Machine Learning with Python

MTH786U/P 2023/24

Detailed solutions Coursework 1

Nicola Perra, Queen Mary University of London (QMUL)

n.perra@qmul.ac.uk

Problem 1

Problem 1. Let A be a 2 matrix

$$A = \begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix}.$$

(i) Find eigenvalues, eigenvectors and eigenvalue decomposition of matrix A.

(ii) Let $\vec{x}$ be a two-dimensional column-vector. Write the product $A\vec{x}$ in terms of eigenvectors of matrix A.

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

(iv) Let $\vec{x}$ be a two-dimensional column-vector. Write the product $A\vec{x}$ in terms of singular vectors of matrix A.

(i) Find eigenvalues, eigenvectors and eigenvalue decomposition of matrix A.

$$A = \begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix}$$

(i) Find eigenvalues, eigenvectors and eigenvalue decomposition of matrix A.

$$A = \begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix}$$

The eigenvalues of matrix A can be found by solving $\det(A - \lambda I) = 0$. In our case one has

$$\det(A - \lambda I) = \det \begin{pmatrix} 3 - \lambda & 4 \\ 0 & 5 - \lambda \end{pmatrix} = (3 - \lambda)(5 - \lambda) = 0 \Rightarrow \lambda_{1,2} = 3, 5.$$

(i) Find eigenvalues, eigenvectors and eigenvalue decomposition of matrix A.

Corresponding eigenvectors can be found by solving $Au = \lambda u$ for the values of λ found above.

$$Au^{(1)} = 3u^{(1)} \Rightarrow \begin{cases} 3u_1^{(1)} + & 4u_2^{(1)} = 3u_1^{(1)} \\ & 5u_2^{(1)} = 3u_2^{(1)} \end{cases} \Rightarrow u^{(1)} = (1,0)^T .$$

$$Au^{(2)} = 5u^{(2)} \Rightarrow \begin{cases} 3u_1^{(2)} + & 4u_2^{(2)} = 5u_1^{(2)} \\ & 5u_2^{(2)} = 5u_2^{(2)} \end{cases} \Rightarrow u^{(2)} = (2,1)^T .$$

$$\begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} u_1^{(1)} \\ u_1^{(2)} \end{pmatrix} = \lambda_1 \begin{pmatrix} u_1^{(1)} \\ u_1^{(2)} \end{pmatrix} = \begin{pmatrix} 3u_1^{(1)} \\ 3u_1^{(2)} \end{pmatrix}$$

(i) Find eigenvalues, eigenvectors and eigenvalue decomposition of matrix A.

One can build an eigenvalue decomposition of a matrix A by writing

$$A = Q\Lambda Q^{-1},$$

where $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix whose diagonal elements are just eigenvalues of matrix A and Q is the matrix whose i-th column is an eigenvector corresponding to λ_i. In our case one has

$$\Lambda = \begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad Q^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}.$$

Therefore, the eigenvalue decomposition of matrix A has a form

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}.$$

(ii) Let $\vec{x}$ be a two-dimensional column-vector. Write the product $A\vec{x}$ in terms of eigenvectors of matrix A.

(ii) Let $\vec{x}$ be a two-dimensional column-vector. Write the product $A\vec{x}$ in terms of eigenvectors of matrix A.

Let v be an arbitrary vector. Then, because eigenvalues of matrix A are linearly independent, one can find α and β such that

$$v = \alpha u^{(1)} + \beta u^{(2)}. \tag{1}$$

(ii) Let $\vec{x}$ be a two-dimensional column-vector. Write the product $A\vec{x}$ in terms of eigenvectors of matrix A.

Let v be an arbitrary vector. Then, because eigenvalues of matrix A are linearly independent, one can find α and β such that

$$v = \alpha u^{(1)} + \beta u^{(2)}. \tag{1}$$

Matrix A when applied to vector v produces

$$Av = \alpha\lambda_1 u^{(1)} + \beta\lambda_2 u^{(2)}.$$

(ii) Let $\vec{x}$ be a two-dimensional column-vector. Write the product $A\vec{x}$ in terms of eigenvectors of matrix A.

Let v be an arbitrary vector. Then, because eigenvalues of matrix A are linearly independent, one can find α and β such that

$$v = \alpha u^{(1)} + \beta u^{(2)}. \tag{1}$$

Matrix A when applied to vector v produces

$$Av = \alpha\lambda_1 u^{(1)} + \beta\lambda_2 u^{(2)}.$$

$$Av = A(\alpha u^{(1)} + \beta u^{(2)}) = \alpha\lambda_1 u^{(1)} + \beta\lambda_2 u^{(2)}$$

(ii) Let $\vec{x}$ be a two-dimensional column-vector. Write the product $A\vec{x}$ in terms of eigenvectors of matrix A.

$$(u^{(1)})^\top v = (u^{(1)})^\top(\alpha u^{(1)} + \beta u^{(2)}) = \alpha\|u^{(1)}\|^2 + \beta\langle u^{(1)}, u^{(2)}\rangle = \langle u^{(1)}, v\rangle$$

(ii) Let $\vec{x}$ be a two-dimensional column-vector. Write the product $A\vec{x}$ in terms of eigenvectors of matrix A.

$$(u^{(1)})^\top v = (u^{(1)})^\top(\alpha u^{(1)} + \beta u^{(2)}) = \alpha\|u^{(1)}\|^2 + \beta\langle u^{(1)}, u^{(2)}\rangle = \langle u^{(1)}, v\rangle$$

$$(u^{(2)})^\top v = (u^{(2)})^\top v(\alpha u^{(1)} + \beta u^{(2)}) = \beta\|u^{(2)}\|^2 + \alpha\langle u^{(2)}, u^{(1)}\rangle = \langle u^{(2)}, v\rangle$$

(ii) Let $\vec{x}$ be a two-dimensional column-vector. Write the product $A\vec{x}$ in terms of eigenvectors of matrix A.

