Seminar assignment 1

Each seminar assignment consists of two parts, one part consisting of $3 - 5$ elementary problems for a maximum of 10 points from each assignment. For the second part consisting of problems in applications you are to choose one of three problems to solve. This part can give up to 5 points from each assignment. Each seminar assignment can give a maximum of 15 points, to pass you will need at least 20 points total from both the assignments.

The first part consists of elementary questions to make sure that you have understood the basic material of the course while the second part consists of one larger application example.

Solutions can either be handed in by email to *christopher.engstrom@mdh.se* or *karl.lundengard@mdh.se* or you can hand in handwritten solutions either during the lectures or in the box outside Christopher and Karl's room $U3 - 185$.

The solutions should be submitted at the latest the 21st of december.

1 Part 1

In the first part you are to solve and hand in solutions to the questions. You are allowed to use computer software to check your results, but your calculations as well as your result should be included in the answers for full points.

1.1

Let A and B be defined as follows

$$A = \begin{bmatrix} 1 & -1 & -1 \\ 2 & -1 & -2 \\ 0 & -1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Then A and B will be similar in the mathematical sense $A = S^{-1}BS$. A consequence if this is that A and B have the same eigenvalues.

- (1) Write down the companion matrix to the polynomial $p(x) = x^3 - x$.
- (1) Use the companion matrix to find the roots of $p(x)$
  Hint: your answer to 1.1 in assignment 1 will be useful here.

1.2

Consider the matrix equation

$$XB = C \tag{1}$$

where

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}, \; B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \; C = \begin{bmatrix} 17 & 26 \\ 4 & 6 \end{bmatrix}$$

- (1) Use the Kroenecker product to rewrite equation (1) as

$$Mx = c$$

where M is a 4×4 matrix and x and c are vectors with 4 elements.

- (1) Find the matrix X in equation (1).

1.3

- (1) Calculate the $\|A\|_{(4)}$ where A is the matrix in problem 1.1.

1.4

- (1) Write the definition of a Inner product space and give one example of one Inner product space.

1.5

- (1) Show that the function $T : \mathbb{R}^3 \to \mathbb{R}^2$ is a linear transformation:

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y - z \\ 2x + z \end{bmatrix}$$

1.6

- (1) What is the QR-factorization of a matrix? Give a short description of the matrices in the factorization.

- (1) Given $\vec{u}, \vec{v}$ below:

$$\vec{u} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Calculate the projection of $\vec{v}$ on $\vec{u}$:

$$\text{proj}_{\vec{u}}(\vec{v}) = \frac{\langle \vec{v}, \vec{u} \rangle}{||\vec{u}||^2} \vec{u}$$

- (1) Use the Gram-Schmidt process to calculate the QR-factorization of

$$A = \begin{bmatrix} 2 & 1 & -1 \\ -2 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix}$$

Part 2

In this second part you are to choose **ONE** example where you attempt to solve the questions presented. If you hand in answers to more than one choice you will get points corresponding to the choice which would give the least total points.

You are allowed to use computer software (but it shouldn't be needed) to solve the questions as long as you make it clear what you are doing and give an example of what could be used to solve it by hand. For example you could write "I solved the linear system $Ax = b$ using Matlabs 'fsolve', another method would have been to use Gaussian elimination and solving the found triangular system".

Remark: There are two different QR-factorizations of a matrix. If we QR-factorize a $n \times r$ matrix A in the way you have discussed during the lecture the Q matrix is square $n \times n$ and the triangular matrix is non-square $n \times r$. An alternative QR-factorization is taking Q non-square $n \times r$ and R square $r \times r$. These two ways are closely related and if A is square they are exactly the same. If you have calculated the first version (using Gram-Schmidt or some software) you can get the second version by taking the first r rows of R to get the new R and the first r columns of Q to get the new Q.

$$\textbf{Example: } A = \begin{bmatrix} 9 & 12 \\ 9 & 10 \\ 3 & 4 \end{bmatrix} = \underbrace{\begin{bmatrix} 3 & 3 & 1 \\ 3 & 2 & 2 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}}_{\text{first version}} = \underbrace{\begin{bmatrix} 3 & 3 \\ 3 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 0 & 2 \end{bmatrix}}_{\text{second version}}$$

For the problems in the second part of the assignment that involves QR-factorization we recommend the second version.

1.7 Estimating wind power using LSM

In recent years wind power has become more and more popular. It is still a limited power resource since wind turbines need to be placed in areas with specific wind characteristics in order to be sustainable, both from an environmental and financial perspective.

