Applied Numerical Methods : LAB8

GOYENS Florentin & WEICKER David

10^{th} December 2015

Introduction

This report presents results for the eighth lab for the course Applied Numerical Method. The problem consists of computing the solution of an underdetermined system using SVD decomposition.

The system to solve comes from the numerical discretization of the Fredholm's integral equation and will sometimes be truncated to a certain rank r.

Building the system

The first first to do is to build the linear system to be solved. The system will come from the discretization of an integral.

Let us define the unknowns. Because we are working on the interval $[0,6]$ with $N = 60$ intervals, with a constant stepsize $h = \frac{6}{N} = 0.1$, we have that the unknows are an approximation of the function p at different point, i.e. :

$$p_i \approx p(x = ih)$$

That gives 61 unknows $(p_0, p_1, ..., p_N, p_{N+1})$. But as stated in the homework, the function $p(x)$ is zero outside $a < x < b$ so we can conclude that $p(a) = p(0) = 0$ and $p(b) = p(6) = 0$. And therefore, $p_0 = p_{N+1} = 0$. That leaves 59 unknowns as predicted $(p_1, p_2, ..., p_N)$.

We secondly have to build the matrix A. It comes form the discretization of the integral using the trapezoidal rule. We have data from 36 measurements, so 36 equations and each is given by (where K is the kernel function given) :

$$\begin{aligned}\int_0^6 K(x, y_i)p(x)dx &= \sum_{j=0}^{59} \int_{hj}^{h(j+1)} K(x, y_i)p(x)dx \\ &\approx \sum_{j=0}^{59} \frac{1}{2h}(K(hj, y_i)p_j + K(h(j+1), y_i)p_{j+1}) \\ &= \frac{1}{h}\sum_{j=1}^{59} K(hj, y_i)p_j\end{aligned}$$

The last equality is found because $p_0 = p_{N+1} = 0$. So the matrix A can be defined by :

$$A = [a_{ij}]$$

$$a_{ij} = \frac{1}{h}K(hj, y_i)$$

for $i = 1, ..., 36$ and $j = 1, ...59$

The next thing to do is get the data vector $f = (f(y_1), ..., f(y_{36}))^T$. In "real" applications, f is the measured data. But here, we will generate it from the given function p (and sometimes add some perturbations). In order to do this, we suppose that the $f(y)$ is a good modeling of our measuring device.

$$\begin{aligned} f_i = f(y_i) &= \int_0^6 K(x, y_i)p(x)dx \\ &= \int_0^6 K(x, y_i)(0.8cos(\pi\frac{x}{6}) - 0.4cos(\pi\frac{x}{2}) + 1)dx \end{aligned}$$

This integral could be performed analytically but we used the Matlab built-in function *integral* instead (which is practically the same since the absolute tolerance of this function is 10^{-10}). The subfunction *data* (available at the end of this report) performs this integration and returns a vector containing the data f.

This yields the following system to be solved in the next section :

$$Ap = f$$

SVD factorization for solving the system

This section focuses on solving the underdetermined system derived in the previous section. To do this, SVD factorization will be used.

We know that our system is composed by 36 equations for 59 unknowns. It is thus indeed underdetermined.

Let us recap the strategy used to solve the system. We start with

$$Ap = f.$$

We get the decomposition

$$A = USV^T.$$

This yields

$$\begin{aligned} USV^Tp &= f && (1) \\ S\underbrace{V^Tp}_{z} &= \underbrace{U^Tf}_{d} && (2) \end{aligned}$$

The matrix S is diagonal and contains the singular values. We decide to keep the r first singular values and set the other ones to zero. This correspond to looking for a least square solution in a subspace of $Col(A)$ but this reduces the sensitivity of the problem to data perturbations.

We set $z_i = d_i/s_i$ for $i = 1 : r$. The other $z_i = 0$ for $i = r+1, \ldots, N$ give the minimum norm solution. Finally we recover the solution $p = Vz$ because we set $z = V^Tp$ in the beginning.

We look at the results we get without adding any noise to the data f. The codes are available at the end of the report. Here starts the game of finding the value of r that works best. We must use enough information but not too much because the sensitivity of the system might cause problems. We first look at figure 1 where we tried the values $r = \begin{bmatrix} 1 & 10 & 20 & 30 \end{bmatrix}$ We see that $r = 1$ does not suffer any oscillation but clearly does not use enough informations. The value $r = 10$ gives an acceptable result. Oscillations start to appear at $r = 20$ and there are many more at $r = 30$.

