Homework 2

Due date: Feb 17, 2017, 11:55PM.

Bibliography: Course lecture notes Lessons 8-11.

Note: Save this file as `Homework2Solution.tm` in your `/home/student/courses/MATH547/homework` directory. Then carry out your work on solving the problems in TeXmacs. Submit your final version through Sakai. If you've understood the theoretical concepts reviewed in the Background section and lecture notes, each problem should require about 30 minutes to complete, hence the entire homework can be completed within 4 hours.

1 Background

1.1 Objectives, basic concepts

In Homework 2 you apply the Fundamental Theorem of Linear Algebra, and the Gram-Schmidt procedure to solve the problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\| \quad (1)$$

for $\boldsymbol{A}\in\mathbb{R}^{m\times n}, \boldsymbol{b}\in\mathbb{R}^m$, and 2-norm $\|\|:\mathbb{R}^m\to\mathbb{R}_+$, $\|\boldsymbol{y}\|^2=\boldsymbol{y}^T\boldsymbol{y}$. If the solution to (1) leads to $\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|=0$, then $\boldsymbol{x}$ is a solution of the linear system $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$, often denoted as $\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}$, where $\boldsymbol{A}^{-1}$ is the inverse of $\boldsymbol{A}$. If the solution to (1) is such that $\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|>0$, then $\boldsymbol{x}$ is a least-squares solution, or a generalized solution of the linear system, denoted as $\boldsymbol{x}=\boldsymbol{A}^+\boldsymbol{b}$, where $\boldsymbol{A}^+$ is the pseudoinverse of $\boldsymbol{A}$. Least-squares solutions appear in many contexts, and in this homework we study data fitting as an example.

Least-squares solutions are found through projection onto the approximation space $C(\boldsymbol{A})$. The projection is found by orthonormalization of $\boldsymbol{A}$, through the Gram-Schmidt algorithm that furnishes a factorization of $\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}$. The projection operator is subsequently expressed as $\boldsymbol{P}=\boldsymbol{Q}\boldsymbol{Q}^T$, and the least-squares best approximation is given by the solution of the system $\boldsymbol{R}\boldsymbol{x}=\boldsymbol{Q}^T\boldsymbol{b}$ (Figure 1).

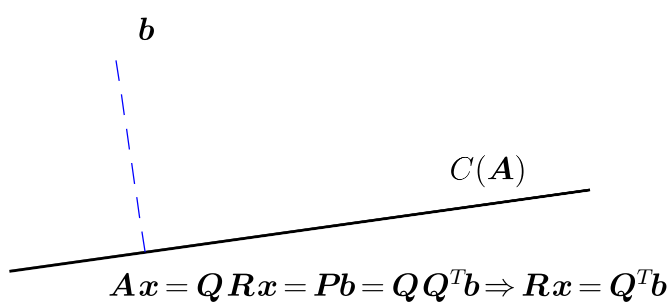

Figure 1. Projection solution of least squares problem

2 Theory problems

2.1 Orthogonal projection

Solve exercises 5.1.27-28. from the textbook, p.217. Solution to exercise 5.1.26 is presented as an example. Complete the rest.

Ex5.1.26 Solution. Place the vectors defining the matrix in matrix $\boldsymbol{A}$

$$\boldsymbol{A}=(\ \boldsymbol{a}_1\ \ \boldsymbol{a}_2\ )=\begin{pmatrix} 2 & 3 \\ 3 & -6 \\ 6 & 2 \end{pmatrix}.$$

Apply Gram-Schmidt to construct an orthonormal basis set for $C(\boldsymbol{A})$

$$\boldsymbol{q}_1 = \frac{1}{\|\boldsymbol{a}_1\|}\boldsymbol{a}_1 = \frac{1}{7}\begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix},$$

$$\boldsymbol{v}_2 = \boldsymbol{a}_2 - (\boldsymbol{q}_1^T\boldsymbol{a}_2)\boldsymbol{q}_1 = \begin{pmatrix} 3 \\ -6 \\ 2 \end{pmatrix} - 0\cdot\boldsymbol{q}_1 = \begin{pmatrix} 3 \\ -6 \\ 2 \end{pmatrix},$$

$$\boldsymbol{q}_2 = \frac{1}{\|\boldsymbol{a}_2\|}\boldsymbol{a}_2 = \frac{1}{7}\begin{pmatrix} 3 \\ -6 \\ 2 \end{pmatrix}.$$