A large wind turbine can typically have a hub height of 80 m. Performing wind measurements at that height in many places in order to find suitable spots can be complicated and tedious. Something that would be helpful would be a method that allowed the wind speed at 80 m to be predicted from measurements done at a more convenient height, typically 10 m. one way of creating such a method is taking more detailed measurements at one site and fit the measured data to some approximate model for wind speed. Hopefully the wind characteristics will be similar near the site and the rest of the area can be evaluated by measurement at a lower height.

In table 1 wind speed measurements for one such site are given.

v (m/s)	6.4	6.7	7.5	7.8	8.8
z (m)	10	22	45	60	75

Table 1: Table of wind speeds, v, at different heights, z.

- (1) One way to model wind speed would be to assume that it increases linearly with the height, $v(z) = a + bz$. Find the parameters a and b using the least square method.
- (1) A physically more sensible model is to assume that the relation between wind speed and height is a so called *power law*

  $$v(z) = v_r \left(\frac{z}{z_r} \right)^{\alpha}$$

  Where v_r is the wind speed at 10 m and $z_r = 10$ m is the reference height. Estimate the parameter α using the least square method.
  Hint: try using logarithms.

Another important circumstance for wind power is that the wind speed (and thereby power generation) varies over time. If this variation can be predicted then it is easier to plan how other energy sources can compensate when the wind turbines can not supply enough power. During different times of the year the wind can be more or less unpredictable. To account for this the wind speed at a specific time during the year can be measured several times and from this we can calculate the average wind speed $\bar{v}$ and the variance of the wind speed σ^2. An example of this kind of measurements can be found in table 2. If the average wind speed has high variance this means that it is difficult to predict. In these areas we can allow a larger difference between the prediction and the average and still get a sensible prediction than we can in areas with small variance. To compensate for this we can use the weighted square method.

$$A^\top W A c = A^\top W \bar{v}$$

Time t	Average wind speed $\bar{v}$ (m/s)	Wind speed variance σ^2
1	5.1	1.2
3	6.2	3.1
5	5.9	0.8
7	4.8	1.2
9	4.5	1.0
11	5.3	1.1

Table 2: Table of average wind speeds, $\bar{v}$, and wind speed variance, σ^2, for different times, t.

where $\bar{v}$ is a vector containing all the average wind speeds, c is a vector of parameter for the function we want to fit to the data and W is a weight matrix that look like this

$$W = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_6^2} \end{bmatrix}$$

Suppose we want to model the variation over time using a 2nd degree polynomial.

$$\bar{v}(t) = c_0 + c_1 t + c_2 t^2$$

- (1) Find a root of W (any root will do).
- (1) Show that if you set

$$\tilde{A} = W^{\frac{1}{2}} A \text{ and } \tilde{v} = W^{\frac{1}{2}} \bar{v}$$

  You get a normal least square method equation

$$\tilde{A}^\top \tilde{A} c = \tilde{A}^\top \tilde{v}$$

Using the QR-factorization of $\tilde{A} = QR$ we can rewrite the least square method equation as

$$Rc = Q\tilde{v}$$

You do **not** need to show this.

- (1) Use the Gram-Schmidt process to calculate the QR-factorization of $\tilde{A}$ and solve $Rc = Q^\top \tilde{v}$ (Note the remark at the beginning of part 2).

1.8 Path finding for a robot

There is a lot of interest and effort put into constructing robots that can perform tasks that are dangerous or impossible for humans to perform. One such task could be to into small areas such as sewers, caves or collapsed buildings. In these kind of circumstances it can be difficult to keep contact with the robot since wires would get stuck and wireless communication might be blocked by rock, earth or concrete. Therefore a robot that can find its own path through an area could be very useful.

Suppose you have a robot that has a sensor that allows it to identify 4 points that should lie on the optimal path through an area. We assume that any nice and smooth curve that comes near these 4 points is going to be close to the optimal path.

In a test area the robot identifies the four points listed in table 3 as points on the optimal path.

x	0.3	0.9	2.1	2.7
y	1.0	1.1	0.2	-0.9

Table 3: Table of points along the optimal path.

- (1) Use the least square method to find a straight line that comes as close to the 4 points as possible.
- (1) Use the least square method to find a 2nd degree polynomial curve that comes as close to the 4 points as possible. Compare the residuals for the two curves you have calculated in this and the previous problem.
- (1) Without calculating any numbers try to predict what residuals you are going to get if you choose a 3rd degree polynomial as the path.
  Hint: you should not calculate any numbers, but you can still write matrices and equations.

