Inverse problems course, spring 2014 Exercise 4 (February 11-14, 2014)
University of Helsinki
Department of Mathematics and Statistics
Samuli Siltanen, Esa Niemi and Teemu Saksala

Note that this exercise has more than one page and contains *three* types of exercises: theoretical (T), computational (M) and LaTeX reporting (L).

Please complete the theoretical exercises (marked with T) before the exercise session and be prepared to present your solution there.

T1. Show that $A^TA + \delta I$ is always an invertible matrix for $\delta > 0$.
Hint: note that A^TA is symmetric and study the eigenvalues of $A^TA + \delta I$.

Answer:

Let $\delta > 0$ and $A : \mathbb{R}^n \to \mathbb{R}^k$. Since matrix A^TA is symmetric it holds due the Spectral theorem that there exists an orthonormal basis $(v_i)_{i=1}^n$ of eigen vectors of A^TA. Let λ_i be an eigen value of related to v_i. Calculate

$$Av_i \cdot Av_j = v_i \cdot A^TAv_j = \lambda_j(v_i \cdot v_j) = \lambda_j\delta_{ij} \tag{1}$$

Equation (1) proves that $\lambda_i \geq 0$ for any i. Calculate

$$(A^TA + \delta I)v_i = A^TAv_i\delta v_i = \underbrace{(\lambda_i + \delta)}_{>0} v_i. \tag{2}$$

The equation (2) proves that mapping $A^TA + \delta I$ has precisely n eigen values and they all are strictly positive. Since it holds for any linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$, that L does not have a zero eigen valuea iff it is one-to-one iff it is onto, we will deduce that $A^TA + \delta I$ is invertible.

T2. Write the following generalized Tikhonov regularized solution in stacked form:

$$\arg\min_{z\in\mathbb{R}^n}\{\|Az - m\|^2 + \delta\|L(z - z^*)\|^2\}.$$

Here the fixed vector $z^* \in \mathbb{R}^n$ is assumed to be known *a priori*.

Answer:

Suppose that $T_\delta(m) \in \mathbb{R}^n$ is the

$$\arg\min_{z\in\mathbb{R}^n}\{\|Az - m\|^2 + \delta\|L(z - z^*)\|^2\}.$$

Then for every $w \in \mathbb{R}^n$ it holds, that

$$\frac{d}{dt}\left\{\|A(T_\delta(m) + tw) - m\|^2 + \delta\|L(T_\delta(m) - z^* + tw)\|^2\right\}_{t=0} = 0.$$

If one does similar computations as in page 67 in textbook one notes that

$$(A^TA + \delta L^TL)T_\delta(m) = A^Tm + \delta L^TLf^*. \tag{3}$$

Consider equation

$$\widetilde{A}\mathbf{f} := \begin{bmatrix} A \\ \sqrt{\delta}L \end{bmatrix} \mathbf{f} = \begin{bmatrix} \mathbf{m} \\ \sqrt{\delta}\, Lf^* \end{bmatrix} =: \widetilde{\mathbf{m}}. \tag{4}$$

Clearly we have that

$$\widetilde{A}^T = \begin{bmatrix} A \\ \sqrt{\delta}L \end{bmatrix}^T = \begin{bmatrix} A^T & \sqrt{\delta}L^T \end{bmatrix}.$$

Now multiply equation (4) with $\widetilde{A}^T$ and replace $\mathbf{f}$ by $T_\delta(m)$ to get equation (3). This proves that (4) is our stacked form.

T3. Let's study the tomographic "ghosts" discussed by Smith, Solmon and Wagner in their classic 1977 paper. For simplicity, we concentrate on the case of one projection direction only.

(a) Choose the direction vector to be $\vec{\theta}_1 = [1, 0]^T$. Define a polynomial by $q(\xi) = \langle \xi, \vec{\theta}_1 \rangle$. Show that q vanishes on the line $\vec{\theta}_1^\perp$ perpendicular to $\vec{\theta}_1$.

(b) Write down the differential operator Q obtained by replacing ξ_j by $-i\partial/\partial x_j$ in the expression for q.

(c) Take any compactly supported, infinitely smooth function $g \in C_0^\infty(\mathbb{R}^2)$ and set $f := Qg$. Determine such a direction $\vec{\theta}$ that $\mathcal{R}f(s, \vec{\theta}) = 0$ for all $s \in \mathbb{R}$. This means that f is invisible in that particular X-ray projection image.

