Assignment 5 - Advanced Image Processing

Kartik Gokhale, Pulkit Agarwal

April 2022

Contents

Note - To obtain the results reported in this document, run the MATLAB files specified alongside the section headers above.

1 Problem 1

1.1 Part A

For an image I and a linear operator $f_{i,k}$, the convolution $f_{i,k} \cdot I$ can be rewritten as $F_{i,k}v$, where v is the vectorized form of I, $\cdot$ represents convolution, and the RHS is simple matrix product.

- First Term:
$$\rho(f_{i,k} \cdot I_1) = \rho(F_{i,k}\text{vector}(I_1))$$
This gives $A_{j\rightarrow} = F_{i,k}$ and $b_j = 0$.
- Second Term:
$$\rho(f_{i,k} \cdot (I - I_1)) = \rho(-f_{i,k} \cdot I_1 - (-f_{i,k} \cdot I)) = \rho(-F_{i,k}\text{vector}(I_1) - (-F_{i,k}\text{vector}(I)))$$
This gives $A_{j\rightarrow} = -F_{i,k}$ and $b_j = -F_{i,k}\text{vector}(I)$.
- Third Term:
$$\lambda\rho(f_{i,k} \cdot I_1 - f_{i,k} \cdot I) = \lambda\rho(F_{i,k}\text{vector}(I_1) - F_{i,k}\text{vector}(I))$$
This gives $A_{j\rightarrow} = F_{i,k}$ and $b_j = F_{i,k}\text{vector}(I)$.
- Fourth Term:
$$\lambda\rho(f_{i,k} \cdot I_1) = \lambda\rho(F_{i,k}\text{vector}(I_1))$$
This gives $A_{j\rightarrow} = F_{i,k}$ and $b_j = 0$.

1.2 Part B

Equation (6) comprises of the following two expressions:

$$\sum_{i,k} \rho(f_{i,k} \cdot I_1) + \rho(f_{i,k} \cdot (I - I_1))$$

and

$$\lambda \sum_{i \in S_1,k} \rho(f_{i,k} \cdot I_1 - f_{i,k} \cdot I) + \lambda \sum_{i \in S_2,k} \rho(f_{i,k} \cdot I_1)$$

The first term represents the prior, while the second term is obtained from the likelihood.

The paper uses a prior on images that is based on the sparsity of derivative filters. They have defined a distribution over the images by assuming that derivative filters are independent over space and orientation. This gives the prior over images as

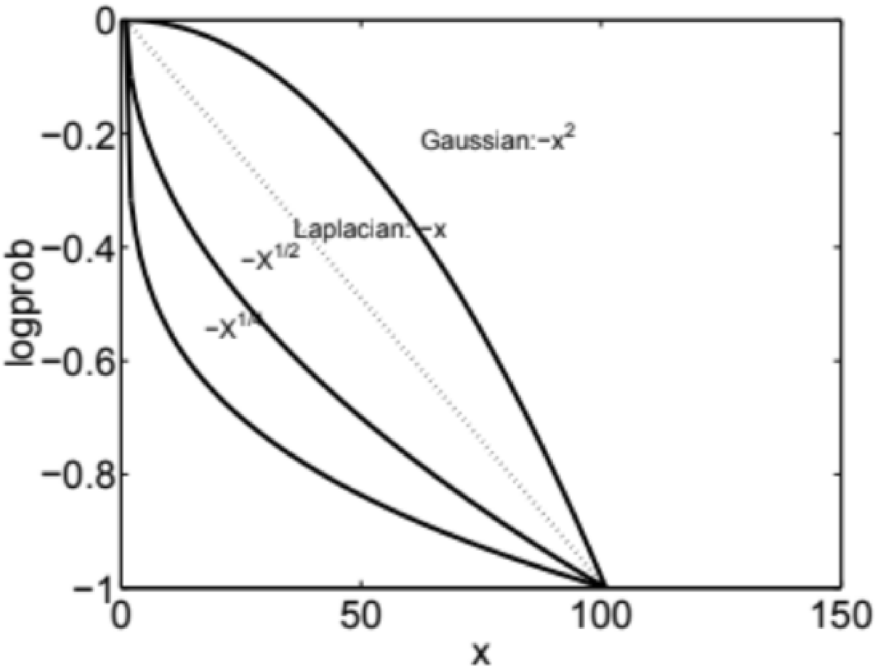
$$P(I) \approx \prod_{i,k} P(f_{i,k} \cdot I)$$

where $f \cdot I$ denotes the inner product between a linear filter f and an image I, and $f_{i,k}$ is the k^{th} derivative filter centered on pixel i.