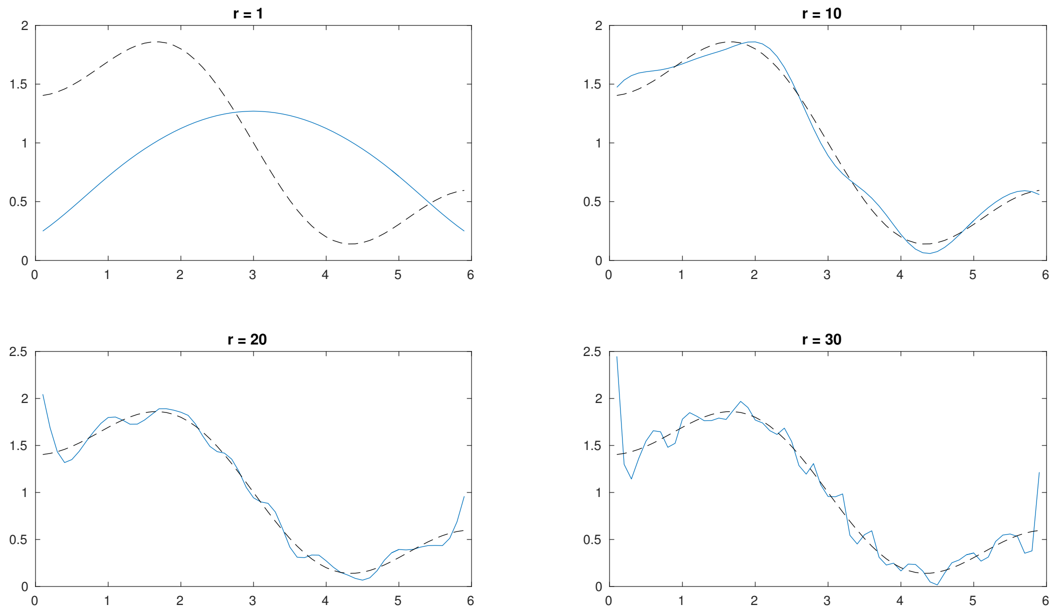

Figure 1: Results for some values of r

Printing the results for other values of r, we see that the best fitting curve is obtained around $r = 10$. On figure 2 we have the best results possible. To choose a value for r, one might have to fix a criteria. It could be, for example, the square error or the maximum difference between the two curves. We won't go into this kind of details but it seems that $r = 9$ or $r = 10$ both give equally great approximation.

Adding perturbations

We are now going to introduce some noise on the measured data. So our vector f is redefined for $i = 1, ..., 36$ by :

$$f_i = f(y_i) + g_i$$

Where g_i's come from a normal distribution with standard deviation 0.01.

The Matlab code *svdPert.m* (available at the end of the report) contains the code used to obtain the plots presented here.

Figure 3 shows the solution computed with the same values of r as for the problem without any perturbation. Comparing to figure **??**, we can see that for the lower values of r ($r = 1$ or $r = 10$), the approximation is quite similar.

On the other hand, for $r = 20$ and $r = 30$, adding perturbations greatly affect the quality of the solution. This is because the problem is ill-posed and thus small variations on the vector f will greatly influence the solution.

Because of the perturbations, we feel, after looking at graphs for different ranks r, that the "best" approximation is when $r = 7$. Figure 4 shows the plot.

We can note here that, because $r = 7$, we are only using 7 singular values (out of the 36 possible!). Although this might seem low, we can see that the approximation is quite good.

