Problem Sheet 8 - Solutions

1. We can write the overdetermined linear system as

$$\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix} x = \begin{pmatrix}m_1\\m_2\\\vdots\\m_n\end{pmatrix}.$$

In order to solve it system we shall use normal equations. Denote $\mathbf{A} = \begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix}m_1\\m_2\\\vdots\\m_n\end{pmatrix}$.

Then $\mathbf{A}^\top\mathbf{A} = n$, and $\mathbf{A}^\top b = m_1 + \ldots + m_n$. Therefore

$$\mathbf{A}^\top\mathbf{A}x = \mathbf{A}^\top\mathbf{b} \Rightarrow nx = m_1 + \ldots + m_n.$$

Therefore, the least squares solution is the arithmetic mean of the measurements

$$x = \frac{m_1 + \ldots + m_n}{n}.$$

2. **a)** We have

$$y_i = f(t_i) = \alpha e^{\beta t_i}.$$

This is the governing non-linear system of equations.

b) Let $g(t) = \log f(t)$, which is well defined since $f(t) > 0$ for all t. Therefore, we have

$$g(t) = \log f(t) = \log\alpha + \log e^{\beta t} = a + \beta t,$$

where we have defined $a = \log\alpha$. In an accordance with g we also, define $b_i :\ = \log y_i$. Consequently, this freshly linearised problem can now be written as an overdetermined system of linear equations,

$$\mathbf{A} = \begin{pmatrix}1 & t_1\\1 & t_2\\\vdots & \vdots\\1 & t_n\end{pmatrix},\ \mathbf{x} = \begin{pmatrix}a\\\beta\end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}, \text{ so that } \mathbf{Ax} = \mathbf{b}.$$

c) Let us compute the parameters of governing normal equations $\mathbf{A}^\top\mathbf{Ax} = \mathbf{A}^\top\mathbf{b}$. We have

$$\mathbf{A}^\top\mathbf{A} = \begin{pmatrix}n & \sum_{i=1}^n t_i\\ \sum_{i=1}^n t_i & \sum_{i=1}^n t_i^2\end{pmatrix} \text{ and } \mathbf{A}^\top\mathbf{b} = \begin{pmatrix}\sum_{i=1}^n b_i\\ \sum_{i=1}^n t_i b_i\end{pmatrix}.$$

Now, once we find the least squares solution $\mathbf{x} = \begin{pmatrix}a\\\beta\end{pmatrix}$, we take $\alpha = e^a$ and β to be our approximative parameters for f.

d)

```matlab
% Problem 2d
% INPUT
% t, y - parameters, such that f(t(i)) = y(i)
% OUTPUT
% alpha, beta - values such that alpha*exp(beta*t) approximates f, thru
% least squares.

function [alpha, beta] = ExpoFuncFit(t, y)
    % Compute the solution of the linearised problem
    x = linearregression(t, log(y));
    % Set the proper output values
    alpha = exp(x(2));
    beta = x(1);
end

function x = linearregression(t,y)
% Solution of linear regression problem (fitting of a line to data) for
% data points \Blue{$(t_i,y_i)$}, \Blue{$i=1,\ldots,n$} passed in the
% \emph{column vectors} \texttt{t} and \texttt{y}.
% The return value is a 2-vector, containing the slope of the fitted line
% in x(1) and its offset in x(2)
n = length(t); if (length(y) ~= n), error('data␣size␣mismatch'); end
% Coefficient matrix of \textbf{overdetermined linear system}
A = [t,ones(n,1)];
% \Red{Determine least squares solution by using MATLAB's $\backslash$
% operator}
x = A\y;
end
```

3. a) Since function f is a linear combination of g_1 and g_2, it can be written as

$$f(x) = \alpha g_1(x) + \beta g_2(x) = \alpha 2^x + \beta 2^{-x},$$

where we have to approximate $\alpha, \beta \in \mathbb{R}$. We have

$$\begin{aligned}
\frac{1}{4}\alpha + 4\beta + 8 &= r_1 \\
\frac{1}{2}\alpha + 2\beta + 4 &= r_2 \\
\alpha + \beta + 2 &= r_3 \\
2\alpha + \frac{1}{2}b - 4 &= r_4 \\
4\alpha + \frac{1}{4}b - 12 &= r_5.
\end{aligned}$$