The likelihood in the paper is given by the Laplacian mixture model, with distribution

$$P(x) = \frac{\pi_1}{2s_1} e^{\frac{-|x|}{s_1}} + \frac{\pi_2}{2s_2} e^{\frac{-|x|}{s_2}}$$

1.3 Part C



The paper exploits the statistics of natural images. It is known that the log-histograms of the derivative filters have a sparse distribution. They lie below the straight line connecting the minimal and maximal values. The Gaussian distribution is not sparse because it is always above the straight line. The Laplacian distribution is exactly at the border between sparse and non-sparse distributions (shown in the image above). This prior has been used because the sparse nature of derivative filter outputs is a robust property of natural images and this paper focuses on removing reflections from natural images.

2 Problem 2 - Q2.m

2.1 MAP Estimate of Signal

We have

$$y = \Phi x + \eta \text{ , where } \Phi \sim \mathcal{N}(0, \frac{1}{m}) \text{ and } \eta \sim \mathcal{N}(0, \sigma^2)$$

Using Bayes' Rule, the maximum a-posteriori estimate is given by

$$x_{\text{MAP}} = \arg\max_x P(y|x)P(x)$$

Here $P(x)$ is the Gaussian distribution with covariance matrix Σ_x and mean 0. On the other hand, $P(y|x)$ is given by the distribution of η, with mean shifted by Φx. Since $\eta \sim \mathcal{N}(0, \sigma^2)$, so $P(y|x) = \mathcal{N}(\Phi x, \sigma^2)$. This gives us

$$x_{\text{MAP}} = \arg\max_x e^{-\frac{||y-\Phi x||^2}{2\sigma^2}} \times \frac{e^{-\frac{1}{2}x^T \Sigma_x^{-1} x}}{(2\pi)^{\frac{n}{2}} |\Sigma_x|^{\frac{1}{2}}}$$

As the covariance matrix does not depend on x, so we do not need to consider it's determinant above. Taking negative logarithm of the above equation then gives us

$$x_{\text{MAP}} = \arg\min_x \left(\frac{(y - \Phi x)^T (y - \Phi x)}{2\sigma^2} + \frac{1}{2} x^T \Sigma_x^{-1} x \right)$$

We differentiate the above expression with respect to x, by using the following two results:

