Exercise 11.1

We have

$$\begin{aligned} A A^{+} A &= (USV^T)(VS^{-1}U^T)(USV^T) \\ &= US(V^TV)S^{-1}(U^TU)SV^T \\ &= US\,I\,S^{-1}\,I\,S\,V^T \\ &= U(SS^{-1})SV^T \\ &= U\,I\,SV^T \\ &= USV^T \\ &= A. \end{aligned}$$

Also,

$$\begin{aligned} A^{+} A A^{+} &= (VS^{-1}U^T)(USV^T)(VS^{-1}U^T) \\ &= VS^{-1}(U^TU)S(V^TV)S^{-1}U^T \\ &= VS^{-1}\,I\,S\,I\,S^{-1}U^T \\ &= V(S^{-1}S)S^{-1}U^T \\ &= V\,I\,S^{-1}U^T \\ &= VS^{-1}U^T \\ &= A^{+}. \end{aligned}$$

Exercise 11.2

(1) A vector x is a solution to the least squares problem (P1) iff it satisfies the normal equations

$$A^T A x = A^T b$$

that is,

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} x = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

So

$$\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} x = \begin{bmatrix} b_1 - b_2 \\ b_2 - b_1 \end{bmatrix}.$$

Hence

$$x = \begin{bmatrix} \frac{b_1-b_2}{2} \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}$$

(2) $X(b) = \left\{ \begin{bmatrix} \frac{b_1-b_2}{2} + t \\ t \end{bmatrix} : t \in \mathbb{R} \right\}.$

Hence $\|x\|^2 = \left(\frac{b_1-b_2}{2} + t\right)^2 + t^2$

$$= 2t^2 + (b_1 - b_2)t + \left(\frac{b_1-b_2}{2}\right)^2,$$

which is minimized when $t = \dfrac{b_2 - b_1}{4}$.

Thus $\hat{x}(b) = \begin{bmatrix} \frac{b_1-b_2}{4} \\ \frac{b_2-b_1}{4} \end{bmatrix}$ is the optimal solution to (P2).

(Alternately, following the procedure described in this chapter, we could solve the equation for u

$$AA^T u = A\bar{x}$$

where $\bar{x}$ is any solution to (P1), and then calculate $\hat{x} = A^T u$. We have

$$\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} u = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} \frac{b_1-b_2}{2} \\ 0 \end{bmatrix}}_{(t=0)} = \begin{bmatrix} \frac{b_1-b_2}{2} \\ \frac{b_2-b_1}{2} \end{bmatrix} \quad (*)$$

So $u = \begin{bmatrix} \frac{b_1-b_2}{4} \\ 0 \end{bmatrix}$ is a solution to $(*)$,

and hence $\hat{x} = A^T u = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{b_1-b_2}{4} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{b_1-b_2}{4} \\ \frac{b_2-b_1}{4} \end{bmatrix}.$)

(3) $\hat{x}(b) = \begin{bmatrix} \frac{b_1 - b_2}{4} \\ \frac{b_2 - b_1}{4} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = A^+ b$, where

$A^+ := \begin{bmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{bmatrix}$.

(4)
$$\begin{aligned} \|Ax - b\|^2 + \epsilon \|x\|^2 &= (Ax-b)^T (Ax-b) + \epsilon x^T x \\ &= x^T A^T A x - 2 b^T A x + b^T b + \epsilon x^T x \\ &= x^T (A^T A + \epsilon I) x - 2 b^T A x + b^T b \\ &= \frac{1}{2} x^T H x + c^T x + c_0, \end{aligned}$$

where $H := 2(A^T A + \epsilon I)$, $c = -2A^T b$ and $c_0 = b^T b$.
So $\tilde{x}$ is an optimal solution to (P3) iff
$$H\tilde{x} = -c$$
that is,
$$2(A^T A + \epsilon I)\tilde{x} = 2A^T b.$$
Hence the unique solution is
$$\tilde{x} = (A^T A + \epsilon I)^{-1} A^T b$$
($A^T A + \epsilon I$ is invertible since it is square and its kernel is 0 : if $(A^T A + \epsilon I)x = 0$, then $x^T(A^T A + \epsilon I)x = 0$, and so $\underbrace{(Ax)^T Ax}_{\geq 0} + \underset{>0}{\epsilon} \underbrace{x^T x}_{\geq 0} = 0$.
Thus $x = 0$.)