$$(u^{(1)})^{\top} v = (u^{(1)})^{\top}(\alpha u^{(1)} + \beta u^{(2)}) = \alpha \|u^{(1)}\|^2 + \beta \langle u^{(1)}, u^{(2)} \rangle = \langle u^{(1)}, v \rangle$$

$$(u^{(2)})^{\top} v = (u^{(2)})^{\top} v(\alpha u^{(1)} + \beta u^{(2)}) = \beta \|u^{(2)}\|^2 + \alpha \langle u^{(2)}, u^{(1)} \rangle = \langle u^{(2)}, v \rangle$$

$$\begin{pmatrix} \|u^{(1)}\|^2 & \langle u^{(1)}, u^{(2)} \rangle \\ \langle u^{(2)}, u^{(1)} \rangle & \|u^{(2)}\|^2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \langle u^{(1)}, v \rangle \\ \langle u^{(2)}, v \rangle \end{pmatrix}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

The singular values of matrix A are defined as positive square roots of eigenvalues of $A^T A$. Let

$$B := A^T A = \begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix}.$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

The singular values of matrix A are defined as positive square roots of eigenvalues of $A^T A$. Let

$$B := A^T A = \begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix}.$$

Then corresponding singular values could be found by solving

$$\det\left(B - \sigma^2 I\right) = \det\begin{pmatrix} 9-\sigma^2 & 12 \\ 12 & 41-\sigma^2 \end{pmatrix} = \left(\sigma^2 - 41\right)\left(\sigma^2 - 9\right) - 144 = 0.$$

Solutions are then given by $\sigma = \sqrt{5}, 3\sqrt{5}$.

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^\top A v^{(1)} = \sigma_1^2 v^{(1)}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^\top A v^{(1)} = \sigma_1^2 v^{(1)}$$

$$\begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix} = 5 \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^{\top}Av^{(1)} = \sigma_1^2 v^{(1)}$$

$$\begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix} = 5 \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix}$$

$$9v_1^{(1)} + 12v_1^{(2)} = 5v_1^{(1)} \rightarrow$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^\top A v^{(1)} = \sigma_1^2 v^{(1)}$$

$$\begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix} = 5 \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix}$$

$$9v_1^{(1)} + 12v_1^{(2)} = 5v_1^{(1)} \rightarrow \qquad 4v_1^{(1)} + 12v_1^{(2)} = 0 \rightarrow$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^{\top} A v^{(1)} = \sigma_1^2 v^{(1)}$$

$$\begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix} = 5 \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix}$$

$$9v_1^{(1)} + 12v_1^{(2)} = 5v_1^{(1)} \rightarrow \quad 4v_1^{(1)} + 12v_1^{(2)} = 0 \rightarrow \quad v_1^{(1)} = -3v_1^{(2)}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^\top A v^{(1)} = \sigma_1^2 v^{(1)}$$

$$\begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix} = 5 \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix}$$

$$9v_1^{(1)} + 12v_1^{(2)} = 5v_1^{(1)} \rightarrow \quad 4v_1^{(1)} + 12v_1^{(2)} = 0 \rightarrow \quad v_1^{(1)} = -3v_1^{(2)}$$

$$12v_1^{(1)} + 41v_1^{(2)} = 5v_1^{(2)} \rightarrow$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^\top A v^{(1)} = \sigma_1^2 v^{(1)}$$

$$\begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix} = 5 \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix}$$

$$9v_1^{(1)} + 12v_1^{(2)} = 5v_1^{(1)} \rightarrow \qquad 4v_1^{(1)} + 12v_1^{(2)} = 0 \rightarrow \qquad v_1^{(1)} = -3v_1^{(2)}$$

$$12v_1^{(1)} + 41v_1^{(2)} = 5v_1^{(2)} \rightarrow \qquad 12v_1^{(1)} + 36v_1^{(2)} = 0 \rightarrow$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^\top A v^{(1)} = \sigma_1^2 v^{(1)}$$

$$\begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix} = 5 \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix}$$

$$9v_1^{(1)} + 12v_1^{(2)} = 5v_1^{(1)} \rightarrow \quad 4v_1^{(1)} + 12v_1^{(2)} = 0 \rightarrow \quad v_1^{(1)} = -3v_1^{(2)}$$

$$12v_1^{(1)} + 41v_1^{(2)} = 5v_1^{(2)} \rightarrow \quad 12v_1^{(1)} + 36v_1^{(2)} = 0 \rightarrow \quad v_1^{(1)} = -3v_1^{(2)}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^\top A v^{(1)} = \sigma_1^2 v^{(1)}$$

$$\begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix} = 5 \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix}$$

$$9v_1^{(1)} + 12v_1^{(2)} = 5v_1^{(1)} \rightarrow \quad 4v_1^{(1)} + 12v_1^{(2)} = 0 \rightarrow \quad v_1^{(1)} = -3v_1^{(2)}$$

$$12v_1^{(1)} + 41v_1^{(2)} = 5v_1^{(2)} \rightarrow \quad 12v_1^{(1)} + 36v_1^{(2)} = 0 \rightarrow \quad v_1^{(1)} = -3v_1^{(2)}$$

$$v_1 = (-3c, c)^\top$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^{\top}Av^{(1)} = \sigma_1^2 v^{(1)}$$

$$\begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix} = 5 \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix}$$

$$9v_1^{(1)} + 12v_1^{(2)} = 5v_1^{(1)} \rightarrow \qquad 4v_1^{(1)} + 12v_1^{(2)} = 0 \rightarrow \qquad v_1^{(1)} = -3v_1^{(2)}$$

$$12v_1^{(1)} + 41v_1^{(2)} = 5v_1^{(2)} \rightarrow \qquad 12v_1^{(1)} + 36v_1^{(2)} = 0 \rightarrow \qquad v_1^{(1)} = -3v_1^{(2)}$$

