Norwegian University of Science and Technology
Department of Mathematics

Solutions to exercise set 10

1 **a)** We follow the algorithm to find z_1 and z_2. All the $h_i = 0.1$ and consequently all $u_i = 0.4$. From $b_i = (y_{i+1} - y_i)/h_i$ and $v_i = 6(b_i - b_{i-1})$ we easily compute $\mathbf{b} = [b_0, b_1, b_2]^T = [3.36512902.90587632.4182345]^T$ and $\mathbf{v} = [v_1, v_2]^T = [-2.7555162, -2.9258508]^T$. This gives the tridiagonal linear system

$$\begin{bmatrix} 0.4 & 0.1 \\ 0.1 & 0.4 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} -2.7555162 \\ -2.9258508 \end{bmatrix},$$

which is easily solved for $z_1 = -5.397476$ and $z_2 = -5.965258$. Insertion of this and the other known quantitites into the cubic equation formula for $S_i(x)$ gives.

$$\begin{aligned}
S_0(x) &= \frac{-5.397476}{6 \cdot 0.1}(x-0.1)^3 + \left(\frac{-0.28398668}{0.1} - \frac{0.1}{6}(-5.397476)\right)(x-0.1) \\
&\quad + \left(\frac{-0.62049958}{0.1}\right)(0.2-x) \\
S_1(x) &= \frac{-5.965258}{6 \cdot 0.1}(x-0.2)^3 + \frac{-5.397476}{6 \cdot 0.1}(0.3-x)^3 \\
&\quad + \left(\frac{0.00660095}{0.1} - \frac{0.1}{6}(-5.965258)\right)(x-0.2) \\
&\quad + \left(\frac{-0.28398668}{0.1} - \frac{0.1}{6}(-5.397476)\right)(0.3-x) \\
S_2(x) &= \frac{-5.965258}{6 \cdot 0.1}(0.4-x)^3 + \left(\frac{0.24842440}{0.1}\right)(x-0.3) \\
&\quad + \left(\frac{0.00660095}{0.1} - \frac{0.1}{6}(-5.965258)\right)(0.4-x)
\end{aligned}$$

or when simplified

$$S(x) = \begin{cases} S_0(x) = -8.9957933x^3 + 2.6987380x^2 + 3.1852131x - 0.95701248, & x \in [0.1, 0.2], \\ S_1(x) = -0.94630333x^3 - 2.1309560x^2 + 4.1511519x - 1.0214084, & x \in [0.2, 0.3], \\ S_2(x) = 9.9420966x^3 - 11.930516x^2 + 7.0910199x - 1.3153952, & x \in [0.3, 0.4]. \end{cases}$$

b) We begin with the function itself.

$$\begin{aligned}
S(0.25) = S_1(0.25) &= -0.94630333 \cdot 0.25^3 - 2.1309560 \cdot 0.25^2 + 4.1511519 \cdot 0.25 \\
&\quad - 1.0214084 = -0.13159116, \\
f(0.25) &= 0.25 \cos 0.25 - 2 \cdot 0.25^2 + 3 \cdot 0.25 - 1 = -0.13277189, \\
|S(0.25 &- f(0.25)| = 1.1807 \times 10^{-3}.
\end{aligned}$$

It is trivial to compute $f'(x) = \cos x - x \sin x - 4x + 3$, and to differentiate the spline function $S_1(x)$, so for the derivative

$$\begin{aligned} S'(0.25) &= S_1'(0.25) = -3 \cdot 0.94630333 \cdot 0.25^2 - 2 \cdot 2.1309560 \cdot 0.25 + 4.1511519 \\ &= 2.9082421 \\ f'(0.25) &= \cos 0.25 - 0.25 \sin 0.25 - 4 \cdot 0.25 + 3 = 2.9070614, \\ &|S'(0.25 - f'(0.25)| = 1.1806 \times 10^{-3}. \end{aligned}$$

We see that the accuracy of the approximation is very similar in both cases. Note that this is not generally the case.

$\boxed{2}$ Since S is a spline of degree k, it is $k-1$-times continuously differentiable, and therefore S' is $k-2$-times continuously differentiable (continuous in case $k = 2$). Furthermore, there exists a partition $a = x_0 < x_1 < \ldots < x_n = b$ such that the restriction of S to each interval (x_{j-1}, x_j), $j = 1, \ldots, n$, is a polynomial of degree k. Hence the restriction of S' to (x_{j-1}, x_j), $j = 1, \ldots, n$, is a polynomial of degree $k-1$, and therefore S' is a spline of degree $k-1$.