as the equations for the values of the error vector $\mathbf{r} = \|\mathbf{b} - \mathbf{A}\mathbf{x}^*\|$, where

$$\mathbf{A} = \begin{pmatrix} \frac{1}{4} & 4 \\ \frac{1}{2} & 2 \\ 1 & 1 \\ 2 & \frac{1}{2} \\ 4 & \frac{1}{4} \end{pmatrix}, \ \mathbf{b} = \begin{pmatrix} -8 \\ -4 \\ -2 \\ 4 \\ 12 \end{pmatrix} \ \text{and} \ \mathbf{x} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

b) We have

$$\mathbf{A}^\top \mathbf{A} = \begin{pmatrix} \frac{341}{16} & 5 \\ 5 & \frac{341}{16} \end{pmatrix}, \mathbf{A}^\top \mathbf{b} = \begin{pmatrix} 50 \\ -37 \end{pmatrix}.$$

c) Solving normal equations $\mathbf{A}^\top \mathbf{A}\mathbf{x} = \mathbf{A}^\top \mathbf{b}$ yields

$$\alpha = \frac{3680}{1263} \quad \text{and} \quad b = -\frac{3056}{1263}.$$

4. **a)** There are $\binom{n}{2}$ values of d_{ij}.

b) Let us define

$$\mathbf{A} = \begin{pmatrix} -1 & 1 & 0 & \cdots & & & \\ -1 & 0 & 1 & 0 & \cdots & & \\ 0 & -1 & 1 & 0 & \cdots & & \\ -1 & 0 & 0 & 1 & 0 & \cdots & \\ 0 & -1 & 0 & 1 & 0 & \cdots & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

and

$$\mathbf{d} = \begin{pmatrix} d_{21} \\ d_{31} \\ d_{32} \\ d_{41} \\ d_{42} \\ \vdots \end{pmatrix}$$

Then our overdetermined system is $\mathbf{A}\mathbf{x} = \mathbf{d}$, where $\mathbf{x} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix}$. In conclusion, $\mathbf{A} \in \mathbb{R}^{\binom{n}{2}, n}$, $\mathbf{d} \in \mathbb{R}^{\binom{n}{2}}$ and $\mathbf{x} \in \mathbb{R}^n$.

c) Let us compute the $\ker(A)$. We want to find all $\mathbf{x}$ such that $\mathbf{A}\mathbf{x} = \mathbf{0}$. We have

$$\begin{pmatrix} -1 & 1 & 0 & \cdots & & & \\ -1 & 0 & 1 & 0 & \cdots & & \\ 0 & -1 & 1 & 0 & \cdots & & \\ -1 & 0 & 0 & 1 & 0 & \cdots & \\ 0 & -1 & 0 & 1 & 0 & \cdots & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

This is equivalent to having $p_i - p_j = 0$ for all $n \geq i > j \geq 1$. In other words, we have $p_i = p_j$ for all i, j. Therefore, $\mathbf{A}\mathbf{x} = \mathbf{0}$ if and only if $p_i = \alpha$, for all $i = 1, \ldots, n$ and an arbitrary $\alpha \in \mathbb{R}$. Thus $\ker(A) = \mathsf{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right\}$. This implies that $\dim(\ker(A)) = 1$, therefore, $\operatorname{rank}(A) < n$, which means that the equation $\mathbf{A}\mathbf{x} = \mathbf{d}$ will not have a unique solution, since it violates the condition 3.9.2.E.

Let us consider a different system, Instead of considering p_i let us shift the whole system by p_1, in other words, define $\tilde{p}_i = p_i - p_1$. Notice that we now still have

$$\tilde{p}_i - \tilde{p}_j = p_i - p_1 - p_j + p_1 = d_{ij}.$$

We also have $\tilde{p}_1 = 0$. Therefore, solving

$$\mathbf{A} = \begin{pmatrix} -1 & 1 & 0 & \cdots & & & \\ -1 & 0 & 1 & 0 & \cdots & & \\ 0 & -1 & 1 & 0 & \cdots & & \\ -1 & 0 & 0 & 1 & 0 & \cdots & \\ 0 & -1 & 0 & 1 & 0 & \cdots & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} 0 \\ \tilde{p}_2 \\ \vdots \\ \tilde{p}_n \end{pmatrix} = \begin{pmatrix} d_{21} \\ d_{31} \\ d_{32} \\ d_{41} \\ d_{42} \\ \vdots \end{pmatrix}$$