$$v_1 = (-3c, c)^{\top}$$

$$v_1 = (-3\frac{\sqrt{10}}{10}, \frac{\sqrt{10}}{10})^{\top}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^\top A v^{(2)} = \sigma_2^2 v^{(2)}$$

$$\begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} \begin{pmatrix} v_2^{(1)} \\ v_2^{(2)} \end{pmatrix} = 45 \begin{pmatrix} v_2^{(1)} \\ v_2^{(2)} \end{pmatrix}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^\top A v^{(2)} = \sigma_2^2 v^{(2)}$$

$$\begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} \begin{pmatrix} v_2^{(1)} \\ v_2^{(2)} \end{pmatrix} = 45 \begin{pmatrix} v_2^{(1)} \\ v_2^{(2)} \end{pmatrix}$$

$$9v_2^{(1)} + 12v_2^{(2)} = 45v_2^{(1)} \rightarrow -36v_2^{(1)} + 12v_2^{(2)} = 0 \rightarrow v_2^{(1)} = \frac{v_2^{(2)}}{3}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^\top A v^{(2)} = \sigma_2^2 v^{(2)}$$

$$\begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} \begin{pmatrix} v_2^{(1)} \\ v_2^{(2)} \end{pmatrix} = 45 \begin{pmatrix} v_2^{(1)} \\ v_2^{(2)} \end{pmatrix}$$

$$9v_2^{(1)} + 12v_2^{(2)} = 45v_2^{(1)} \rightarrow -36v_2^{(1)} + 12v_2^{(2)} = 0 \rightarrow v_2^{(1)} = \frac{v_2^{(2)}}{3}$$

$$12v_2^{(1)} + 41v_2^{(2)} = 45v_2^{(2)} \rightarrow 12v_2^{(1)} - 4v_2^{(2)} = 0 \rightarrow v_2^{(1)} = \frac{v_2^{(2)}}{3}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^\top A v^{(2)} = \sigma_2^2 v^{(2)}$$

$$\begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} \begin{pmatrix} v_2^{(1)} \\ v_2^{(2)} \end{pmatrix} = 45 \begin{pmatrix} v_2^{(1)} \\ v_2^{(2)} \end{pmatrix}$$

$$9v_2^{(1)} + 12v_2^{(2)} = 45v_2^{(1)} \rightarrow -36v_2^{(1)} + 12v_2^{(2)} = 0 \rightarrow v_2^{(1)} = \frac{v_2^{(2)}}{3}$$

$$12v_2^{(1)} + 41v_2^{(2)} = 45v_2^{(2)} \rightarrow 12v_2^{(1)} - 4v_2^{(2)} = 0 \rightarrow v_2^{(1)} = \frac{v_2^{(2)}}{3}$$

$$v_2 = (c, 3c)^\top$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$A^\top A v^{(2)} = \sigma_2^2 v^{(2)}$$

$$\begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} \begin{pmatrix} v_2^{(1)} \\ v_2^{(2)} \end{pmatrix} = 45 \begin{pmatrix} v_2^{(1)} \\ v_2^{(2)} \end{pmatrix}$$

$$9v_2^{(1)} + 12v_2^{(2)} = 45v_2^{(1)} \rightarrow -36v_2^{(1)} + 12v_2^{(2)} = 0 \rightarrow v_2^{(1)} = \frac{v_2^{(2)}}{3}$$

$$12v_2^{(1)} + 41v_2^{(2)} = 45v_2^{(2)} \rightarrow 12v_2^{(1)} - 4v_2^{(2)} = 0 \rightarrow v_2^{(1)} = \frac{v_2^{(2)}}{3}$$

$$v_2 = (c, 3c)^\top$$

$$v_1 = (\frac{\sqrt{10}}{10}, 3\frac{\sqrt{10}}{10})^\top$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$AA^\top u^{(1)} = \sigma_1^2 u^{(1)}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$AA^\top u^{(1)} = \sigma_1^2 u^{(1)}$$

$$u^{(1)} = \sigma_1^{-1} A v^{(1)}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$AA^\top u^{(1)} = \sigma_1^2 u^{(1)}$$

$$u^{(1)} = \sigma_1^{-1} A v^{(1)}$$

$$u^{(1)} = \frac{1}{\sqrt{5}} \begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} -3\frac{\sqrt{10}}{10} \\ \frac{\sqrt{10}}{10} \end{pmatrix}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$AA^\top u^{(1)} = \sigma_1^2 u^{(1)}$$

$$u^{(1)} = \sigma_1^{-1} A v^{(1)}$$

$$u^{(1)} = \frac{1}{\sqrt{5}} \begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} -3\frac{\sqrt{10}}{10} \\ \frac{\sqrt{10}}{10} \end{pmatrix}$$

$$u^{(1)} = \frac{\sqrt{10}}{10\sqrt{5}} \begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} -3 \\ 1 \end{pmatrix} = \frac{\sqrt{2}}{2} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$AA^\top u^{(2)} = \sigma_2^2 u^{(2)}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$AA^\top u^{(2)} = \sigma_2^2 u^{(2)}$$

$$u^{(2)} = \sigma_2^{-1} A v^{(2)}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$AA^\top u^{(2)} = \sigma_2^2 u^{(2)}$$

$$u^{(2)} = \sigma_2^{-1} A v^{(2)}$$

$$u^{(2)} = \frac{1}{3\sqrt{5}} \begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{10}}{10} \\ 3\frac{\sqrt{10}}{10} \end{pmatrix}$$

(iii) Find singular values, right and left singular vectors and singular value decomposition of matrix A.

$$AA^\top u^{(2)} = \sigma_2^2 u^{(2)}$$

$$u^{(2)} = \sigma_2^{-1} A v^{(2)}$$

$$u^{(2)} = \frac{1}{3\sqrt{5}} \begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{10}}{10} \\ 3\frac{\sqrt{10}}{10} \end{pmatrix}$$

$$u^{(2)} = \frac{\sqrt{10}}{30\sqrt{5}} \begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

