MTH786, Semester A, 2023/24

Coursework 1 **N. Perra**

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.

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 $f(\mathbf{x})$.

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

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

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

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

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

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

. These should be functions of $x_1, x_2, \ldots, x_s$ and $y_1, y_2, \ldots, y_s$.

Hint: you should obtain

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

(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 $\omega_1 = \omega_1^*$.

(iv) Find $\min_{\omega_0, \omega_1 \in \mathbb{R}} f(\omega_0, \omega_1)$.

Problem 4. As you may have seen in the Lecture 2, the function considered in previous question is a mean-squared error function of a linear regression for data set $\{x_k, y_k\}_{k=1}^{s}$. Corresponding values ω_0^*, ω_1^* are the coefficients of a linear regression model for data set $\{x_k, y_k\}_{k=1}^{s}$. Using the above result find the linear regression, i.e. coefficients ω_0^*, ω_1^* for the following people's height/weight data.

Weight	162.31	183.93	154.34	187.50	187.06	173.42
Height	68.78	68.79	68.50	68.62	68.25	68.49

Now let us add one more data point and recalculate the regression.

Weight	162.31	183.93	154.34	187.50	187.06	173.42	192.34
Height	68.78	68.79	68.50	68.62	68.25	68.49	68.14

Can you explain the origin of such a difference between two results?