Another feature of the robot is that it can fire a small flare through holes to light up areas or to make it easier to locate the robot if contact is lost with the operator for two long. The rockets are cheap an light and can therefore not fly entirely straight. Test show that the altitude of the rocket can be accurately modeled by the equation

$$h(t) = h_0 + \alpha e^{0.7t} + k\sin(5t)$$

where h_0 is the initial altitude, α is a parameter that determines how fast fuel is burnt and k is a parameter dependent on the shape of the rockets fins.

In the test area we want the robot to fire the rocket in a trajectory that passes near the four points in table 4.

- (1) The condition that the rocket should pass through the correct height at each point can be written as a linear equation system. Write this system on matrix form

$$Ac = h$$

t	1.0	2.0	3.0	4.0
h	2.5	4.3	9.0	16.3

Table 4: The desired altitude of the rocket, h, at different times, t.

where $c = [h_0 \ \alpha \ k]^\top$.

- (1) We can find the least square solution to $Ac = h$ by solving the normal equations

$$A^\top Ac = A^\top h$$

Using the QR-factorization of $A = QR$ we can rewrite the normal equations as

$$Rc = Q^\top h$$

Use the Gram-Schmidt process to calculate the QR-factorization of A and solve $Rc = Q^\top h$ (Note the remark at the beginning of part 2).

1.9 Estimating production function parameters

The *production function* is a concept in economics that can be used to describe and analyze systems of different sizes, from a firm to a whole industry, nation or economy. An interesting property for production functions is constant elasticity of substitution (CES). Elasticity of substitution can be described as a measure of how easy it is to substitute one input to a producing system with another input. One example of inputs you might want to substitute for one another is capital and labor. The general CES-production function for capital and labor is written

$$p = f(ak^r + (1-a)l^r)^{\frac{1}{r}}$$

Here p is the production, k is the capital, l is the labor, f is the productivity factor and a is the share parameter.
Two special cases of this production function is the *linear production* function and the *Cobb-Douglas function.*

$$\text{Linear production function: } q = f \cdot (ak + (1-a)l)$$
$$\text{Cobb-Douglas function: } q = f \cdot k^a \cdot l^{1-a}$$

These two functions can be derived from the CES-function by letting $r = 1$ and $r \to \infty$ respectively.
In table 5 you can find the relative production of 6 different factories. All the factories produce the same good but use different methods of production, some use labor-intensive methods and some use capital-intensive methods.

- (1) Use the least square method to estimate the parameters a and f in the linear production function.

Labor l	Capital k	Production p	Variance σ^2
1.0	1.0	1.2	0.1
1.1	0.9	2.3	0.4
1.3	0.8	0.8	0.2
1.8	0.5	1.2	0.5
0.6	2.1	1.0	0.7
0.8	1.6	1.1	0.4

Table 5: Table of relative labor, capital and production for 6 different factories.

- (1) Use the least square method to estimate the parameters a and f in the Cobb-Douglas production function.
  Hint: try using logarithms.

The production shown in column 3 in table 5 is the average production of the plant. The actual production varies from day to day. This is reflected in column 4 where the variances of the production is listed. A factory with a small variance gives a better estimate of the production function. In the two estimations you have already done this is not taken into account. To take this property of the factories into account we will use the following matrix

$$W = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_6^2} \end{bmatrix}$$

This matrix will be taken as the weight matrix in the weighted least square method $A^\top WAi = A^\top Wp$, here i is a vector containing the coefficients in the linear productivity function. Then the residuals for a factory with large variance will be smaller and the residuals of a factory with small variance will be enlarged. This means that the will be fitted closer to the points corresponding to factories with low variance.

- (1) Find a root of W (any root will do).
- (1) Show that if you set

$$\tilde{A} = W^{\frac{1}{2}}A \text{ and } \tilde{p} = W^{\frac{1}{2}}p$$

  You get a normal least square method equation

$$\tilde{A}^\top \tilde{A}i = \tilde{A}^\top \tilde{p}$$

Using the QR-factorization of $\tilde{A} = QR$ we can rewrite the least square method equation as

$$Ri = Q^\top \tilde{p}$$

You do **not** need to show this.

- (1) Use the Gram-Schmidt process to calculate the QR-factorization of $\tilde{A}$ and solve $Ri = Q^\top \tilde{p}$ (Note the remark at the beginning of part 2).