(iv) Let $\vec{x}$ be a two-dimensional column-vector. Write the product $A\vec{x}$ in terms of singular vectors of matrix A.

$$Aw = U\Sigma V^T w$$

$$(Aw)^{(1)} = \sigma_1 u^{(1)} \langle (v^{(1)})^T, w \rangle$$

$$(Aw)^{(2)} = \sigma_2 u^{(2)} \langle (v^{(2)})^T, w \rangle$$

Problem 2

Problem 2. Let the function $f : \mathbb{R}^2 \to \mathbb{R}$ be defined as

$$f\left(\mathbf{x} = (x_1, x_2)^\top\right) = \frac{1}{2}\langle \mathbf{x}, \mathbf{A}\mathbf{x}\rangle + \langle \mathbf{u}, \mathbf{x}\rangle,$$

where $\mathbf{A}$ is a real, symmetric, positive definite 2×2 matrix, and $\mathbf{u}$ is a real vector of length 2. In other words

$$f\left(\mathbf{x} = (x_1, x_2)^\top\right) = \frac{1}{2}ax_1^2 + bx_1x_2 + \frac{1}{2}cx_2^2 + vx_1 + wx_2,$$

where

$$\mathbf{A} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \text{ and } \mathbf{u} = (v, w)^\top.$$

Our goal is to find such a vector $\mathbf{x}^*$ that minimises $f(\mathbf{x})$.

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.
2. Thus proof that the gradient is zero at $\mathbf{x}^* = -\mathbf{A}^{-1}\mathbf{u}$.
3. Evaluate $f(\mathbf{x}^*)$.
4. Calculate the Hessian $H_f(\mathbf{x}^*)$ and show it is positive. This will finish the proof of $\mathbf{x}^*$ being a minimizer point of $f(\mathbf{x})$.

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

$$f(x) = \frac{1}{2}\langle x, Ax \rangle + \langle u, x \rangle$$

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

$$f(x) = \frac{1}{2}\langle x, Ax\rangle + \langle u, x\rangle$$

$$f(x) = \frac{1}{2}\sum_i x_i^\top \sum_j A_{ij}x_j + \sum_i u_i^\top x_i$$

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

$$f(x) = \frac{1}{2}\langle x, Ax\rangle + \langle u, x\rangle$$

$$f(x) = \frac{1}{2}\sum_i x_i^\top \sum_j A_{ij}x_j + \sum_i u_i^\top x_i$$

$$\partial_{x_p} f(x) = \frac{1}{2}\left(\partial_{x_p}\sum_i x_i^\top\right)\sum_j A_{ij}x_j + \frac{1}{2}\sum_i x_i^\top\left(\partial_{x_p}\sum_j A_{ij}x_j\right) + \partial_{x_p}\sum_i u_i^\top x_i$$

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

$$\partial_{x_p} f(x) = \frac{1}{2}\left(\partial_{x_p} \sum_i x_i^\top\right) \sum_j A_{ij} x_j + \frac{1}{2} \sum_i x_i^\top \left(\partial_{x_p} \sum_j A_{ij} x_j\right) + \partial_{x_p} \sum_i u_i^\top x_i$$

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

$$\partial_{x_p} f(x) = \frac{1}{2}\left(\partial_{x_p} \sum_{i} x_i^{\top}\right) \sum_{j} A_{ij} x_j + \frac{1}{2} \sum_{i} x_i^{\top} \left(\partial_{x_p} \sum_{j} A_{ij} x_j\right) + \partial_{x_p} \sum_{i} u_i^{\top} x_i$$

i=p

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

$$\partial_{x_p} f(x) = \frac{1}{2}\left(\partial_{x_p} \sum_{i} x_i^\top\right) \sum_{j} A_{ij} x_j + \frac{1}{2} \sum_{i} x_i^\top \left(\partial_{x_p} \sum_{j} A_{ij} x_j\right) + \partial_{x_p} \sum_{i} u_i^\top x_i$$

i=p

j=p

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

$$\partial_{x_p} f(x) = \frac{1}{2}\left(\partial_{x_p} \sum_{i} x_i^\top\right) \sum_{j} A_{ij} x_j + \frac{1}{2} \sum_{i} x_i^\top \left(\partial_{x_p} \sum_{j} A_{ij} x_j\right) + \partial_{x_p} \sum_{i} u_i^\top x_i$$

i=p j=p i=p

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

$$\partial_{x_p} f(x) = \frac{1}{2}\left(\partial_{x_p} \sum_{i} x_i^\top\right) \sum_{j} A_{ij} x_j + \frac{1}{2} \sum_{i} x_i^\top \left(\partial_{x_p} \sum_{j} A_{ij} x_j\right) + \partial_{x_p} \sum_{i} u_i^\top x_i$$

i=p   j=p   i=p

$$\partial_{x_p} f(x) = \frac{1}{2} \sum_{j} A_{pj} x_j + \frac{1}{2} \sum_{i} x_i^\top A_{ip} + u_p$$

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

$$\partial_{x_p} f(x) = \frac{1}{2}\left(\partial_{x_p} \underset{i=p}{\sum_i} x_i^{\top}\right) \sum_j A_{ij} x_j + \frac{1}{2} \sum_i x_i^{\top} \left(\partial_{x_p} \underset{j=p}{\sum_j} A_{ij} x_j\right) + \partial_{x_p} \sum_i \underset{i=p}{u_i^{\top} x_i}$$

$$\partial_{x_p} f(x) = \frac{1}{2} \sum_j A_{pj} x_j + \frac{1}{2} \sum_i x_i^{\top} A_{ip} + u_p$$

$$\partial_{x_p} f(x) = \frac{1}{2} \sum_j A_{pj} x_j + \frac{1}{2} \sum_i A_{pi}^{\top} x_i + u_p$$