In our case,
$$A^T A + \epsilon I = \begin{bmatrix} 2+\epsilon & -2 \\ -2 & 2+\epsilon \end{bmatrix} \quad \text{and} \quad A^T b = \begin{bmatrix} b_1 - b_2 \\ b_2 - b_1 \end{bmatrix}$$

Consequently
$$\begin{aligned} \tilde{x} &= \frac{1}{(2+\epsilon)^2 - 2^2} \begin{bmatrix} 2+\epsilon & 2 \\ 2 & 2+\epsilon \end{bmatrix} \begin{bmatrix} b_1 - b_2 \\ b_2 - b_1 \end{bmatrix} \\ &= \frac{1}{\epsilon(4+\epsilon)} \begin{bmatrix} 2+\epsilon & 2 \\ 2 & 2+\epsilon \end{bmatrix} \begin{bmatrix} b_1 - b_2 \\ b_2 - b_1 \end{bmatrix} \\ &= \frac{1}{\epsilon(4+\epsilon)} \begin{bmatrix} \epsilon(b_1 - b_2) \\ \epsilon(b_2 - b_1) \end{bmatrix} \\ &= \begin{bmatrix} \frac{b_1 - b_2}{4+\epsilon} \\ \frac{b_2 - b_1}{4+\epsilon} \end{bmatrix}. \end{aligned}$$

(5) From the above, it follows that

$$\tilde{x}(b, \epsilon) = \begin{bmatrix} \frac{b_1 - b_2}{4+\epsilon} \\ \frac{b_2 - b_1}{4+\epsilon} \end{bmatrix} = \begin{bmatrix} \frac{1}{4+\epsilon} & \frac{-1}{4+\epsilon} \\ -\frac{1}{4+\epsilon} & \frac{1}{4+\epsilon} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \tilde{A}_\epsilon \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

where

$$\tilde{A}_\epsilon = \begin{bmatrix} \frac{1}{4+\epsilon} & -\frac{1}{4+\epsilon} \\ -\frac{1}{4+\epsilon} & \frac{1}{4+\epsilon} \end{bmatrix}.$$

As $\epsilon \to 0$, $\frac{1}{4+\epsilon} \longrightarrow \frac{1}{4}$, and so it follows that for all i, j,

$$[\tilde{A}_\epsilon]_{ij} \to [A]_{ij}.$$

Exercise 11.3

(The problem can be formulated as:

$$\begin{cases} \text{minimize} & \frac{1}{2}\|u-v\|^2 = \frac{1}{2}(u-v)^T(u-v) \\ \text{subject to} & Ru = p, \\ & Sv = q. \end{cases}$$

The Lagrange method gives the system

$$\begin{bmatrix} H & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix}$$

where

$$H = \begin{bmatrix} I_4 & -I_4 \\ -I_4 & I_4 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} R & 0 \\ 0 & S \end{bmatrix}.$$

So the above is a system of $(4+4+3+3=)14$ equations in $(4+4+3+3=)14$ unknowns, and so a computation by hand is not feasible. So we try a different method.)

We note that $Ru = p$ iff $u = u_0 + g x_1$, where

$$u_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \quad g = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}$$

and x_1 is an arbitrary real number.

Similarly $Sv = q$ iff $v = v_0 + h x_2$, where

$$v_0 = \begin{bmatrix} 0 \\ 2 \\ 2 \\ 2 \end{bmatrix}, \quad h = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

and x_2 is an arbitrary real number.

(Note that $\ker R = \operatorname{span} g$ and $\ker S = \operatorname{span} h$.)