$\boxed{3}$ The optimal linear $f(x) = ax + b$ solves the normal equations

$$\begin{aligned} a \sum_k x_k^2 + b \sum_k x_k &= \sum_k x_k y_k, \\ a \sum_k x_k + nb &= \sum_k y_k, \end{aligned}$$

where n is the number of data points (x_k, y_k). In this case we obtain the equations

$$\begin{aligned} 72a + 16b &= 70, \\ 16a + 6b &= 16, \end{aligned}$$

with the solution

$$a = \frac{41}{44} \qquad \text{and} \qquad b = \frac{2}{11},$$

that is,

$$f(x) = \frac{41}{44} x + \frac{2}{11}.$$

$\boxed{4}$ **a)** The least squares problem we want to solve in this situation is the optimisation problem

$$\sum_i (a x_i^2 + b x_i + c - y_i)^2 \to \min.$$

The normal equations can now be obtained by differentiating the left hand side function with respect to a, b, and c, and setting the different derivatives to 0. Thus we obtain

$$\begin{aligned} \partial_a : &\qquad 2 \sum_i x_i^2 (a x_i^2 + b x_i + c - y_i) = 0, \\ \partial_b : &\qquad 2 \sum_i x_i (a x_i^2 + b x_i + c - y_i) = 0, \\ \partial_c : &\qquad 2 \sum_i (a x_i^2 + b x_i + c - y_i) = 0, \end{aligned}$$

which can be rewritten as

$$\begin{aligned} a\sum_i x_i^4 + b\sum_i x_i^3 + c\sum_i x_i^2 &= \sum_i x_i^2 y_i, \\ a\sum_i x_i^3 + b\sum_i x_i^2 + c\sum_i x_i &= \sum_i x_i y_i, \\ a\sum_i x_i^2 + b\sum_i x_i + cn &= \sum_i y_i. \end{aligned}$$

A different way for obtaining the same equations is to rewrite the model as an approximate linear equation of the form $At \approx y$,

$$\begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \approx \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Then the normal equations are $A^T At = A^T$, which in this situation read as

$$\begin{pmatrix} x_1^2 & x_2^2 & \dots & x_n^2 \\ x_1 & x_2 & \dots & x_n \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} x_1^2 & x_2^2 & \dots & x_n^2 \\ x_1 & x_2 & \dots & x_n \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Moreover, this simplifies to

$$\begin{pmatrix} \sum_i x_i^4 & \sum_i x_i^3 & \sum_i x_i^2 \\ \sum_i x_i^3 & \sum_i x_i^2 & \sum_i x_i \\ \sum_i x_i^2 & \sum_i x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_i x_i^2 y_i \\ \sum_i x_i y_i \\ \sum_i y_i \end{pmatrix},$$

which is (not too surprisingly) the same system as we have obtain by differentiation.

b) For the particular given data we obtain the system

$$\begin{pmatrix} 114 & 26 & 18 \\ 26 & 18 & 2 \\ 18 & 2 & 6 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 153 \\ 37 \\ 24 \end{pmatrix}$$

with the solution

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \frac{103}{80} \\ \frac{3}{16} \\ \frac{3}{40} \end{pmatrix}.$$

That is,

$$f(x) = \frac{103}{80}x^2 + \frac{3}{16}x + \frac{3}{40}.$$

$\boxed{5}$ There are (at least) two different ways for approaching this problem:

- The solution of the least squares problem is the solution of the equation $A^T Ax = A^T b$. Using the fact that $A = QR$ and therefore $A^T = R^T Q^T$, we can write this as
$$R^T Q^T QRx = R^T Q^T b.$$
Since Q is orthogonal, we have $Q^T Q = \mathrm{Id}$. Moreover, by definition $Q^T b = y$. Thus the normal equation is equivalent to
$$R^T Rx = R^T y.$$
Using the specific form of the matrix R and the vector y, we rewrite this as
$$\begin{pmatrix} \hat{R}^T & 0 \end{pmatrix} \begin{pmatrix} \hat{R} \\ 0 \end{pmatrix} x = \begin{pmatrix} \hat{R}^T & 0 \end{pmatrix} \begin{pmatrix} \hat{y} \\ z \end{pmatrix},$$
which simplifies to
$$\hat{R}^T \hat{R}x = \hat{R}^T \hat{y}. \tag{1}$$
Now note that the matrix A was assumed to have full rank, which implies that also R has full rank (as Q is an invertible matrix). This, however, means that the square matrix $\hat{R}$ has full rank and is therefore invertible. Thus we can multiply both sides of the equation (1) with $(\hat{R}^T)^{-1}$, and obtain
$$\hat{R}x = \hat{y},$$
which was to show.
- Another approach is to use the fact that the Euclidean norm is invariant under orthogonal transormations, and therefore
$$\|Ax - b\|_2^2 = \|QRx - b\|_2^2 = \|Q(Rx - Q^{-1}b)\|_2^2 = \|Rx - Q^{-1}b\|_2^2.$$
Because Q is orthogonal, we moreover have
$$Q^{-1}b = Q^T b = y.$$
Next we use the particular form of R and y and obtain
$$\|Rx - y\|_2^2 = \left\| \begin{pmatrix} \hat{R} \\ 0 \end{pmatrix} x - \begin{pmatrix} \hat{y} \\ z \end{pmatrix} \right\|_2^2 = \left\| \begin{pmatrix} \hat{R}x - \hat{y} \\ -z \end{pmatrix} \right\|_2^2.$$
Summarising all the above equations and using Pythagoras' theorem, we see that
$$\|Ax - b\|_2^2 = \|\hat{R}x - \hat{y}\|_2^2 + \|z\|_2^2.$$
Since the second term on the right hand side does not depend on x, the minimisation of $\|Ax - b\|_2^2$ is therefore equivalent to the minimisation of $\|\hat{R}x - \hat{y}\|_2^2$. Because $\hat{R}$ is invertible (as A has full rank), the minimum of this latter functional is attained if (and only if)
$$\hat{R}x = \hat{y},$$
which, again, proves that the proposed procedure yields the correct result.