Homework 7

MA430

As always, give full justification for your answers!

1. Read Chapter 1, Chapter 6, and anything else I handed out (e.g., section I to III of the paper I wrote with Deborah Walter).

2. Consider a one-dimensional version of the problem of locating radio frequency (RF) sources from the paper I wrote with Dr. Walter. Specifically, suppose there are emitters scattered in the interval $[0, 1]$ on the real line. Possible locations for the emitters are $x_j = (j - 0.5)/N$ for $1 \leq j \leq N$ (equispaced points in the interval). We have RF sensors at altitude h (in the y direction) and with x coordinates $x = a_i$, $1 \leq i \leq M$ (so there are M sensors). The distance r_{ij} from the ith sensor to the jth potential emitter location is then

$$r_{ij} = \sqrt{(a_i - x_j)^2 + h^2}.$$

Assume the emitter at position x_j has power $p_j \geq 0$. Then the total received signal strength by sensor i is

$$d_i = \sum_{j=1}^{N} \frac{p_j}{r_{ij}^2}$$

for $1 \leq i \leq M$, if the signal attenuates as $1/r^2$ and normalizing constants are set to 1. We can write this as

$$\mathbf{Ap} = \mathbf{d}$$

where $\mathbf{p} = [p_1, \ldots, p_N]^T$, $\mathbf{d} = [d_1, \ldots, d_M]^T$, and $\mathbf{A}$ has entries

$$A_{ij} = \frac{1}{r_{ij}^2} = \frac{1}{(a_i - x_j)^2 + h^2}.$$

By the way, in this problem you can play around with the various tolerances and iteration caps—maybe you'll find a better combination than I suggest below.

(a) Let $N = 50$, and consider the case $M = 10$ with sensors at $x = a_i$ where the a_i are given by $a_i = (i - 0.5)/10$, $1 \leq i \leq 10$ (equispaced) and the sensors are at altitude $h = 0.2$. In Matlab, set up the sensing matrix. Then let $\mathbf{p}^*$ be the "true" (sparse) power vector with nonzero entries $p^*_{17} = 1.2$ and $p^*_{36} = 0.5$. Compute the data $\mathbf{d} = \mathbf{Ap}^*$, then try to recover $\mathbf{p}^*$ using both OMP and BP (my OMP routine, and the BP routine in the L^1 magic package). Cap the iterations for OMP at 10 with a tolerance of 10^{-4}. You can let BP have 100 iterations, same tolerance. Plot the results (try a stem plot).

What is the coherence of the sensing matrix?

(b) Repeat part (a), but now add a bit of noise to the data vector $\mathbf{d}$, as

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

This adds noise to $\mathbf{d}$ at about a 10 percent level, in that $\|\mathbf{e}\|_2 \approx 0.1\|\mathbf{d}\|_2$, on average.

(c) Repeat part (b) but now with a "noise aware" version of basis pursuit, or with OMP but using a much looser tolerance. Specifically, in OMP use tolerance 10.0 (this is about 1.5 times the noise level).

For basis pursuit recovery use the L^1 magic BPDN routine

$$\texttt{xp} = \texttt{l1qc\_logbarrier(x0, A, [], d, epsilon, 1e} - \texttt{3)};$$

where "epsilon" is $0.15\|\mathbf{d}\|_2$ the noise level, in the Euclidean norm, comparable to what we used for OMP). Here x0 is an initial guess, say `x0 = minnorm(A,d);` (the minimum norm solution).

3. Suppose I have data that falls into 3 classes. I collect 2 feature vectors for each class, amalgamate them into a dictionary $\mathbf{D}$ given by

$$\mathbf{D} = \left[\begin{array}{rrrrrr} 1.1 & 3.02 & 0.07 & 0.03 & -1.01 & -2.00 \\ 1.05 & 2.96 & 2.02 & -3.01 & 1.03 & 1.98 \end{array} \right].$$

Columns 1 and 2 are class 1, 3 and 4 are class 2, columns 5 and 6 are class 3. I then collect feature vector $\mathbf{x} = [-2.1, 2.4]^T$. Find a sparse solution $\boldsymbol{\alpha}$ to $\mathbf{D}\boldsymbol{\alpha} = \mathbf{x}$ using OMP, then classify $\mathbf{x}$ using the method of the text.