is equivalent to solving $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \mathbf{d}$, where

$$\tilde{\mathbf{A}} = \begin{pmatrix} 1 & 0 & \cdots & & & & \\ 0 & 1 & 0 & \cdots & & & \\ -1 & 1 & 0 & \cdots & & & \\ 0 & 0 & 1 & 0 & \cdots & & \\ -1 & 0 & 1 & 0 & \cdots & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

is created by removing the first column of $\mathbf{A}$ and $\tilde{\mathbf{x}} = \begin{pmatrix} \tilde{p}_2 \\ \tilde{p}_3 \\ \vdots \\ \tilde{p}_n \end{pmatrix}$. This modified matrix $\binom{n}{2} \times (n-1)$ matrix $\tilde{\mathbf{A}}$ has an empty kernel, since solving $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \mathbf{0}$ gives $\tilde{p}_2 = 0, \tilde{p}_3 = 0, \ldots$ and so on. Therefore, $\tilde{\mathbf{A}}$ is a matrix of full rank (as $\dim(\ker(A)) = 0$), and it admits a unique least squares solution.

d) For $n = 5$, the matrix of the modified system is given by

$$\tilde{\mathbf{A}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix}$$

so that the overdetermined system is given by

$$\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \begin{pmatrix} d_{21} \\ d_{31} \\ d_{32} \\ d_{41} \\ d_{42} \\ d_{43} \\ d_{51} \\ d_{52} \\ d_{53} \\ d_{54} \end{pmatrix}.$$

Let us write down the normal equations. We compute

$$\tilde{\mathbf{A}}^\top \tilde{\mathbf{A}} = \begin{pmatrix} 4 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 \\ -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & 4 \end{pmatrix}$$

and

$$\tilde{\mathbf{A}}^\top \mathbf{d} = \begin{pmatrix} d_{21} - d_{32} - d_{42} - d_{52} \\ d_{31} + d_{32} - d_{43} - d_{53} \\ d_{41} + d_{42} + d_{43} - d_{54} \\ d_{51} + d_{52} + d_{53} + d_{54} \end{pmatrix}.$$

Hence, the normal equations are $\tilde{\mathbf{A}}^\top \tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{A}}^\top \mathbf{d}$.

e)

```
% Problem 4e
% INPUT
% D - a strictly lower triangular matrix
% OUTPUT
% p - vector containing the shited values, \tilde{p}_2,...,\tilde{p}_n

function p = RoadLengths(D)

% Find the indices of non-zero entries
[I,J] = find(D' > 0);
% Determine the size of A
m = size(D, 1);
n = length(I);
% Build A, in a sparse form
A = sparse([(1:n)';(1:n)'],[I;J],[-ones(n,1);ones(n,1)],n,m);
% Remove A's first column to ensure uniqueness of our least squares solution
A = A(:,2:end);
% Extract the right hand side vector
d = nonzeros(D');
% Finally, solve the equation
p = A\d;
```

5. a) Let us show that

$$\begin{pmatrix} \mathbf{A}^\top & \mathbf{0} \\ \mathbf{I} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{b} \end{pmatrix} \tag{1}$$

implies $\mathbf{A}^\top\mathbf{A}\mathbf{x} = \mathbf{A}^\top\mathbf{b}$. Multiplying (1) through we have

$$\begin{aligned} \mathbf{A}^\top\mathbf{r} &= \mathbf{0} \\ \mathbf{r} + \mathbf{A}\mathbf{x} &= \mathbf{b}. \end{aligned}$$

Multiplying $\mathbf{r} + \mathbf{A}\mathbf{x} = \mathbf{b}$ by $\mathbf{A}^\top$ from the left we have

$$\mathbf{A}^\top\mathbf{r} + \mathbf{A}^\top\mathbf{A}\mathbf{x} = \mathbf{A}^\top\mathbf{b}.$$

Plugging in $\mathbf{A}^\top\mathbf{r} = \mathbf{0}$ into the preceding equation gives $\mathbf{A}^\top\mathbf{A}\mathbf{x} = \mathbf{A}^\top\mathbf{b}$.

