# Exercises for Week 6

## 1 LSQ problem (by hand)

1. Show that the objective function $\rho(\boldsymbol{x}) = \|\boldsymbol{y} - A\boldsymbol{x}\|_2^2$ in the linear LSQ problem in convex.

2. Consider the rank deficient case in the SVD approach to obtain the LSQ solution, see lecture slides page 27.

   (a) Show that $\boldsymbol{x}^* = \sum_{i=1}^{r} \frac{\boldsymbol{u}_i^T \boldsymbol{y}}{\sigma_i} \boldsymbol{v}_i$ is a minimizer of the LSQ problem.

   (b) Find $\|\boldsymbol{x}^*\|_2^2$ and show that $\boldsymbol{x}^*$ is the solution to the LSQ problem with minimal $l_2$-norm.

## 2 Linear least squares fit (in Matlab)

We are given 5 data points

| $i$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $t_i$ | -2 | -1 | 0 | 1 | 2 |
| $y_i$ | 0 | 15 | 25 | 50 | 110 |

and we wish to compute a least squares fit with the model function

$$\phi(\boldsymbol{x}, t) = \sum_{j=1}^{3} x_j f_j(t), \qquad f_1(t) = 1, \quad f_2(t) = t, \quad f_3(t) = t^2 - 2.$$

1. Form the $5 \times 3$ coefficient matrix $A$ and the normal equations.

2. Write a Matlab function

```
function x = linearLSQ(A, y)
```

   to implement the QR approach to compute the least squares solution $\boldsymbol{x}^*$.

3. Compute the residual $\boldsymbol{r}^*$ at the least squares solution $\boldsymbol{x}^*$, and use this to estimate the standard deviation $\varsigma$ of the errors in the data according to $\|\boldsymbol{r}^*\|_2^2 \approx (m - n)\,\varsigma^2$.

## 3 Trigonometrical fit

| date | 11/1 | 25/1 | 26/1 | 31/1 | 2/2 | 5/2 | 16/2 | 27/2 | 16/3 | 20/3 |
|---|---|---|---|---|---|---|---|---|---|---|
| $x$ | 11 | 25 | 26 | 31 | 33 | 36 | 47 | 58 | 75 | 79 |
| $y$ | 160 | 140 | 138 | 130 | 125 | 120 | 95 | 72 | 27 | 17 |

Mr. Madsen have recorded the time of the sunrise from his kitchen window, and the results are shown in the above table. Here, each element in $x$ gives the number of days starting from 1st January, and each element in $y$ gives the time of the sunrise in minutes after 6am. Note: the values of $y$ do not take the "daylight saving time" into account.

Since the movement of the sun to the earth can be assumed as periodic, we use trigonometric functions to fit these data:

$$F(x) = c_0 + c_1 \sin(\omega x) + c_2 \cos(\omega x) \ , \qquad \omega = 2\pi/365, \tag{1}$$

where $\omega$ is chosen such that the period of $F$ is 365 days. Compute the least-squares fit by calling `LinearLSQ`, and plot the fit function (as a line) with the data points (as points).

Now based on our fit function $F$ we can predict when the sunrise is the earliest. Basically, we need calculate the function values for $x = 80, 81, 82, \ldots, 365$, and find the minimun in $F(80), F(81), F(82), \ldots, F(365)$. Is it more or less what you expected?

## 4 Polynomial fit

In this exercise, we try to fit the data in `data_exe4.mat` with a polynomial of degree $d$,

$$\phi(\boldsymbol{x}, t) = x_1 t^d + x_2 t^{d-1} + \cdots + x_d t + x_{d+1}.$$

Note that the coefficient vector $\boldsymbol{x}$ has $n = d+1$ elements.

1. Set $n = 2, 3, 4, 5$ and 6, respectively, and compute the least-squares fits by calling `LinearLSQ`. Plot the fit functions (as a line) with the data points (as points). Visually, which value of $n$ gives the best fit function?

2. The proper choice of $n$ can be quantified by using autocorrelation test. We can estimate the standard deviation $\varsigma$ of the errors in the data according to $\|\boldsymbol{r}^*\|_2^2 \approx (m-n)\,\varsigma^2$. If $n$ is too small, $\varsigma_n$ will decrease for larger $n$. The estimate $\varsigma_n$ will settle at an almost constant value when we have passed the optimal value of $n$. This is where the level of approximation error is below the noise level. Compute the standard deviation $\varsigma_n$ for $n = 2, 3, 4, 5$ and 6, and plot the values of $\varsigma_n$ as points. According to the plot, which $n$ would be the optimal choice?