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

$$\partial_{x_p} f(x) = \frac{1}{2}\left(\partial_{x_p} \sum_{i} x_i^\top\right) \sum_j A_{ij} x_j + \frac{1}{2} \sum_i x_i^\top \left(\partial_{x_p} \sum_j A_{ij} x_j\right) + \partial_{x_p} \sum_i u_i^\top x_i$$

i=p j=p i=p

$$\partial_{x_p} f(x) = \frac{1}{2} \sum_j A_{pj} x_j + \frac{1}{2} \sum_i x_i^\top A_{ip} + u_p$$

$$\partial_{x_p} f(x) = \frac{1}{2} \sum_j A_{pj} x_j + \frac{1}{2} \sum_i A_{pi}^\top x_i + u_p$$

$$\partial_{x_p} f(x) = \left(\frac{1}{2}(A + A^\top)x + u\right)_p$$

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

$$\partial_{x_p} f(x) = \left(\frac{1}{2}(A + A^\top)x + u\right)_p$$

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

$$\partial_{x_p} f(x) = \left(\frac{1}{2}(A + A^\top)x + u \right)_p$$

$$\nabla f(x) = \frac{1}{2}(A + A^\top)x + u$$

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

$$\partial_{x_p} f(x) = \left(\frac{1}{2}(A + A^\top)x + u \right)_p$$

$$\nabla f(x) = \frac{1}{2}(A + A^\top)x + u$$

where $\mathbf{A}$ is a real, symmetric, positive definite 2×2 matrix,

1. Show that the gradient ∇f is given by $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{u}$.

$$\partial_{x_p} f(x) = \left(\frac{1}{2}(A + A^\top)x + u \right)_p$$

$$\nabla f(x) = \frac{1}{2}(A + A^\top)x + u$$

where $\mathbf{A}$ is a real, symmetric, positive definite 2×2 matrix,

$$\nabla f(x) = Ax + u$$

2. Thus proof that the gradient is zero at $\mathbf{x}^* = -\mathbf{A}^{-1}\mathbf{u}$.

2. Thus proof that the gradient is zero at $\mathbf{x}^* = -\mathbf{A}^{-1}\mathbf{u}$.

$$\nabla f(x) = Ax + u$$

2. Thus proof that the gradient is zero at $\mathbf{x}^* = -\mathbf{A}^{-1}\mathbf{u}$.

$$\nabla f(x) = Ax + u$$

$$\nabla f(x) = 0 = Ax + u = 0$$

2. Thus proof that the gradient is zero at $\mathbf{x}^* = -\mathbf{A}^{-1}\mathbf{u}$.

$$\nabla f(x) = Ax + u$$

$$\nabla f(x) = 0 = Ax + u = 0$$

$$Ax + u = 0$$

2. Thus proof that the gradient is zero at $\mathbf{x}^* = -\mathbf{A}^{-1}\mathbf{u}$.

$$\nabla f(x) = Ax + u$$

$$\nabla f(x) = 0 = Ax + u = 0$$

$$Ax + u = 0$$

$$x^* = -A^{-1}u$$

2. Thus proof that the gradient is zero at $\mathbf{x}^* = -\mathbf{A}^{-1}\mathbf{u}$.

$$\nabla f(x) = Ax + u$$

$$\nabla f(x) = 0 = Ax + u = 0$$

$$Ax + u = 0$$

$$x^* = -A^{-1}u$$

Note how so far we did not use the expression of A!

3. Evaluate $f(\mathbf{x}^*)$.

3. Evaluate $f(\mathbf{x}^*)$.

$$f(x) = \frac{1}{2}\langle x, Ax\rangle + \langle u, x\rangle$$

3. Evaluate $f(\mathbf{x}^*)$.

$$f(x) = \frac{1}{2}\langle x, Ax\rangle + \langle u, x\rangle$$

$$x^* = -A^{-1}u$$

3. Evaluate $f(\mathbf{x}^*)$.

$$f(x) = \frac{1}{2}\langle x, Ax\rangle + \langle u, x\rangle$$

$$x^* = -A^{-1}u$$

$$f(x^*) = \frac{1}{2}\langle x^*, Ax^*\rangle + \langle u, x^*\rangle$$

3. Evaluate $f(\mathbf{x}^*)$.

$$f(x) = \frac{1}{2}\langle x, Ax\rangle + \langle u, x\rangle$$

$$x^* = -A^{-1}u$$

$$f(x^*) = \frac{1}{2}\langle x^*, Ax^*\rangle + \langle u, x^*\rangle \qquad \rightarrow f(x^*) = \frac{1}{2}\langle -A^{-1}u, -AA^{-1}u\rangle + \langle u, -A^{-1}u\rangle$$

3. Evaluate $f(\mathbf{x}^*)$.

$$f(x) = \frac{1}{2}\langle x, Ax\rangle + \langle u, x\rangle$$

$$x^* = -A^{-1}u$$

$$f(x^*) = \frac{1}{2}\langle x^*, Ax^*\rangle + \langle u, x^*\rangle \qquad \to f(x^*) = \frac{1}{2}\langle -A^{-1}u, -AA^{-1}u\rangle + \langle u, -A^{-1}u\rangle$$

$$f(x^*) = \frac{1}{2}\langle A^{-1}u, u\rangle - \langle u, A^{-1}u\rangle$$

3. Evaluate $f(\mathbf{x}^*)$.

$$f(x) = \frac{1}{2}\langle x, Ax \rangle + \langle u, x \rangle$$

$$x^* = -A^{-1}u$$

$$f(x^*) = \frac{1}{2}\langle x^*, Ax^* \rangle + \langle u, x^* \rangle \qquad \rightarrow f(x^*) = \frac{1}{2}\langle -A^{-1}u, -AA^{-1}u \rangle + \langle u, -A^{-1}u \rangle$$

$$f(x^*) = \frac{1}{2}\langle A^{-1}u, u \rangle - \langle u, A^{-1}u \rangle = -\frac{1}{2}\langle A^{-1}u, u \rangle$$

4. Calculate the Hessian $H_f(\mathbf{x}^*)$ and show it is positive. This will finish the proof of $\mathbf{x}^*$ being a minimizer point of $f(\mathbf{x})$.

