Inverse problems course, Exercise 5 (for the week starting on February 27, 2017)
University of Helsinki
Department of Mathematics and Statistics
Samuli Siltanen, Markus Juvonen, Minh Mach, Santeri Kaupinmäki and Alexander Meaney
Related book sections (Mueller & Siltanen 2012): Section 6.

Theoretical exercises:

T1. Compute the derivative $g'(t)$ when the function $g : \mathbb{R} \to \mathbb{R}$ is given by

$$\text{(a) } g(t) = \sqrt{t^2 + \beta}, \qquad \text{(b) } g(t) = \frac{1}{\beta} \log(\cosh(\beta t)),$$

where $\beta > 0$.

T2. Let $x \in \mathbb{R}^3$ and $\beta > 0$. Compute the gradient of the approximate total variation penalty functional

$$\Phi_{\mathrm{TV}}(x) = \sum_{\ell=1}^{2} |x_{\ell+1} - x_\ell|_\beta,$$

where $|t|_\beta = g(t)$ as in exercise T1(a).

T3. Read the Wikipedia article on steepest descent minimization and explain in your own words how it works.

Matlab exercises:

M1. (a) Plot the absolute value function $|x|$ for $-M \leq x \leq M$, starting with $M = 2$. Add to the same picture plots of function $g(x)$ as in T1(a) with various choices of parameter β. Choose M and values of β so that the similarities and differences between functions $|x|_\beta$ and $|x|$ become clear.

(b) Make a similar plot using g from T1(b).

M2. **Sparsity-based regularization parameter choice.** Consider Total Variation regularization, or in other words, finding the vector $\mathbf{f}_\alpha \in \mathbb{R}^n$ that gives the smallest possible value to the functional $\|A\mathbf{f} - \widetilde{\mathbf{m}}\|_2^2 + \alpha\|L\mathbf{f}\|_1$ with

$$L = \frac{1}{\Delta x}\begin{bmatrix} -1 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & -1 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \ddots & & \\ \vdots & & & & & \ddots & \\ 0 & & \cdots & & 0 & -1 & 1 & 0 \\ 0 & & \cdots & & 0 & 0 & -1 & 1 \end{bmatrix}. \tag{1}$$

Here $\alpha > 0$ is the so-called regularization parameter giving the trade-off between measurement data and *a priori* information about $\mathbf{f}$. Choosing the optimal value for α is one of the central challenges in the mathematics of inverse problems. Here we discuss one possible approach.

The effect of using the 1-norm $\|\cdot\|_1$ is that after minimization, the vector inside the 1-norm has most of its elements equal to zero. This is called *sparsity*, and it is related to *compressed sensing*, a hot topic in signal processing research.

Since our regularization term is $\|L\mathbf{f}\|_1$, we should consider sparsity in the derivative $f'(x)$. In your 1D convolution simulation, take the continuum model f to be a piecewise constant function with 8 jumps (that is, four flat areas different from zero background). Assume that we have *a priori* information about the sparsity of $f'(x)$. In this case it means that we know the number of jumps to be 8.

(a) Take $n = 64$, use quadratic programming to solve the inverse problem with a variety of α values, and choose the α that results in as close to 8 jumps as possible in the reconstruction. Repeat for a higher noise level.

(b) Repeat (a) with $n = 128$. Is the relative error in the reconstructions smaller or larger compared to $n = 64$?

M3. Go to this page and download the Matlab routines under the headline *Compute reconstruction using matrix-free iterative (approximate) total variation regularization.* Compute TV regularized one-dimensional deconvolutions by simplifying those files from the 2D tomography case to the 1D deconvolution case. You should use the smooth TV penalty of Problem T2.

How do the results compare to the reconstructions computed with the quadratic programming approach (`deconv9_TVreg_comp.m`)?