Homework 4

MA430

As always, give full justification for your answers!

1. Suppose I have 10 marbles, nominal mass 10 grams each, but a few marbles may be defective and have mass that's slightly off. Let x_i denote the deviation of the ith marble from nominal. I weigh a subset consisting of marbles $\{3, 6, 10\}$ and get 30.13. I can conclude that

$$x_3 + x_6 + x_{10} = 0.13$$

(the set had mass 30.13 grams, had 3 marbles, an excess of 0.13 grams). Suppose I weigh additional subsets consisting of marbles $\{1, 4, 6, 7\}$, $\{1, 2, 8, 9, 10\}$, $\{2, 3, 4, 5, 8, 10\}$, $\{1, 2, 5, 6, 7\}$, and $\{3, 7, 10\}$, and obtain masses $40.00, 49.63, 59.76$, 50.00 and 30.00 grams, respectively.

   (a) Write out the other five relevant equations and formulate this as a system $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is 6×10.

   (b) What is coherence of $\mathbf{A}$? What sparsity solutions are guaranteed to be unique?

   (c) Find a minimal norm solution to $\mathbf{Ax} = \mathbf{b}$. What is its sparsity?

   (d) Run Matching Pursuit on the system; limit the number of iterations to 3, and use tolerance 1.0×10^{-5}. What is the result? Is the fit to the data perfect?

   (e) Run Orthogonal Matching Pursuit on the system; limit the number of iterations to 3, and use tolerance 1.0×10^{-5}. What is the result? Is the fit to the data perfect?

2. Load in the matrix in the file `Amat.mat`; it's a 20×100 matrix, standard normal random entries.

   (a) Compute the coherence of $\mathbf{A}$. What sparsity solutions to $\mathbf{Ax} = \mathbf{b}$ are assured to be unique?

   (b) Let $\mathbf{x}^*$ be the 3-sparse vector with nonzero components $x^*_{13} = 1.0, x^*_{33} = -0.82, x^*_{56} = 0.94$. Synthesize data $\mathbf{b} = \mathbf{Ax}^*$. Then run MP on $\mathbf{Ax} = \mathbf{b}$ using 10 iterations max to try to recover $\mathbf{x}^*$, with tolerance 0.1 (since $\|\mathbf{b}\|_2 \approx 6.83$, this should fit $\mathbf{b}$ pretty well.)

What are the nonzero components in your estimate? Make a plot of the solution vector.

What is the percentage error, as $\|\mathbf{x} - \mathbf{x}^*\|_2/\|\mathbf{x}^*\|_2$? How many iterations did MP actually run?

(c) Repeat the last problem with OMP.

(d) Repeat the last problem using the minimum norm solution.

(e) Suppose the data vector $\mathbf{b}$ is a bit noisy. Add noise to $\mathbf{b}$ as

$$\texttt{e} = \texttt{0.1} * \texttt{norm}(\texttt{b})/\texttt{sqrt}(\texttt{20}) * \texttt{randn}(\texttt{size}(\texttt{b}));$$
$$\texttt{b} = \texttt{b} + \texttt{e};$$

This adds noise to $\mathbf{b}$ at about a 10 percent level, in that $\|\mathbf{e}\|_2 \approx 0.1\|\mathbf{b}\|_2$, on average. Repeat the recovery of $\mathbf{x}$ using MP and OMP, same parameters and tolerances, and plot the results and COMMENT.

(f) Suppose the solution isn't really 3-sparse, but has small additional components; in fact, maybe ALL of the components are nonzero. Take the original $\mathbf{x}^*$ that had 3 big components and add

$$\texttt{x} = \texttt{x} + \texttt{0.01} * \texttt{randn}(\texttt{size}(\texttt{x}));$$

so all components of $\mathbf{x}^*$ will be in the range of ± 0.02, except for the three large ones. Then generate data $\mathbf{b} = \mathbf{A}\mathbf{x}^*$, don't add noise to $\mathbf{b}$, though, and recover $\mathbf{x}^*$ with MP and OMP, same tolerances as above. Plot the results and COMMENT.

3. Suppose $\mathbf{A}$ is a $2 \times n$ matrix with the restriction that all entries in $\mathbf{A}$ must have nonnegative components. What is the smallest possible coherence that $\mathbf{A}$ can have (as a function of n)? Hint: mimic the argument the argument from the last homework, except now the column vectors $\mathbf{a}_i$ of $\mathbf{A}$ must lie in the first quadrant.

   Then make up 1000 matrices of dimension 2×4 (nonnegative entries), compute the coherence of each, and find the minimum and median value for $\mu(\mathbf{A})$ for your 1000 matrices.

4. Make up a random 100×1000 matrix in Matlab (use `A = randn(100,1000);`). Compute the coherence of $\mathbf{A}$. What sparsity of solutions are guaranteed to be unique and recoverable with OMP?

Make up a "true" solution $\mathbf{x}^* \in \mathbb{R}^{1000}$ with components $x_i^* = i$ for $1 \leq i \leq 10$, all other $x_i^* = 0$. Synthesize data $\mathbf{b} = \mathbf{A}\mathbf{x}^*$. Then try to recover $\mathbf{x}^*$ from $\mathbf{A}$ and $\mathbf{b}$ using both MP and OMP. In each case use at most 50 iterations, tolerance 1.0×10^{-6}.

Comment!