4. Calculate the Hessian $H_f(\mathbf{x}^*)$ and show it is positive. This will finish the proof of $\mathbf{x}^*$ being a minimizer point of $f(\mathbf{x})$.

$$\mathbf{H}_f = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1\,\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1\,\partial x_n} \\[2.2ex] \dfrac{\partial^2 f}{\partial x_2\,\partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2\,\partial x_n} \\[2.2ex] \vdots & \vdots & \ddots & \vdots \\[2.2ex] \dfrac{\partial^2 f}{\partial x_n\,\partial x_1} & \dfrac{\partial^2 f}{\partial x_n\,\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

4. Calculate the Hessian $H_f(\mathbf{x}^*)$ and show it is positive. This will finish the proof of $\mathbf{x}^*$ being a minimizer point of $f(\mathbf{x})$.

4. Calculate the Hessian $H_f(\mathbf{x}^*)$ and show it is positive. This will finish the proof of $\mathbf{x}^*$ being a minimizer point of $f(\mathbf{x})$.

$$H_f = \begin{pmatrix} \partial_{x_1}(\nabla f)_{x_1} & \partial_{x_1}(\nabla f)_{x_2} \\ \partial_{x_2}(\nabla f)_{x_1} & \partial_{x_2}(\nabla f)_{x_2} \end{pmatrix}$$

4. Calculate the Hessian $H_f(\mathbf{x}^*)$ and show it is positive. This will finish the proof of $\mathbf{x}^*$ being a minimizer point of $f(\mathbf{x})$.

$$H_f = \begin{pmatrix} \partial_{x_1}(\nabla f)_{x_1} & \partial_{x_1}(\nabla f)_{x_2} \\ \partial_{x_2}(\nabla f)_{x_1} & \partial_{x_2}(\nabla f)_{x_2} \end{pmatrix}$$

$$\partial_{x_p} f(x) = \sum_j A_{pj} x_j + u_p$$

4. Calculate the Hessian $H_f(\mathbf{x}^*)$ and show it is positive. This will finish the proof of $\mathbf{x}^*$ being a minimizer point of $f(\mathbf{x})$.

$$H_f = \begin{pmatrix} \partial_{x_1}(\nabla f)_{x_1} & \partial_{x_1}(\nabla f)_{x_2} \\ \partial_{x_2}(\nabla f)_{x_1} & \partial_{x_2}(\nabla f)_{x_2} \end{pmatrix}$$

$$\partial_{x_p} f(x) = \sum_j A_{pj} x_j + u_p$$

$$H_f = \begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{pmatrix}$$

4. Calculate the Hessian $H_f(\mathbf{x}^*)$ and show it is positive. This will finish the proof of $\mathbf{x}^*$ being a minimizer point of $f(\mathbf{x})$.

$$H_f = \begin{pmatrix} \partial_{x_1}(\nabla f)_{x_1} & \partial_{x_1}(\nabla f)_{x_2} \\ \partial_{x_2}(\nabla f)_{x_1} & \partial_{x_2}(\nabla f)_{x_2} \end{pmatrix}$$

$$\partial_{x_p} f(x) = \sum_j A_{pj} x_j + u_p$$

$$H_f = \begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{pmatrix}$$

where $\mathbf{A}$ is a real, symmetric, positive definite 2×2 matrix,

Problem 3

Problem 3. Consider the following function of real arguments ω_0, ω_1.

$$f(\omega_0, \omega_1) = \frac{1}{2s} \sum_{k=1}^{s} (\omega_0 + \omega_1 x_k - y_k)^2, \tag{1}$$

where $x_1, x_2, \ldots, x_s$ and $y_1, y_2, \ldots, y_s$ are real-valued constants. This function measures the mean squared error of a linear approximation for the data points $\{(x_i, y_i)\}_{k=1}^{s}$.

(i) Find partial derivatives $\frac{\partial}{\partial \omega_0} f(\omega_0, \omega_1)$, $\frac{\partial}{\partial \omega_1} f(\omega_0, \omega_1)$, $\frac{\partial^2}{\partial \omega_0^2} f(\omega_0, \omega_1)$, $\frac{\partial^2}{\partial \omega_0 \partial \omega_1} f(\omega_0, \omega_1)$, $\frac{\partial^2}{\partial \omega_1^2} f(\omega_0, \omega_1)$.

Hint: you should be able to show that

$$\frac{\partial}{\partial \omega_0} f(\omega_0, \omega_1) = \omega_0 + \overline{\mathbf{x}} w_1 - \overline{\mathbf{y}}, \qquad \frac{\partial}{\partial \omega_1} f(\omega_0, \omega_1) = \overline{\mathbf{x}} \omega_0 + \overline{\mathbf{x^2}} w_1 - \overline{\mathbf{xy}}.$$

where

$$\overline{\mathbf{x}} = \frac{1}{s} \sum_{k=1}^{s} x_k, \quad \overline{\mathbf{x^2}} = \frac{1}{s} \sum_{k=1}^{s} x_k^2, \quad \overline{\mathbf{y}} = \frac{1}{s} \sum_{k=1}^{s} y_k, \quad \overline{\mathbf{xy}} = \frac{1}{s} \sum_{k=1}^{s} x_k y_k.$$

$$f\left(\omega_0, \omega_1\right) = \frac{1}{2s} \sum_{k=1}^{s} \left(\omega_0 + \omega_1 x_k - y_k\right)^2 ,$$

$$f\left(\omega_0, \omega_1\right) = \frac{1}{2s} \sum_{k=1}^{s} \left(\omega_0 + \omega_1 x_k - y_k\right)^2 ,$$

$$\frac{\partial f}{\partial \omega_0} = \frac{1}{s} \sum_{k=1}^{s} \left(\omega_0 + \omega_1 x_k - y_k\right) = \omega_0 + \omega_1 \overline{\mathbf{x}} - \overline{\mathbf{y}},$$