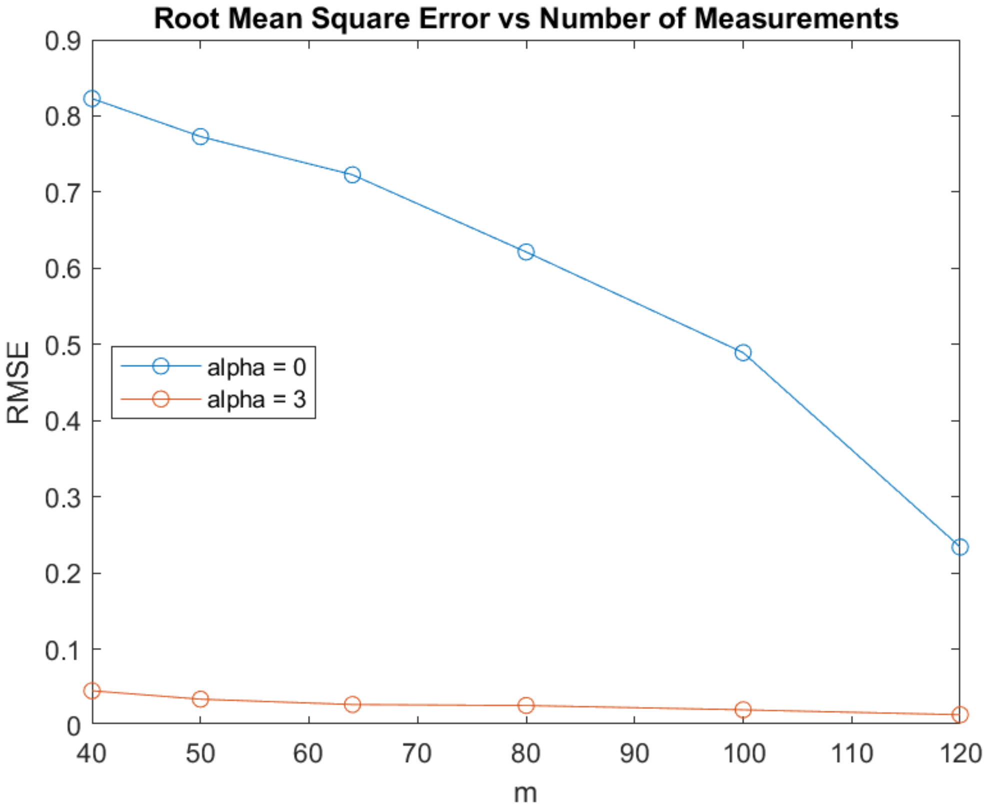
$$\frac{\delta}{\delta x} x^T A x = 2Ax, \frac{\delta}{\delta x}(y - Ax)(y - Ax)^T = 2A^T A x - 2A^T y$$

Equating the derivative to 0 (to find the minimum of the expression) leads to

$$\frac{1}{2\sigma^2}(2\Phi^T \Phi x_{\text{MAP}} - 2\Phi^T y) + \frac{1}{2}(2\Sigma_x^{-1} x_{\text{MAP}}) = 0$$

$$\implies \left(\frac{1}{\sigma^2} \Phi^T \Phi + \Sigma_x^{-1} \right) x_{\text{MAP}} = \frac{1}{\sigma^2} \Phi^T y \implies \boxed{x_{\text{MAP}} = (\Phi^T \Phi + \sigma^2 \Sigma_x^{-1})^{-1} \Phi^T y}$$

2.2 Graph and Results



The above graph shows a plot between the base 10 logarithm of the average root mean square error, and the different number of measurements (m) taken. As we can see, when α is increased from 0 to 3, RMSE decreases if m is kept same, or equivalently, the reconstruction performance improves with higher α.

Note that, for larger α, the eigenvalues of the covariance matrix decrease at a faster rate. In particular, smaller eigenvalue means smaller variation about the mean of x, and since the mean is 0, so it means smaller value at that index of x. Thus, as α increases, the number of near-zero values in x also increase, making the original signal sparser. This means that we will need lesser number of measurements to get the same error factor in the reconstructed signal for larger α. In other words, for same m, we will have a smaller error term when α is greater, as shown in the graph.

3 Problem 3

3.1 Part A

This follows from the definition of X^*.

$$X^* = \min_X ||X||_*, A(X) = b$$

Thus, for any X_0 satisfying A(X) = b, we have $||X_0||_* \geq ||X^*||_*$

3.2 Part B

Lemma 2.3 states that if A and B be matrices of the same dimension and If AB' = 0 and A'B = 0 then

$$||A + B||_* = ||A||_* + ||B||_*$$

We know from triangle inequality applied to part A that

$$||X_0||_* \geq ||X_0 + R_c||_* - ||R_0||_*$$

Since, we have chosen matrices R_0, R_c which exist according to Lemma 3.4, we have $X_0 R_c' = X_0' R_c = 0$. Thus, the conditions for invoking Lemma 2.3 are satisfied and we obtain

$$||X_0 + R_c||_* - ||R_0||_* \geq ||X_0||_* + ||R_c||_* - ||R_0||_*$$

3.3 Part C

By definition of SVD, we have each singular value in I_i greater than or equal to any singular value in I_{i+1} since the singular values are non-increasing. Consider

$$\frac{1}{3r} \sum_{j \in I_i} \sigma_j$$

Let σ_m be the smallest singular value in the above sum. Thus, we have

$$\begin{aligned} &\geq \frac{1}{3r} \sum_{j \in I_i} \sigma_m \\ &= \frac{1}{3r} 3r \cdot \sigma_m = \sigma_m \end{aligned}$$

