Problem set 5: Image reconstruction and estimation theory

Theoretical image science
Mats Persson,

February 2023, corrected May 2023

Discussing solutions is encouraged but you must **individually** write and hand in your solutions. For numerical computations and plots, you can use the programming language of your choice, e.g. Matlab or Python.

These homework problems build on material covered in all lectures, in particular the two last ones.

1 Sparsity-based denoising (3p)

Variational reconstruction problems seldom have closed-form exact solutions but here we will study one that does. Let us once more return to the problem of a noisy $N \times N$ image of the night sky. Since only a minor fraction of the image can be assumed to contain visible stars, it makes sense to look for an image that is sparse in the pixel basis, i.e. as few pixels as possible are nonzero. Since this gives a computationally difficult problem, we make use of the observation that the following denoising problem will typically converge to a sparse image:

$$\hat{\mathbf{g}} = \mathrm{argmin}_{\mathbf{g}} \|\mathbf{g} - \mathbf{g}_{\mathrm{meas}}\|_2^2 + \lambda \|\mathbf{g}\|_1 \tag{1}$$

Here, $\mathbf{g}_{\mathrm{meas}}$ is the actual measured image, reformatted into a vector, and $\hat{\mathbf{g}}$ is the denoised image. Furthermore, λ denotes a regularization parameter and $\|\mathbf{x}\|_p = (\sum_i x_i^p)^{1/p}$ denotes the l^p norm.

Show that the solution to 1 is given by

$$\hat{g}_i = \begin{cases} g_i - \lambda/2, \ g_{\mathrm{meas}} \geq \lambda/2 \\ 0, -\lambda/2 \leq g_{\mathrm{meas}} < \lambda/2 \\ g_i + \lambda/2, \ g_{\mathrm{meas}} < -\lambda/2 \end{cases} \tag{2}$$

This is referred to as a "soft thresholding" of the original image. Can you explain (without formulas) why such an operation will lead to a sparse image?

2 Deconvolution (4p)

The file `HW5_deconvolution.mat` contains an image $\hat{\mathbf{g}}$ corrupted by blurring with a psf h(x,y) (also included) and noise. (If you are using python you can load this with `scipy.io.loadmat`.) Deconvolve this image by

$$\hat{\mathbf{g}} = \text{argmin}_{\mathbf{g}} \|\mathcal{H}\mathbf{g} - \mathbf{g}_{\text{meas}}\|_2^2 + \lambda \sum_i \sum_j w_{ij}\phi(g_i - g_j) \tag{3}$$

where $\mathcal{H}$ is an operator that convolves the image with h(x,y).

$$w_{ij} = \begin{cases} 1, \text{if pixels i and j are abutting neighbors} \\ 1/\sqrt{2}, \text{if pixels i and j are diagonal neighbors} \\ 0, \text{if i=j or if i and j are not neighbors} \end{cases} \tag{4}$$

and ϕ is the Huber penalty

$$\phi(x) = \begin{cases} x^2/2, |x| \leq x_0 \\ x_0|x| - x_0^2/2, |x| > x_0 \end{cases} \tag{5}$$

a) Write a function that takes an input image and calculates $\sum_i \sum_j w_{ij}\phi(g_i - g_j)$. Verify that this function gives the correct output value $(4(1 + 1/\sqrt{2}))$ when $x_0 = 2$ and $\mathbf{g}$ is a zero-filled image with one (non-boundary) pixel being equal to 1. (g_i are components of $\mathbf{g}$.)
b) Use an optimization routine of your choice (you are not expected to implement it yourself) to solve (3). In `Matlab` you can use `fminunc` (you may want to increase the maximum number of function evaluations to about 100000), and in `python`, you can use `scipy.optimize.fmin_bfgs`. Use $\lambda = 0.02$ and try both $x_0 = 0.02$ and $x = 1$. Explain the difference in the result between these two cases.
You can implement the convolution using for example conv2 in matlab or scipy.signal.convolve2d in python. Make sure that the size of the output image is the same as the size of the input image.
Note: Due to the small size of the image, this should converge in a few minutes. But to speed up the run time it can be useful to solve the problem for a downsized version of the image while debugging your program.

3 The Cramér-Rao lower bound for velocity estimation (3p)

Consider a video sequence of one-dimensional images of an object with center position $x(t)$ as function of time t. For simplicity we assume that $x(t)$ is known to be at the origin $x = 0$ at $t =$ and move with constant velocity v, i.e. $x(t) = vt$. The imaging task is to measure v. Each frame in the video sequence is acquired with a 1D array of pixels centered at $x_i = i\Delta x$ for $i = -I/2 \ldots I/2 - 1$, at time points $t_k = k\Delta t$ for $k = 0 \cdots K - 1$. The measured image value at pixel i at time k is $g_{ik} = \langle g_{ik} \rangle + n_{ik}$ where n_{ik} is a Gaussian random variable with mean 0 and standard deviation σ (equidistributed and independent for different i and k).

Let $h(x)$ denote the image of the object centered at $x = 0$.

Use the Cramér-Rao Lower Bound to show that any unbiased estimator of v must satisfy

$$\mathrm{V}[v] \geq \sigma^2 \frac{1}{\sum_{i=-I/2}^{I/2-1} \sum_{k=0}^{K-1} h'(x_i - vt_k)^2 t_k^2} \tag{6}$$