$$f(\omega_0, \omega_1) = \frac{1}{2s} \sum_{k=1}^{s} (\omega_0 + \omega_1 x_k - y_k)^2 \, ,$$

$$\frac{\partial f}{\partial \omega_0} \quad = \quad \frac{1}{s} \sum_{k=1}^{s} (\omega_0 + \omega_1 x_k - y_k) = \omega_0 + \omega_1 \overline{\mathbf{x}} - \overline{\mathbf{y}},$$

$$\frac{\partial^2 f}{\partial \omega_0^2} \quad = \quad \left(\frac{\partial}{\partial \omega_0} \omega_0 + \omega_1 \overline{\mathbf{x}} - \overline{\mathbf{y}} \right) = 1,$$

$$f(\omega_0, \omega_1) = \frac{1}{2s} \sum_{k=1}^{s} (\omega_0 + \omega_1 x_k - y_k)^2 ,$$

$$f(\omega_0, \omega_1) = \frac{1}{2s} \sum_{k=1}^{s} (\omega_0 + \omega_1 x_k - y_k)^2 \,,$$

$$\frac{\partial f}{\partial \omega_1} = \frac{1}{s} \sum_{k=1}^{s} x_k (\omega_0 + \omega_1 x_k - y_k) = \omega_0 \overline{\mathbf{x}} + \omega_1 \overline{\mathbf{x^2}} - \overline{\mathbf{xy}},$$

$$f\left(\omega_0, \omega_1\right) = \frac{1}{2s}\sum_{k=1}^{s}\left(\omega_0 + \omega_1 x_k - y_k\right)^2,$$

$$\frac{\partial f}{\partial \omega_1} = \frac{1}{s}\sum_{k=1}^{s} x_k\left(\omega_0 + \omega_1 x_k - y_k\right) = \omega_0\overline{\mathbf{x}} + \omega_1\overline{\mathbf{x^2}} - \overline{\mathbf{xy}},$$

$$\frac{\partial^2 f}{\partial \omega_1^2} = \frac{\partial}{\partial \omega_1}\left(\omega_0\overline{\mathbf{x}} + \omega_1\overline{\mathbf{x^2}} - \overline{\mathbf{xy}}\right) = \overline{\mathbf{x^2}},$$

$$f\left(\omega_0, \omega_1\right) = \frac{1}{2s} \sum_{k=1}^{s} \left(\omega_0 + \omega_1 x_k - y_k\right)^2,$$

$$\frac{\partial f}{\partial \omega_1} = \frac{1}{s} \sum_{k=1}^{s} x_k \left(\omega_0 + \omega_1 x_k - y_k\right) = \omega_0 \overline{\mathbf{x}} + \omega_1 \overline{\mathbf{x^2}} - \overline{\mathbf{xy}},$$

$$\frac{\partial^2 f}{\partial \omega_1^2} = \frac{\partial}{\partial \omega_1} \left(\omega_0 \overline{\mathbf{x}} + \omega_1 \overline{\mathbf{x^2}} - \overline{\mathbf{xy}}\right) = \overline{\mathbf{x^2}},$$

$$\frac{\partial^2 f}{\partial \omega_0 \partial \omega_1} = \frac{\partial}{\partial \omega_1} \left(\omega_0 + \omega_1 \overline{\mathbf{x}} - \overline{\mathbf{y}}\right) = \overline{\mathbf{x}},$$

(ii) Using the above results, find values ω_0^*, ω_1^* such that

$$\nabla f\left(\omega_0^*, \omega_1^*\right) = 0$$

(ii) Using the above results, find values ω_0^*, ω_1^* such that

$$\nabla f\left(\omega_0^*, \omega_1^*\right) = 0$$

$$\frac{\partial f}{\partial \omega_0} = \frac{1}{s}\sum_{k=1}^{s}\left(\omega_0 + \omega_1 x_k - y_k\right) = \omega_0 + \omega_1 \overline{\mathbf{x}} - \overline{\mathbf{y}},$$

(ii) Using the above results, find values ω_0^*, ω_1^* such that

$$\nabla f\left(\omega_0^*, \omega_1^*\right) = 0$$

$$\frac{\partial f}{\partial \omega_0} = \frac{1}{s}\sum_{k=1}^{s}\left(\omega_0 + \omega_1 x_k - y_k\right) = \omega_0 + \omega_1 \overline{\mathbf{x}} - \overline{\mathbf{y}},$$

$$\frac{\partial f}{\partial \omega_1} = \frac{1}{s}\sum_{k=1}^{s} x_k\left(\omega_0 + \omega_1 x_k - y_k\right) = \omega_0 \overline{\mathbf{x}} + \omega_1 \overline{\mathbf{x^2}} - \overline{\mathbf{xy}},$$

(ii) Using the above results, find values ω_0^*, ω_1^* such that

$$\nabla f\left(\omega_0^*, \omega_1^*\right) = 0$$

$$\frac{\partial f}{\partial \omega_0} = \frac{1}{s}\sum_{k=1}^{s}\left(\omega_0 + \omega_1 x_k - y_k\right) = \omega_0 + \omega_1 \overline{\mathbf{x}} - \overline{\mathbf{y}},$$

$$\frac{\partial f}{\partial \omega_1} = \frac{1}{s}\sum_{k=1}^{s} x_k\left(\omega_0 + \omega_1 x_k - y_k\right) = \omega_0 \overline{\mathbf{x}} + \omega_1 \overline{\mathbf{x^2}} - \overline{\mathbf{xy}},$$

$$\begin{cases} \omega_0^* + \quad \omega_1^* \overline{\mathbf{x}} = \overline{\mathbf{y}} \\ \omega_0^* \overline{\mathbf{x}} + \quad \omega_1^* \overline{\mathbf{x^2}} = \overline{\mathbf{xy}} \end{cases}.$$