Since any singular value in I_i greater than or equal to any singular value in I_{i+1}

$$\geq \sigma_k$$

3.4 Part D

Note that

$$\|R_i\|_* = \text{trace}\sqrt{R_i^* R_i} = \sum_{j \in I_i} \sigma_j$$

Also, we know

$$\frac{1}{3r}\sum_{j \in I_i} \sigma_j \geq \sigma_k$$

which becomes

$$\sigma_k \leq \frac{1}{3r}\|R_i\|_*$$

Squaring on both sides (as both the sides are positive quantities, the sign of the inequality doesn't change)

$$\sigma_k^2 \leq \frac{1}{9r^2}\|R_i\|_*^2 \quad \forall k \in I_{i+1}$$

Adding up the above $3r$ inequalities, each corresponding the one singular value of I_{i+1}, we get

$$\sum_{k \in I_{i+1}} \sigma_k^2 \leq \frac{3r}{9r^2}\|R_i\|_*^2$$

But $\|R_{i+1}\|_F^2 = \sum_{k \in I_{i+1}} \sigma_k^2$. Hence we have

$$\|R_{i+1}\|_F^2 \leq \frac{1}{3r}\|R_i\|_*^2$$

3.5 Part E

From Part D, we know that

$$\|R_{i+1}\|_F^2 \leq \frac{1}{3r}\|R_i\|_*^2$$

Taking square root,

$$\|R_{i+1}\|_F \leq \frac{1}{\sqrt{3r}}\|R_i\|_*$$

Repeating the above inequality and adding all of them, we have

$$\sum_{i \geq 1}\|R_{i+1}\|_F \leq \frac{1}{\sqrt{3r}}\sum_{i \geq 1}\|R_i\|_*$$

Let us express $j = i + 1$. Thus $i \geq 1 \implies j \geq 2$

$$\sum_{j \geq 2}\|R_j\|_F \leq \frac{1}{\sqrt{3r}}\sum_{i \geq 1}\|R_i\|_*$$

3.6 Part F

As all the R_j have orthogonal row space and column space, by the sub-addivity property of nuclear norm, we have

$$\sum_{j \geq 1}\|R_j\|_* = \|R_c\|_*$$

From Part B, we know that

$$\|X_0\| \geq ||X_0||_* + ||R_c||_* - ||R_0||_*$$

and thus

$$\|R_0\|_* \geq \|R_c\|_*$$

Hence

$$\sum_{j \geq 2}\|R_j\|_F \leq \frac{1}{\sqrt{3r}}\sum_{j \geq 1}\|R_j\|_* = \frac{1}{\sqrt{3r}}\|R_c\|_* \leq \frac{1}{\sqrt{3r}}\|R_0\|_*$$

3.7 Part G

We know the nuclear norm and Frobenius norm (from linear algebra) for some matrix P are related as

$$\|P\|_* \leq \sqrt{\text{rank}(P)}\ \|P\|_F$$

Since rank of X_0 is r, by design using Lemma 3.4, we have

$$\text{rank}(R_0) \leq 2r$$

Applying to R_0, we get

$$\|R_0\|_* \leq \sqrt{2r}\ \|R_0\|_F$$

We also know, from Part F, that

$$\sum_{j\geq 2} \|R_j\|_F \leq \frac{1}{\sqrt{3r}} \sum_{j\geq 1} \|R_j\|_* = \frac{1}{\sqrt{3r}} \|R_c\|_* \leq \frac{1}{\sqrt{3r}} \|R_0\|_*$$

Thus, comibining with the inequation above

$$\sum_{j\geq 2} \|R_j\|_F \leq \frac{1}{\sqrt{3r}} \sum_{j\geq 1} \|R_j\|_* = \frac{1}{\sqrt{3r}} \|R_c\|_* \leq \frac{1}{\sqrt{3r}} \|R_0\|_* \leq \frac{\sqrt{2r}}{\sqrt{3r}} \|R_0\|_F$$

