Assignment 4: CS 754, Advanced Image Processing

Due: 4th April before 11:55 pm

Remember the honor code while submitting this (and every other) assignment. All members of the group should work on and <u>understand</u> all parts of the assignment. We will adopt a zero-tolerance policy against any violation.

Submission instructions: You should ideally type out all the answers in Word (with the equation editor) or using Latex. In either case, prepare a pdf file. Create a single zip or rar file containing the report, code and sample outputs and name it as follows: A4-IdNumberOfFirstStudent-IdNumberOfSecondStudent.zip. (If you are doing the assignment alone, the name of the zip file is A4-IdNumber.zip). Upload the file on moodle BEFORE 11:55 pm on the due date. The cutoff is 10 am on 5th April after which no assignments will be accepted. Note that only one student per group should upload their work on moodle. Please preserve a copy of all your work until the end of the semester. <u>If you have difficulties, please do not hesitate to seek help from me.</u>

1. Consider a signal x which is sparse in the canonical basis and contains n elements, which is compressively sensed in the form $y = \Phi x + \eta$ where y, the measurement vector, has m elements and Φ is the $m \times n$ sensing matrix. Here η is a vector of noise values that are distributed by $\mathcal{N}(0, \sigma^2)$. One way to recover x from y, Φ is to solve the LASSO problem, based on minimizing $J(x) \triangleq \|y - \Phi x\|^2 + \lambda \|x\|_1$. A crucial issue is to how to choose λ. One purely data-driven technique is called cross-validation. In this technique, out of the m measurements, a random subset of (say) 90 percent of the measurements is called the reconstruction set R, and the remaining measurements constitute the validation set V. Thus V and R are always disjoint sets. The signal x is reconstructed using measurements only from R (and thus only the corresponding rows of Φ) using one out of many different values of λ chosen from a set Λ. Let the estimate using the gth value from Λ be denoted x_g. The corresponding validation error is computed using $VE(g) \triangleq \sum_{i \in V} (y_i - \Phi^i x_g)^2 / |V|$. The value of λ for which the validation error is the least is chosen to be the optimal value of λ. Your job is to implement this technique for the case when $n = 500, m = 200, \|x\|_0 = 18, \sigma = 0.05 \times \sum_{i=1}^{m} |y_i|/m$. Choose $\Lambda = \{5 \times 10^{-4}, 10^{-4}, 5 \times 10^{-3}, 10^{-3}, 0.01, 0.05, 0.1, 0.5, 1, 2, 5\}$. Draw the non-zero elements of x at randomly chosen location, and let their values be drawn randomly from Uniform$(0, 1000)$. The sensing matrix Φ should be drawn from $\pm$Bernoulli with probability of $+1$ being 0.5. Now do as follows. Use the L1-LS solver from `https://web.stanford.edu/~boyd/l1_ls/` for implementing the LASSO.

   (a) Plot a graph of VE versus the logarithm of the values in Λ. Also plot a graph of the RMSE versus the logarithm of the values in Λ, where RMSE is given by $\|x_g - x\|_2 / \|x\|_2$. Comment on the plots. Do the optimal values of λ from the two plots agree?

   (b) What would happen if V and R were not disjoint but coincident sets?

   (c) The validation error is actually a proxy for actual mean squared error. Note that you can never determine the mean squared error since the ground truth x is unknown in an actual application. Which theorem/lemma from the paper `https://ieeexplore.ieee.org/document/6854225` (On the theoretical analysis of cross-validation in compressed sensing) refers to this proxying ability? Explain how.

   (d) In your previous assignment, there was a theorem from the book by Tibshirani and others which gave you a certain value of λ. What is the advantage of this cross-validation method compared to the choice of λ using that theorem? Explain. [10+5+5+5=25 points]

2. Consider that you learned a dictionary D to sparsely represent a certain class S of images - say handwritten alphabet or digit images. How will you convert D to another dictionary which will sparsely represent the following classes of images? Note that you are not allowed to learn the dictionary all over again, as it is time-consuming.

   (a) Class S_1 which consists of images obtained by applying a known derivative filter to the images in S.

   (b) Class S_2 which consists of images obtained by rotating a subset of the images in class S by a known fixed angle α, and the other subset by another known fixed angle β.

   (c) Class S_3 which consists of images obtained by applying an intensity transformation $I_{new}^i(x, y) = \alpha(I_{old}^i(x, y))^2 + \beta(I_{old}^i(x, y)) + \gamma$ to the images in S, where α, β, γ are known.

   (d) Class S_4 which consists of images obtained by applying a known blur kernel to the images in S.

   (e) Class S_5 which consists of images obtained by applying a blur kernel which is known to be a linear combination of blur kernels belonging to a known set B, to the images in S.

   (f) Class S_6 which consists of 1D signals obtained by applying a Radon transform in a known angle θ to the images in S.

   (g) Class S_7 which consists of images obtained by translating a subset of the images in class S by a known fixed offset (x_1, y_1), and the other subset by another known fixed offset (x_2, y_2). Assume appropriate zero-padding and increase in the size of the image canvas owing to the translation. [4+4+4+4+4+6+4=30 points]

3. How will you solve for the minimum of the following objective functions: (1) $J(A_r) = \|A - A_r\|_F^2$, where A is a known $m \times n$ matrix of rank greater than r, and A_r is a rank-r matrix, where $r < m, r < n$. (2) $J(R) = \|A - RB\|_F^2$, where $A \in \mathbb{R}^{n \times m}, B \in \mathbb{R}^{n \times m}, R \in \mathbb{R}^{n \times n}, m > n$ and R is constrained to be orthonormal. Note that A and B are both known.
   In both cases, explain briefly any one situation in image processing where the solution to such an optimization problem is required. [5+5+5+5=20 points]

4. We have studied the non-negative matrix factorization (NMF) technique in our course and examined applications in face recognition. I also described the application to hyperspectral unmixing. Your job is to find a research paper which explores an application of NMF in any task apart from these. You may look up the wikipedia article on this topic. Other interesting applications include stain normalization in pathology. Your job is to answer the following: (1) Mention the title, author list, venue and year of publication of the paper and include a link to it. (2) Which task does the paper apply NMF to? (3) How exactly does the paper solve the problem using NMF? What is the significance of the dictionary and the dictionary coefficients in solving the problem at hand? [15 points]

5. In parallel bean computed tomography, the projection measurements are represented as a single vector $y \sim \text{Poisson}(I_o \exp(-Rf))$, where $y \in \mathbb{R}^m$ with $m = $ number of projection angles $\times$ number of bins per angle; I_o is the power of the incident X-Ray beam; R represents the Radon operator (effectively a $m \times n$ matrix) that computes the projections at the pre-specified known projection angles; and f represents the unknown signal (actually tissue density values) in $\mathbb{R}^n$. If $m < n$, write down a suitable objective function whose minimum would be a good estimate of f given y and R and which accounts for the Poisson noise in y. State the motivation for each term in the objective function. Recall that if $z \sim \text{Poisson}(\lambda)$, then $P(z = k) = \lambda^k e^{-\lambda}/k!$ where k is a non-negative integer. Now suppose that apart from Poisson noise, there was also iid additive Gaussian noise with mean 0 and known standard deviation σ, in y. How would you solve this problem (eg: appropriate preprocessing or suitable change of objective function)? [6+ 4 = 10 points]