I$ be the matrix, whose columns are respectively $(A_i)_{i\in I}$ and $sign(x_I^*) = (sign(x_i^*))_{i\in I}$. Let us prove the following implication

$$||x^*||_0 \leq \frac{1 + 1/M(A)}{2} \implies \forall j \notin I, |A_j^T A_I (A_I^T A_I)^{-1} sign(x_I^*)| \leq 1.$$

First, let us prove the following result on localization of eigenvalues.

Let Q be a $n \times n$ matrix, with entries Q_{ij}. For $i \in \{1, ..., n\}$ let R_i be defined as $R_i = \sum_{j\neq i} |Q_{ij}|$, then every eigenvalue λ of Q, $|\lambda - Q_{ii}| < R_i$.

Let λ be an eigenvalue of Q. Let us choose a corresponding eigenvector x so that one its component x_i is equal to 1 and the others are of absolute value less or equal to 1, thus $x_i = 1$ and $|x_j| \leq 1$ for $j \neq i$. Since $Qx = \lambda x$, in particular $\sum_j Q_{ij}x_j = \lambda x_i = \lambda$, which can be rewritten as $\sum_{j\neq i} Q_{ij}x_j = \lambda - Q_{ii}$. Now by applying a triangle inequality

$$|\lambda - Qii| = |\sum_{j\neq i} Q_{ij}x_j| \leq \sum_{j\neq i} |Q_{ij}||x_j| \leq |\sum_{j\neq i} |Q_{ij}| = R_i,$$

which gives a desired bound.

Now, let $k = ||x^*||_0$ and $Q = A_I^T A_I \in \mathbb{R}^{k\times k}$. Since Q is a Gram matrix, its diagonal elements are equal to 1, because A is a normalized matrix, while non diagonal elements are, in absolute value, smaller than M(A), thus by the previous result, each eiqenvalue λ of Q satisfies the inequalities

$$1-(k-1)M(A) \leq \lambda \leq 1+(k-1)M(A).$$

From previous result we obtain:

$$\frac{1}{1+(k-1)M(A)} \leq \frac{1}{\lambda} \leq \frac{1}{1-(k-1)M(A)}.$$

Since if λ is an eigenvalue of matrix B, then $1/\lambda$ is an eigenvalue of matrix B^{-1}. By restricting k with $\frac{1+1/M(A)}{2}$ and rewriting the second inequality, after simple transformations we get that each eigenvalue of $(A_I^T A_I)^{-1}$ is smaller than $\frac{2}{1+M(A)}$.

Now, one may notice that $A_j^T A_I (A_I^T A_I)^{-1} sign(x_I^*)$ is the scalar product between $A_I^T A_j$ and $(A_I^T A_I)^{-1} sign(x_I^*)$, thus the following inequality holds

$$|A_j^T A_I (A_I^T A_I)^{-1} sign(x_I^*)| \leq ||A_I^T A_j||_2 ||(A_I^T A_I)^{-1} sign(x_I^*)||_2 \leq$$

$$M(A)\sqrt{k}\sqrt{k}||(A_I^T A_I)^{-1}\frac{sign(x_I^*)}{\sqrt{k}}||_2 \leq M(A)k\frac{2}{M(A)+1} \leq$$

$$M(A)\frac{1+1/M(A)}{2}\frac{2}{M(A)+1} = 1.$$

In the inequalities above we restrict each of the components of $A_j^T A_I$ with M(A), since $j \notin supp(x)$, thus the norm of the first vector is smaller than $\sqrt{k}M(A)$. To restrict the second vector, we normalize $sign(x^*)$, so after that we can use a property of norm of a linearly transformed vector, which maximum value is the largest eigenvalue of the matrix of transformation (this property can be proven by calculating conditional maximum of function $x^T(A^T A)x$ with condition $||x||_2 = 1$). After that we restrict k with its assumed upper bound and the largest eigenvalue of $A_j^T A_I$ with upper bound calculated earlier, we obtain a final implication.