3.8 Part H

We have constructed R_0 using Lemma 3.4 which states existence of R_1, R_0 such that

$$\text{rank}(R_0) \leq 2\text{rank}(X_0) = 2r$$

We know $\text{rank}(R_1) \leq 3r$ since we partitioned R_c as a sum of matrices each of rank atmost 3. Hence by sub-additivity property of rank, we have

$$\text{rank}(R_0 + R_1) \leq 2r + 3r = 5r$$

3.9 Part I

We have, by definition of $R_j s$

$$R = R_0 + R_c = R_0 + R_1 + \sum_{j\geq 2} R_j$$

Hence, by triangle inequality,

$$\left\|\mathcal{A}\left(\sum_{j\geq 2} R_j\right)\right\| \leq \sum_{j\geq 2} \|\mathcal{A}(R_j)\|$$

But again using the triangle inequality, we have

$$\|\mathcal{A}(R)\| = \left\|\mathcal{A}\left((R_0 + R_1) + \sum_{j\geq 2} R_j\right)\right\| \geq \|\mathcal{A}(R_0 + R_1)\| - \left\|\mathcal{A}\left(\sum_{j\geq 2} R_j\right)\right\|$$

Hence,

$$\|\mathcal{A}(R)\| =\geq \|\mathcal{A}(R_0 + R_1)\| - \left\|\mathcal{A}\left(\sum_{j\geq 2} R_j\right)\right\| \geq \|\mathcal{A}(R_0 + R_1)\| - \sum_{j\geq 2} \|\mathcal{A}(R_j)\|$$

Thus,

$$\|\mathcal{A}(R)\| \geq \|\mathcal{A}(R_0 + R_1)\| - \sum_{j\geq 2} \|\mathcal{A}(R_j)\|$$

3.10 Part J

Now, we have the restricted isometry property (RIP) which states that

$$(1-\delta_s)\cdot||\theta||_2^2 \leq ||A\theta||_2^2 \leq (1+\delta_s)\cdot||\theta||_2^2$$

As the rank of $R_0 + R_1$ is atmost $5r$, by using the restricted isometry property of $\mathcal{A}$, we have

$$\|\mathcal{A}(R_0+R_1)\| \geq (1-\delta_{5r})\,\|R_0+R_1\|_F$$

Also, we know that the rank of each R_j is atmost 3r $\forall j \geq 2$. Similarly,

$$\sum_{j\geq 2}\|\mathcal{A}(R_j)\| \leq (1+\delta_{3r})\sum_{j\geq 2}\|R_j\|_F \quad \forall j \geq 2$$

Using the above two inequalities along with

$$\|\mathcal{A}(R)\| \geq \|\mathcal{A}(R_0+R_1)\| - \sum_{j\geq 2}\|\mathcal{A}(R_j)\|$$

we get

$$\begin{aligned}\|\mathcal{A}(R)\| &\geq \|\mathcal{A}(R_0+R_1)\| - \sum_{j\geq 2}\|\mathcal{A}(R_j)\| \\ &\geq (1-\delta_{5r})\,\|R_0+R_1\|_F - (1+\delta_{3r})\sum_{j\geq 2}\|R_j\|_F\end{aligned}$$

3.11 Part K

Since $\mathcal{A}$ is an affine transformation and

$$\mathcal{A}(X^*) = b, \mathcal{A}(X_0) = b$$

we have

$$\mathcal{A}(R) = \mathcal{A}(X^* - X_0) = \mathcal{A}(X^*) - \mathcal{A}(X_0) = b - b = 0$$

This is the assumption being referred to here

3.12 Part L

We have

$$\|\mathcal{A}(R)\| \geq \left((1-\delta_{5r}) - \frac{9}{11}(1+\delta_{3r})\right)\|R_0\|_F$$

Since $\mathcal{A}(R) = 0$, $\|\mathcal{A}(R)\| = 0$. Frobenius norm involves adding squared quantities, and thus is always non-negative. Hence,

$$\begin{aligned}(1-\delta_{5r}) - \frac{9}{11}(1+\delta_{3r}) &> 0 \\ 2 - (11\delta_{5r} + 9\delta_{3r}) &> 0 \\ 2 &> 11\delta_{5r} + 9\delta_{3r}\end{aligned}$$

4 Problem 4

4.1 Part A

Incoherence of singular vectors with the canonical basis is required for matrix recovery to avoid cases where the low rank matrix is in the null space of the sampling operator. When random elements of such matrices are sampled most of the time we would observe small variation among the sampled values and we have no way of knowing what were the significantly different values. For example, consider a matrix with all entries zero, except the top left entry, which is set to 1. This matrix cannot be recovered from a sampling of its entries unless we pretty much see all the entries, irrespective of the method used.