So the original problem can be rephrased as follows:

$$\text{minimize} \quad \frac{1}{2}\| u_0 - v_0 + g x_1 - h x_2 \|^2$$

which is a least squares problem with

$$A := \begin{bmatrix} g & -h \end{bmatrix}, \quad \text{and}$$

$$b := v_0 - u_0.$$

The normal equations are:

$$A^TA\,x = A^Tb$$

and so

$$\begin{bmatrix} g^T \\ -h^T \end{bmatrix} \begin{bmatrix} g & -h \end{bmatrix} x = \begin{bmatrix} g^T \\ -h^T \end{bmatrix} (u_0 - v_0),$$

that is

$$\begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ -1 & -1 \end{bmatrix} x = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & -1 \end{bmatrix} \left(\begin{bmatrix} 0 \\ 2 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right)$$

Thus $$\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} x = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ -5 \end{bmatrix},$$

$-1+1+1-2$
$-1-1-1-2$

and so $$x = \begin{bmatrix} -\frac{1}{4} \\ -\frac{5}{4} \end{bmatrix}$$

Consequently

$$\hat{u} = u_0 + g x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} \\ \frac{3}{4} \\ \frac{3}{4} \\ \frac{1}{4} \end{bmatrix}, \quad \text{and}$$

$$\hat{v} = v_0 + h x_2 = \begin{bmatrix} 0 \\ 2 \\ 2 \\ 2 \end{bmatrix} - \frac{5}{4} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{4} \\ \frac{3}{4} \\ \frac{3}{4} \\ \frac{3}{4} \end{bmatrix}.$$

The distance $d(U,V)$ between the two sets (lines) U and V is $d(U,V) = \|\hat{u} - \hat{v}\| = \sqrt{\left(\frac{1}{2}\right)^2 + 0^2 + 0^2 + \left(\frac{1}{2}\right)^2}$

$$= \frac{1}{\sqrt{2}}.$$

Exercise 11.4

We seek a solution to the normal equations

$$A^T A x = A^T b \quad (*).$$

A has linearly independent columns, and so $A^T A$ is invertible and $(*)$ has a unique solution $\hat{x}$.

We have

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

$$= \begin{bmatrix} 6 & -3 \\ -3 & 6 \end{bmatrix} = 3 \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

So $$\hat{x} = \frac{1}{3} \cdot \frac{1}{3} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}$$

$$= \frac{1}{9} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 7 \\ 4 \end{bmatrix}$$

$$= \frac{1}{9} \begin{bmatrix} 18 \\ 15 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 5/3 \end{bmatrix}.$$

Exercise 11.5

Let h_1, h_2, h_3 be the true heights of the hills above sea level. The errors $e_1, \dots, e_6$ in the six measurements are then given as follows:

$$e_1 = h_1 - 1236$$
$$e_2 = h_2 - 1941$$
$$e_3 = h_2 - h_1 - 711$$
$$e_5 = h_3 - h_1 - 1177$$
$$e_6 = h_3 - h_2 - 474.$$

The problem is to minimize $e_1^2 + e_2^2 + e_3^2 + e_4^2 + e_5^2 + e_6^2$, that is

$$\begin{cases} \text{minimize} & \| Ax - b \|^2 \\ \text{subject to} & x \in \mathbb{R}^3, \end{cases}$$

where

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1236 \\ 1941 \\ 2417 \\ 711 \\ 1177 \\ 474 \end{bmatrix} \quad \text{and} \quad x = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}.$$

$\hat{x}$ is optimal iff it satisfies the normal equation

$$A^T A \, \hat{x} = A^T b.$$

We have

$$A^T A = \begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix}$$

and $$A^T b = \begin{bmatrix} -652 \\ 2178 \\ 4068 \end{bmatrix}.$$

Thus adding all equations in the system $A^TA\hat{x} = A^Tb$ gives us that

$$\hat{x}_1 + \hat{x}_2 + \hat{x}_3 = 4068 + 2178 - 652 = 5594.$$

Now adding the equation

$$\hat{x}_1 + \hat{x}_2 + \hat{x}_3 = 5594$$

to each of the equations in the system $A^TA\hat{x} = A^Tb$ yields

$$\hat{x}_1 = \frac{-652 + 5594}{4} = 1235.5,$$

$$\hat{x}_2 = \frac{2178 + 5594}{4} = 1943,$$

$$\hat{x}_3 = \frac{4068 + 5594}{4} = 2415.5.$$

So upon minimizing the least squares error associated with the measurements, the estimated heights of the hills H_1, H_2, H_3 are 1235.5, 1943, 2415.5 m, respectively.