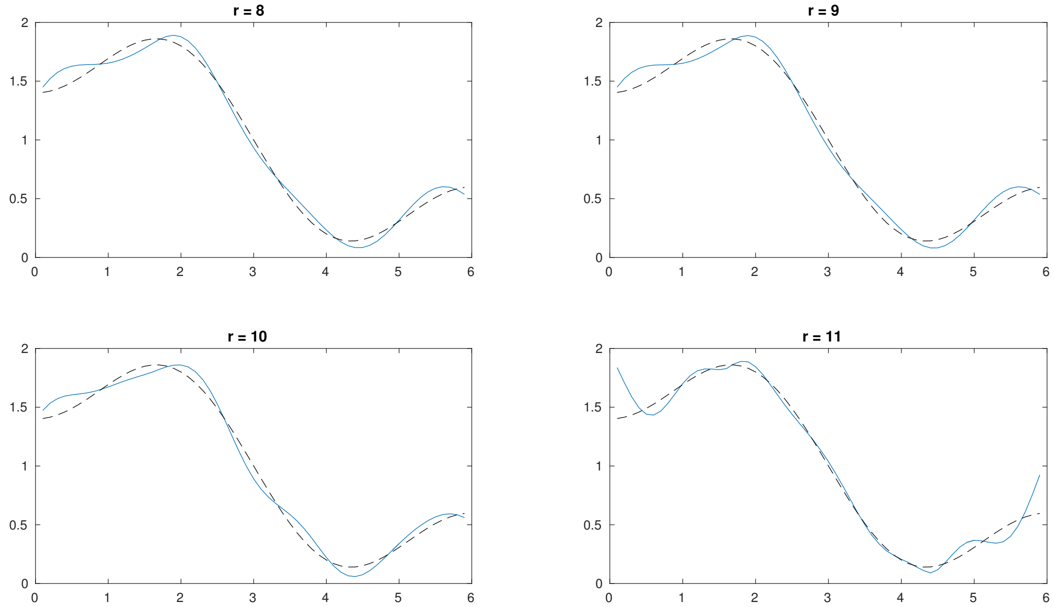

Figure 2: Results for some values of r

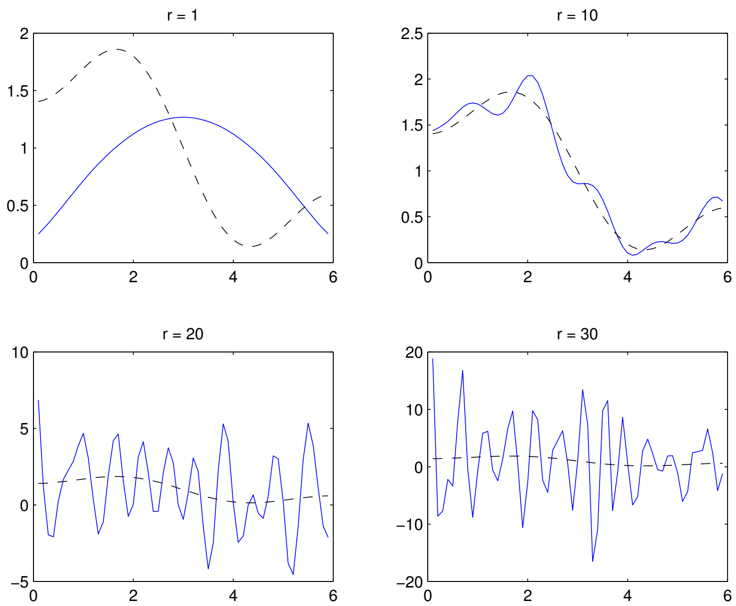

Figure 3: Results for some values of r with perturbations

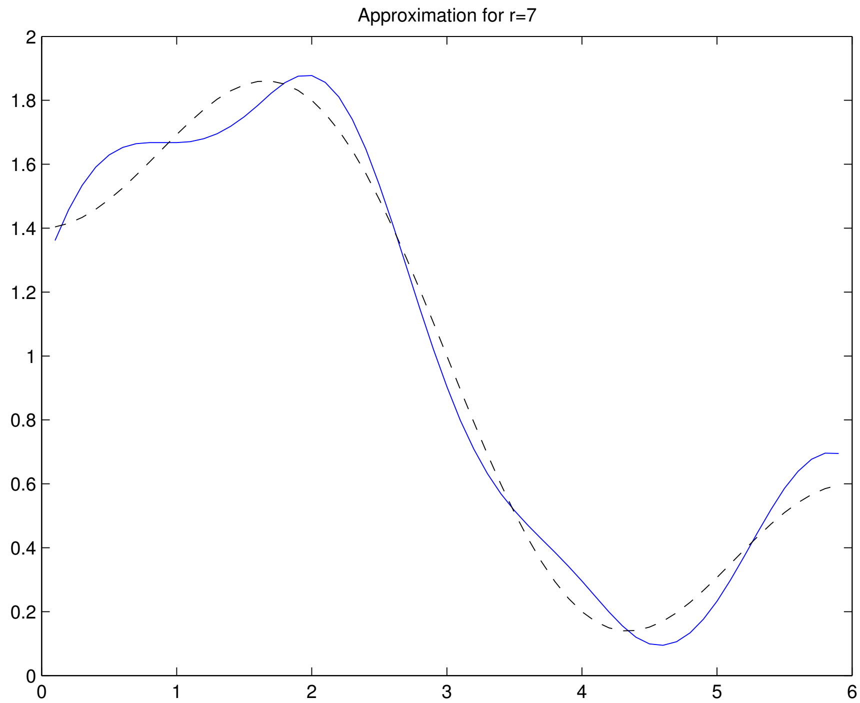

Figure 4: Approximation for $r = 7$