(ii) Using the above results, find values ω_0^*, ω_1^* such that

$$\nabla f\left(\omega_0^*, \omega_1^*\right) = 0$$

$$\frac{\partial f}{\partial \omega_0} = \frac{1}{s}\sum_{k=1}^{s}\left(\omega_0 + \omega_1 x_k - y_k\right) = \omega_0 + \omega_1 \overline{\mathbf{x}} - \overline{\mathbf{y}},$$

$$\frac{\partial f}{\partial \omega_1} = \frac{1}{s}\sum_{k=1}^{s} x_k\left(\omega_0 + \omega_1 x_k - y_k\right) = \omega_0 \overline{\mathbf{x}} + \omega_1 \overline{\mathbf{x^2}} - \overline{\mathbf{xy}},$$

$$\begin{cases} \omega_0^* + \quad \omega_1^* \overline{\mathbf{x}} = \overline{\mathbf{y}} \\ \omega_0^* \overline{\mathbf{x}} + \quad \omega_1^* \overline{\mathbf{x^2}} = \overline{\mathbf{xy}} \end{cases}.$$

$$\omega_0^* = \frac{\overline{\mathbf{y}} \cdot \overline{\mathbf{x^2}} - \overline{\mathbf{x}} \cdot \overline{\mathbf{xy}}}{\overline{\mathbf{x^2}} - \overline{\mathbf{x}}^2}, \quad \omega_1^* = \frac{\overline{\mathbf{xy}} - \overline{\mathbf{x}} \cdot \overline{\mathbf{y}}}{\overline{\mathbf{x^2}} - \overline{\mathbf{x}}^2}.$$

(iii) Using the expressions of second order derivatives obtained in (i) find the value of Hessian of the function f for $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$. Prove it is positive definite and thus show that $f(\omega_0, \omega_1)$ attains its minimum value at $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$.

(iii) Using the expressions of second order derivatives obtained in (i) find the value of Hessian of the function f for $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$. Prove it is positive definite and thus show that $f(\omega_0, \omega_1)$ attains its minimum value at $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$.

The Hessian of function f is equal to

$$H_f(\omega_0, \omega_1) = \begin{pmatrix} 1 & \overline{\mathbf{x}} \\ \overline{\mathbf{x}} & \overline{\mathbf{x}^2} \end{pmatrix}.$$

(iii) Using the expressions of second order derivatives obtained in (i) find the value of Hessian of the function f for $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$. Prove it is positive definite and thus show that $f(\omega_0, \omega_1)$ attains its minimum value at $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$.

The Hessian of function f is equal to

$$H_f(\omega_0, \omega_1) = \begin{pmatrix} 1 & \overline{\mathbf{x}} \\ \overline{\mathbf{x}} & \overline{\mathbf{x}^2} \end{pmatrix}.$$

Trace>0, matrix is symmetric, hence it is positive definite if the determinant is positive

(iii) Using the expressions of second order derivatives obtained in (i) find the value of Hessian of the function f for $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$. Prove it is positive definite and thus show that $f(\omega_0, \omega_1)$ attains its minimum value at $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$.

The Hessian of function f is equal to

$$H_f(\omega_0, \omega_1) = \begin{pmatrix} 1 & \overline{\mathbf{x}} \\ \overline{\mathbf{x}} & \overline{\mathbf{x}^2} \end{pmatrix}.$$

Trace>0, matrix is symmetric, hence it is positive definite if the determinant is positive

$$\det H_f(\omega_0, \omega_1) = \overline{\mathbf{x}^2} - \overline{\mathbf{x}}^2,$$

(iii) Using the expressions of second order derivatives obtained in (i) find the value of Hessian of the function f for $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$. Prove it is positive definite and thus show that $f(\omega_0, \omega_1)$ attains its minimum value at $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$.

$$\det H_f(\omega_0, \omega_1) = \overline{\mathbf{x^2}} - \overline{\mathbf{x}}^2,$$

(iii) Using the expressions of second order derivatives obtained in (i) find the value of Hessian of the function f for $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$. Prove it is positive definite and thus show that $f(\omega_0, \omega_1)$ attains its minimum value at $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$.

$$\det H_f(\omega_0, \omega_1) = \overline{\mathbf{x^2}} - \overline{\mathbf{x}}^2,$$

$$\frac{1}{s}\sum_{i=1}^{s}(x_i - \bar{x})^2 \geq 0$$

(iii) Using the expressions of second order derivatives obtained in (i) find the value of Hessian of the function f for $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$. Prove it is positive definite and thus show that $f(\omega_0, \omega_1)$ attains its minimum value at $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$.

$$\det H_f(\omega_0, \omega_1) = \overline{\mathbf{x^2}} - \overline{\mathbf{x}}^2,$$

$$\frac{1}{s}\sum_{i=1}^{s}(x_i - \bar{x})^2 \geq 0$$

$$\frac{1}{s}\sum_{i=1}^{s}(x_i^2 + \bar{x}^2 - 2x_i\bar{x}) = \frac{1}{s}\sum_{i=1}^{s}x_i^2 + \bar{x}^2 - \frac{2\bar{x}}{s}\sum_{i=1}^{s}x_i \geq 0$$

(iii) Using the expressions of second order derivatives obtained in (i) find the value of Hessian of the function f for $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$. Prove it is positive definite and thus show that $f(\omega_0, \omega_1)$ attains its minimum value at $\omega_0 = \omega_0^*$ and $\omega_1 = \omega_1^*$.

$$\det H_f(\omega_0, \omega_1) = \overline{\mathbf{x^2}} - \overline{\mathbf{x}}^2,$$

$$\frac{1}{s}\sum_{i=1}^{s}(x_i - \bar{x})^2 \geq 0$$

$$\frac{1}{s}\sum_{i=1}^{s}(x_i^2 + \bar{x}^2 - 2x_i\bar{x}) = \frac{1}{s}\sum_{i=1}^{s}x_i^2 + \bar{x}^2 - \frac{2\bar{x}}{s}\sum_{i=1}^{s}x_i \geq 0$$

$$\overline{x^2} + \bar{x}^2 - 2\bar{x}^2 = \overline{x^2} - \bar{x}^2 \geq 0$$