The projector is $\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{Q}^T$ with $\boldsymbol{Q} = (\ \boldsymbol{q}_1 \ \ \boldsymbol{q}_2\ )$

$$\boldsymbol{Q} = \frac{1}{7}\begin{pmatrix} 2 & 3 \\ 3 & -6 \\ 6 & 2 \end{pmatrix} \Rightarrow \boldsymbol{P} = \frac{1}{49}\begin{pmatrix} 2 & 3 \\ 3 & -6 \\ 6 & 2 \end{pmatrix}\begin{pmatrix} 2 & 3 & 6 \\ 3 & -6 & 2 \end{pmatrix} = \frac{1}{49}\begin{pmatrix} 13 & -12 & 18 \\ -12 & 45 & 6 \\ 18 & 6 & 40 \end{pmatrix}.$$

Compute the projection of $\boldsymbol{b}$ onto $C(\boldsymbol{A})$

$$\text{proj}_{C(\boldsymbol{A})}\boldsymbol{b} = \boldsymbol{P}\boldsymbol{b} = \begin{pmatrix} 19 \\ 39 \\ 64 \end{pmatrix}.$$

After carrying out the procedure by hand, check the result of the calculation in Octave

```
octave> A=[2 3; 3 -6; 6 2]; [Q,R]=qr(A,'0'); format rat; Q
   Q =

          -2/7          3/7
          -3/7         -6/7
          -6/7          2/7

octave> P=Q*Q'
   P =

         13/49      -12/49       18/49
        -12/49       45/49        6/49
         18/49        6/49       40/49

octave> b=[49; 49; 49]; P*b
   ans =

            19
            39
            64

octave>
```

Ex5.1.27 Solution.

Ex5.1.28 Solution.

2.2 Working with transposes

Products involving a matrix and its transpose arise very often, and succintly express groupoings of scalar products. Solve exercises 5.3.22-5.3.26 to gain facility in working with transposes. Solution to 5.3.21 is presented as an example

Ex5.3.21 Solution. Denote $\boldsymbol{M}=\boldsymbol{A}^T\boldsymbol{A}$, and compute $\boldsymbol{M}^T=(\boldsymbol{A}^T\boldsymbol{A})^T=\boldsymbol{A}^T\boldsymbol{A}$, hence $\boldsymbol{M}=\boldsymbol{M}^T$.

Ex5.3.22 Solution.

Ex5.3.23 Solution.

Ex5.3.24 Solution.

Ex5.3.25 Solution.

Ex5.3.26 Solution.

2.3 Pseudoinverse matrix

The following exercises introduce the basic properties of pseudoinverses

Ex5.4.8 Solution.

Ex5.4.11 Solution.

2.4 Data fitting

Solve exercises 5.4.31-5.4.34. Solution to 5.4.30 is presented as an example.

Ex5.4.30 Solution. State the problem in matrix form as

$$\min_{\boldsymbol{c}\in\mathbb{R}^2}\|\boldsymbol{A}\boldsymbol{c}-\boldsymbol{b}\|, \boldsymbol{A}=\begin{pmatrix}1 & 0\\ 1 & 0\\ 1 & 1\end{pmatrix}, \boldsymbol{b}=\begin{pmatrix}0\\ 1\\ 1\end{pmatrix}, \boldsymbol{c}=\begin{pmatrix}c_0\\ c_1\end{pmatrix}$$

Compute the $\boldsymbol{Q}\boldsymbol{R}$ factorization of $\boldsymbol{A}$, and the best fit coefficients are the solution of $\boldsymbol{R}\boldsymbol{c}=\boldsymbol{Q}^T\boldsymbol{b}$

```
octave> A=[1 0; 1 0; 1 1]; format rat; [Q,R]=qr(A,'0')
  Q =

    -780/1351   -881/2158
    -780/1351   -881/2158
    -780/1351   3920/4801

  R =

    -1351/780   -780/1351
            0   3920/4801
```

```
octave> b=[0; 1; 1]; c=R\(Q'*b)
  c =

          1/2
          1/2

octave>
```