Answer:

(a) Define $\theta = \vec{\theta}_1$. We first note that polynomial $q : \mathbb{R}^2 \to \mathbb{R}$ is given by formula

$$q(\xi) = \langle \xi, \theta \rangle = \xi \cdot \theta = \xi_1.$$

Let $\alpha = [0, 1]^T$. It holds that line $\{t\alpha : t \in \mathbb{R}\}$ is perpendicular to line defined by θ. Therefore it holds

$$q(t\alpha) = t\alpha_1 = 0.$$

(b) Let $\theta \in S^1$. We define a differential operator Q_θ as

$$Q_\theta = \langle (-i\frac{\partial}{\partial x_j})_{j=1}^2, \theta \rangle = -i\left(\theta_1 \frac{\partial}{\partial x_1} + \theta_2 \frac{\partial}{\partial x_2}\right).$$

If $g \in C^k(\mathbb{R}^2, \mathbb{C}), k \geq 1$ then it holds that

$$Q_\theta : C^k(\mathbb{R}^2) \to C^{k-1}(\mathbb{R}^2),\ Q_\theta(g) = -i\left(\theta_1 \frac{\partial g}{\partial x_1} + \theta_2 \frac{\partial g}{\partial x_2}\right)$$

(c) Let $g \in C_0^\infty(\mathbb{R}^2)$. Then it holds that function $f := Q_\theta(g) \in C_0^\infty(\mathbb{R}^2)$. Radon transform is well defined for any function $h \in C_0^\infty(\mathbb{R}^2)$, since set $\operatorname{supp} h \cap L_{\alpha,s} \subset \mathbb{R}$ is compact where $L_{s,\alpha}$ is the line $x \cdot \alpha = s$.

Define $\gamma(t) := \frac{\theta_2}{\theta_1} t - \frac{s}{\theta_1}$ and $\alpha := [\theta_2, -\theta_1]$. We claim that the Radon transform of f vanishes at each line $L_{s,\alpha}$. Calculate

$$\mathcal{R}f(\alpha, s) = \int_{\mathbb{R}} f(t, \gamma(t))\, dt = \int_{\mathbb{R}} Q_\theta(g)(t, \gamma(t))\, dt$$

$$= -i \int_{\mathbb{R}} \theta_1 \frac{\partial g}{\partial x_1}(t, \gamma(t)) + \theta_2 \frac{\partial g}{\partial x_2}(t, \gamma(t))\, dt = -i\theta_1 \int_{\mathbb{R}} \frac{\partial g}{\partial x_1}(t, \gamma(t)) + \frac{\theta_2}{\theta_1} \frac{\partial g}{\partial x_2}(t, \gamma(t))\, dt$$

$$= -i\theta_1 \int_{\mathbb{R}} \frac{d}{dt}(g(t, \gamma(t)))\, dt = -i\theta_1 \lim_{m \longrightarrow \infty} \int_{-m}^{m} \frac{d}{dt}(g(t, \gamma(t)))\, dt$$

$$= -i\theta_1 \lim_{m \longrightarrow \infty} g(m, \gamma(m)) - g(-m, \gamma(-m)) = 0$$

The second last equality holds due fundamental theorem of analysis. The last holds since we assumed g to be compactly supported. Now we have proven that the Radon transform of f vanishes at line $x \cdot \theta = s$, which was the claim.

You can work on these Matlab exercises (marked with M) in the exercise session.

M1. Apply the L-curve method to the generalized Tikhonov regularization with discrete derivative penalty. In other words, insert the stuff from the file `DC6_TikhonovD_comp.m` into the file `DC5_Tikhonov_Lcurve.m` in a suitable way. Plot the reconstruction you get from the automatic choice of regularization parameter.

Remember to use $\log \|Lx\|$ in the L-curve plot instead of $\log \|x\|$.

Save the plot in Encapsulated PostScript format with filename convplot.eps. The Matlab command you need is

```
>> print -depsc convplot.eps
```

Answer

The computations in file `DC5_Tikhonov_Lcurve.m` work with the standard Tikhonov regularization with the matrix L equal to identity matrix. Modify the file to use the derivative penalty by copying line 27 from file `DC6_TikhonovD_comp.m` onto line 23 in `DC5_Tikhonov_Lcurve.m`. Then run files `DC1_cont_data_comp.m`, `DC2_discretedata_comp.m` and `DC5_Tikhonov_Lcurve.m`.