To minimize the number of measurements needed for recovery, both the left and right singular vectors need to be uncorrelated with the standard basis, since matrices whose column and row spaces have low coherence cannot really be in the null space of the sampling operator. In particular, if sampling matrix M has SVD $U\Sigma V^T$, then there must exist some constant which upper bounds the coherence of U and V.

4.2 Part B

Suppose the constraint is changed to

$$\begin{aligned} \text{minimize} \quad & rank(X) \\ \text{subject to} \quad & f_i^* X g_j = f_i^* M g_j \quad \forall\ (i,j) \in \Omega \end{aligned}$$

Note that there exist orthonormal (unitary for complex matrices) transformations F and G such that for all $i = 1, 2, \ldots n$, we get $e_i = Ff_i$ and $e_i = Gg_i$. Then, we can write

$$f_i^* X g_j = e_i^*(FXG^*)e_j$$

Again we can see that if the left and right singular vectors are highly coherent with f_i and g_j then we will get an almost zero matrix with few non-zero elements. Hence, we need the conditions to hold for this matrix FXG^* now. Thus, all that is needed is that the column and row spaces of M be respectively incoherent with the basis f_i and g_i.

4.3 Part C

The authors have given the example of matrix M as

$$M = \sum_{k=1}^{2} \sigma_k u_k u_k^* \ , \ u_1 = \frac{e_1 + e_2}{\sqrt{2}} \ , \ u_2 = \frac{e_1 - e_2}{\sqrt{2}}$$

for arbitrary singular values. This matrix vanishes everywhere except in the top-left 2×2 corner. So its randomly samples elements might be 0 with high probability, and this makes it hard to recover the matrix. The techniques in the paper require that the singular vectors be significantly spread, or in other words, uncorrelated with the standard basis. Since matrix M does not satisfy this property, hence the method in the paper is unsuccessful for this example.

5 Problem 5

5.1 Details of the Paper

Title: **A unified approach to salient object detection via low rank matrix recovery**
Authors: Xiaohui Shen and Ying Wu
Venue: 2012 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2012
Date: 1 October, 2012
The paper can be accessed here.

5.2 Problem Statement

Salient object detection (SOD) is an important computer vision task aimed at precise detection and segmentation of visually distinctive image regions from the perspective of the human visual system (HVS). The behavior of SOD models is expected to mimic the pre-attentive stage of HVS which guides human attention to the highly interesting regions in the scene. The identified salient regions in images can facilitate subsequent high-level vision tasks for improved efficiency and optimal resource usage. As a preprocessing step, SOD has served many computer vision tasks such as, visual tracking, image captioning, image/video segmentation, and so forth.

5.3 Solving the Problem

The paper describes a method which first performs feature extraction. After feature extraction, they perform image segmentation based on the extracted features by mean-shift clustering. Then, they consider the image as a combination of a background residing in a low dimensional space with salient objects as sparse noises. Therefore, F can be decomposed into two parts $F = L + S$, where L is the low-rank matrix corresponding to the background while S is a sparse matrix representing the salient regions. The low-rank matrix recovery problem can then be formulated as:

$$(L^*, S^*) = \min_{L,S}(rank(L) + \lambda||S||_0)$$

Since the above problem is NP-hard and hard to approximate, one can alternatively solve the convex surrogate:

$$(L^*, S^*) = \min_{L,S}(rank(L) + \lambda||S||_1)$$

which is the low rank matrix recovery problem. L and S can be perfectly recovered, in most cases by the Robust PCA method. Thus, a saliency map is then accordingly generated and normalized to be a gray-scale image with the salient details, solving the problem