The linear best fit is $f(t) = \frac{1}{2} + \frac{1}{2}t$. Sketch the solution

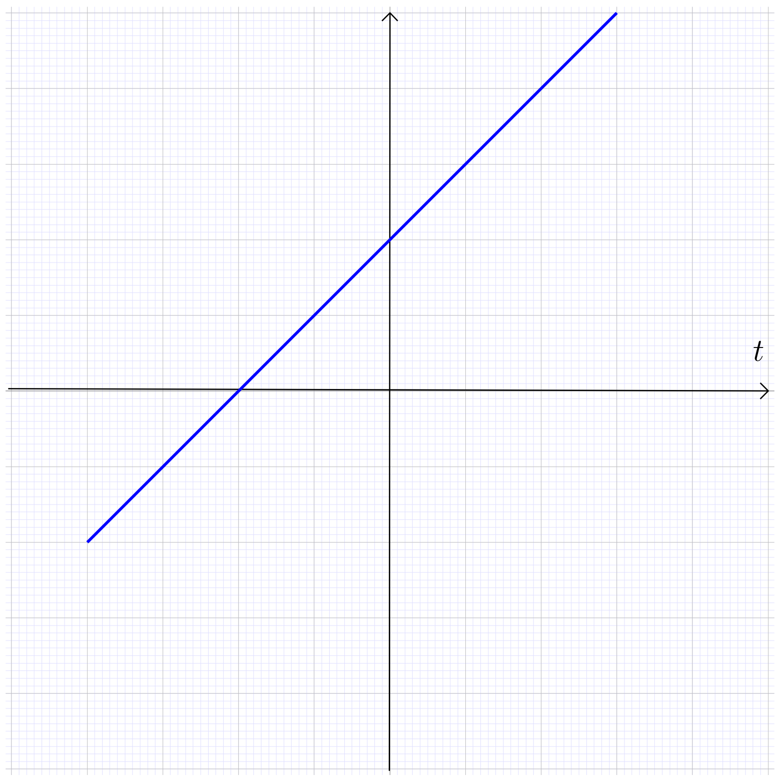

Figure 2.

Ex5.4.31 Solution.

Ex5.4.32 Solution.

Ex5.4.33 Solution.

Ex5.4.34 Solution.

3 Computational problems

We reinforce the theoretical concepts on least squares procedures by consideration of data from electroencephalograms (EEGs). Consult the EEG.tm tutorial.

```
octave> load /home/student/courses/MATH547/lessons/eeg/eeg;
octave> data=EEG.data'; [md nd]=size(data)
  md =  30504
```

```
  nd =  32
octave> pdata=data./max(data)+meshgrid(0:nd-1,0:md-1);
octave> hold on;
        for j=1:nd
          plot(pdata(:,j));
        end;
        hold off;
octave> cd /home/student/courses/MATH547/homework; print -mono -deps eeg.eps;
octave>
```

Divide the overall time recording into nt=30 segments of length m=1000 fron ns=32 sensors.

```
octave> A=reshape(data(1:30000,1:32),1000,30,32); [m nt ns]=size(A)
  m = 1000
  nt = 30
  ns = 32
octave>
```

3.1 Orthogonal projection

Ask the question: are any of the sensors redundant? To answer the question, choose a time slice `i`, and consider the projection of each sensor `j`, `s=A(:,i,j)` onto the space spanned by all the others. Form a matrix of all sensors except sensor `j` by first copying all the signals into a new matrix `B=A(:,i,:)`, and then deleting column `j` through `B(:,i,j)=[]`. Here's a simple example of the syntax:

```
octave> D=[1 2 3; 4 5 6; 7 8 9]
  D =

        1      2      3
        4      5      6
        7      8      9

octave> D(:,2)=[]
  D =

        1      3
        4      6
        7      9

octave>
```

Compute the projection through the same procedure as in the theory question 1

$$\boldsymbol{Q}\boldsymbol{R}=\boldsymbol{B}, \boldsymbol{P}=\boldsymbol{Q}\boldsymbol{Q}^T, \text{proj}_{\boldsymbol{B}}\boldsymbol{s}_j=\boldsymbol{P}\boldsymbol{s}_j,$$

and then compute the ratio of the norm of the projection to that of the original sensor signal

$$f_j=\frac{\|\boldsymbol{P}\boldsymbol{s}_j\|}{\|\boldsymbol{s}_j\|}, \text{for } j=1,\ldots,32.$$

Are any of the sensors redundant? Quantitatively state your criterion for determining whether a sensor is redundant.

3.2 Fourier analysis

Carry out a practical example of the procedure outlined in Exercise 5.4.35. For some time slice, fit a sum of trigonometric functions to the available sensor data

$$s_j(t) \cong f_j(t) = a_0^{(j)} + \sum_{k=1}^{16} \left[a_k^{(j)} \cos\left(\frac{2\pi}{m} kt\right) + b_k^{(j)} \sin\left(\frac{2\pi}{m} kt\right)\right], \text{for } j = 1, \ldots, 32.$$

Use the same procedure as in theory question 4. Choose times `t=0:m-1`. The quantity $p_k^{(j)} = \sqrt{[a_k^{(j)}]^2 + [b_k^{(j)}]^2}$ is the power of frequency k in the signal. Present a table of $p_k^{(j)}$. Which frequency is dominant in each sensor? Do all sensors show the same dominant frequency, i.e., are waves in different regions of the brain synchronous?

Note: Both questions in this section are direct applications of the procedures from theory questions 1 and 4. Be sure you understood the theory completely if you find the applications difficult.