Now, assuming the files were in the original settings, i.e. e.g. the parameter $a = 0.04$ in `DC1_cont_data_comp.m`, we should get a reconstruction with 43% error (Figure 5). This is not an optimal reconstruction since, as can be seen from Figure 3, the chosen regularization parameter (red dot) does not correspond to the corner of the L-curve. This is due to the fact that the parameter choice is not automatic (contrary to what is said in the instructions for this exercise) but chosen manually on line 68 in `DC5_Tikhonov_Lcurve.m`. We may now guess and choose a new value for `best_index`, e.g. `best_index = 250`, and run file `DC5_Tikhonov_Lcurve.m` again. Then we get a better reconstruction than previously but the red dot still is quite far away from the corner of the L-curve. By continuing this trial-and-error procedure we may conclude that an index ~ 300 gives the parameter suggested by the L-curve criterion.

(The choice of the regularization parameter can be automated, for example, by finding the point on the L-curve where the curvature is largest. This can be done e.g. by central difference approximations of the derivatives as follows: replace the line 68 in `DC5_Tikhonov_Lcurve.m` by the lines

```
x1    = log(Ax_m_norms);
x2    = log(Lx_norms);
dx1   = (x1(3:end)-x1(1:end-2))/2;
ddx1  = x1(3:end)-2*x1(2:end-1)+x1(1:end-2);
dx2   = (x2(3:end)-x2(1:end-2))/2;
ddx2  = x2(3:end)-2*x2(2:end-1)+x2(1:end-2);
kappa = (dx1.*ddx2-dx2.*ddx1)./( (dx1.^2 + dx2.^2).^(3/2) );
index = find(kappa == max(kappa));
best_index = index+1;
```

Note, however, that this procedure might be numerically unreliable since the computed L-curve is not really a continuum object but consists of discrete data points with errors that might make the computation of the derivatives unstable. Thus, before finding the point of maximal curvature, it is recommendable to apply some type of local smoothing and interpolation to the L-curve as explained in [1, Sec. 7.5.2].)

To save the plot, simply give the command `>> print -depsc convplot.eps` in the command window after running `DC5_Tikhonov_Lcurve.m`, or add the command `print -depsc convplot.eps` to the end of `DC5_Tikhonov_Lcurve.m` and run the file. Then you have the eps figure named 'convplot.eps' in your working directory.

M2. Repeat exercise M1 with the regularization matrix L replaced by second derivative penalty, i.e. the rows of the matrix L should have the form

$$0, \ldots, 0, 1, -2, 1, 0, \ldots, 0.$$

Find the best regularization parameter according to the L-curve method. Is the reconstruction less or more oscillatory than in M1?

Answer

Define the matrix L corresponding to the second derivative penalty by replacing the previous commands on line 22 (and 23) in `DC5_Tikhonov_Lcurve.m` by

```
L = -2*eye(n)+diag(ones(n-1,1),1)+diag(ones(n-1,1),-1);
```

Again, the optimal `best_index` is ~ 300. The reconstruction seems to be a bit less oscillatory than in M1 as expected since we are now penalizing the second derivative of the reconstruction.

You can work on these LaTeX exercises (marked with L) in the exercise session, or you can complete them beforehand.

L1. Download the file `reportdraft1.tex` from the course webpage. Run it with LaTeX; the result should be a pdf document. This document will serve as the template for the report part of the home exam.

L2. Let us try including the picture you created above (in exercise M1) in the LaTeX document. Add the command

```
\usepackage{graphicx}
```

to the second row of `reportdraft1.tex`. Next use this series of commands to add the picture to the Section *Results*:

```
\begin{picture}(320,250)
\put(0,0){\includegraphics[height=8cm]{convplot.eps}}
\end{picture}
```

Try changing the numbers 320, 250, 0 and 8 above one by one and rerunning LaTeX. Can you figure out the meaning of each of the numbers?

L3. Add the following list of references (actually a list with only one entry) to your LaTeX document:

```
\begin{thebibliography}{99}
\bibitem{kurssikirja}
Mueller, J.L. and Siltanen, S: {\em Linear and Nonlinear Inverse
Problems with Practical Applications.}
SIAM 2012. ISBN 978-1-611972-33-7
\end{thebibliography}
```

Include in the text a reference to the book using the command

```
\cite{kurssikirja}
```

Answer: If you have had problems to add the picture you can try to change the format of picture to some other format or try to load usepackage epstopdf.

References

[1] Hansen P. C., *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*, SIAM 1998.