Conversely, let us assume that $\mathbf{A}^\top\mathbf{A}\mathbf{x} = \mathbf{A}^\top\mathbf{b}$ holds. Define $\mathbf{r} = \mathbf{A}\mathbf{x} - \mathbf{b}$. Recall that we can do that since $\mathbf{A}, \mathbf{x}$ and $\mathbf{b}$ are known to us. Therefore, the what remains to be proven is that such an $\mathbf{r}$ satisfies $\mathbf{A}^\top\mathbf{r} = \mathbf{0}$, but we have

$$\mathbf{A}^\top\mathbf{r} = \mathbf{A}^\top\mathbf{A}\mathbf{x} - \mathbf{A}^\top\mathbf{b} = \mathbf{0},$$

since $\mathbf{x}$ is the solution of our normal equations. Hence, the converse also holds.

b) From **a)** we have $\mathbf{r} = \mathbf{A}\mathbf{x} - \mathbf{b}$.

c) In problems **4.d)** and **4.e)** the overdetermined matrix of the system, $\mathbf{A}$, was sparse, but the matrix $\mathbf{A}^\top\mathbf{A}$ which concerns normal equations, was a dense matrix. Therefore, a major benefit of $\begin{pmatrix} \mathbf{A}^\top & \mathbf{0} \\ \mathbf{I} & \mathbf{A} \end{pmatrix}$ is that it is sparse, provided that the original matrix $\mathbf{A}$ is also sparse.

6. **a)** Take

$$\mathbf{A} = \begin{pmatrix} x_1 & x_2 \\ x_2 & x_1 + x_3 \\ x_3 & x_2 + x_4 \\ \vdots & \vdots \\ x_{m-1} & x_{m-2} + x_m \\ x_m & x_{m-1} \end{pmatrix}, \ \mathbf{x} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \ \text{ and } \ \mathbf{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

Therefore, our overdetermined system is $\mathbf{A}\mathbf{x} = \mathbf{b}$.

b) Denote

$$\mathbf{C} = \mathbf{A}^\top\mathbf{A} = \begin{pmatrix} x_1 & x_2 & \dots & x_m \\ x_2 & x_1 + x_3 & \dots & x_{m-1} \end{pmatrix} \begin{pmatrix} x_1 & x_2 \\ x_2 & x_1 + x_3 \\ x_3 & x_2 + x_4 \\ \vdots & \vdots \\ x_m & x_{m-1} \end{pmatrix}.$$

Then we have $c_{11} = \sum_{i=1}^m x_i^2$. Also,

$$\begin{aligned} c_{12} = c_{21} &= x_1 x_2 + \sum_{i=2}^{m-1} x_i(x_{i-1} + x_{i+1}) + x_{m-1}x_m = x_1x_2 + \sum_{i=2}^{m-1} x_i x_{i-1} + \sum_{i=2}^{m-1} x_i x_{i+1} + x_m x_{m-1} \\ &= \sum_{i=2}^{m} x_i x_{i-1} + \sum_{i=1}^{m-1} x_i x_{i+1} = 2\sum_{i=1}^{m-1} x_i x_{i+1}. \end{aligned}$$

For the least entry of $\mathbf{A}^\top\mathbf{A}$ we have

$$c_{22} = x_2^2 + \sum_{i=2}^{m-1} (x_{i-1} + x_{i+1})^2 + x_m^2 = 2\sum_{i=2}^{m-2} x_i^2 + 2\sum_{i=2}^{m-1} x_{i-1}x_{i+1} + x_1^2 + x_m^2.$$

Let us now compute $\mathbf{A}^\top \mathbf{b}$. We have

$$\mathbf{A}^\top \mathbf{b} = \begin{pmatrix} \sum_{i=1}^{m} x_i y_i \\ x_2 y_1 + x_{m-1} y_m + \sum_{i=2}^{m-1} y_i (x_{i-1} + x_{i+1}) \end{pmatrix}$$

c)

```
% Problem 6c)
% INPUT
% x, y - signals, of equal length
% OUTPUT
% beta, alpha - parameters which give an appropriate least squares solution

function [beta, alpha] = CrosstalkChannel(x, y)

% Here we are addapting the linearregression.m code
n = length(y); if (length(x) ~= n), error('data size mismatch'); end

%B Build the matrix of the overdetermined system
A = [x, [x(2); x(1:n-2)+x(3:n);x(n-1)] ];
% Compute the solution
solution = A\y;
% Assign appropriate values
alpha = solution(1);
beta = solution(2);

end
```