Version MAG/4

EGT2
ENGINEERING TRIPOS PART IIA

Monday 2 May 2022   2 to 3.40

Module 3M1

MATHEMATICAL METHODS

*Answer **all** the questions.*

All questions carry the same number of marks.

*The **approximate** percentage of marks allocated to each part of a question is indicated in the right margin.*

*Write your candidate number **not** your name on the cover sheet.*

STATIONERY REQUIREMENTS
Write on single-sided paper.

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM
CUED approved calculator allowed.
You are allowed access to the Engineering Data Books.

10 minutes reading time is allowed for this paper at the start of the exam.

You may not start to read the questions printed on the subsequent pages of this question paper until instructed to do so.

You may not remove any stationery from the Examination Room.

1 Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & v \\ v & 5 \end{bmatrix}$$

(a) Find the range of v values for which $\mathbf{A}$ is definite positive [20%]

(b) We would like to solve the equation $\mathbf{Ax} = \mathbf{b}$ using a stationary method, in which $\mathbf{A}$ is split into the following two terms: $\mathbf{A} = \mathbf{N} - \mathbf{P}$.

(i) Explain how the method can be used to find the solution. [20%]

(ii) Establish a criterion for the convergence of the method. [20%]

(c) (i) What are the values of v for which the Jacobi method ($\mathbf{N} = \text{diag}(\mathbf{A})$) converges? [20%]

(ii) Would the Richardson iteration ($\mathbf{N} = \mathbf{Id}$) be able to provide solutions beyond this range? [20%]

2 For D-dimensional column vectors $\mathbf{x} \in \mathbb{R}^D$ and $\mathbf{b} \in \mathbb{R}^D$ consider the function of $\mathbf{x}$

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{Q}\mathbf{x} - \mathbf{x}^\top \mathbf{b} + \mathbf{1}^\top \mathbf{g}(\mathbf{x})$$

where the $D \times D$ matrix $\mathbf{Q}$ has elements $q_{i,j}$, each $i = 1, \cdots, D$ and $j = 1, \cdots, D$, the D-dimensional vector of ones is denoted as $\mathbf{1} = [1, 1, \cdots, 1]^\top$ and $\mathbf{g}(\mathbf{x})$ is an element-wise application of the function $g(\cdot)$ which acts on each component of the vector $\mathbf{x}$ such that $\mathbf{g}(\mathbf{x}) = [g(x_1), g(x_2), \cdots, g(x_D)]^\top$.

(a) Derive and justify the necessary condition for a point $\mathbf{x}^* \in \mathbb{R}^D$ to be a strong local minimum of the function $f(\mathbf{x})$. [15%]

(b) Derive and justify the sufficient condition for a point $\mathbf{x}^* \in \mathbb{R}^D$ to be a strong local minimum of the function $f(\mathbf{x})$. [15%]

(c) Derive a steepest descent method for the function $f(\mathbf{x})$ acting on $\mathbf{x} \in \mathbb{R}^D$ and provide an expression for the step-size for each iteration based on a second-order Taylor expansion of $f(\mathbf{x})$. [40%]

(d) For the specific case where $D = 2$, $\mathbf{Q}$ is given as an identity matrix, and the nonlinear term is the exponential function, i.e. $g(\cdot) = \exp(\cdot)$, assess whether $f(\mathbf{x})$ is convex in $\mathbb{R}^2$ and state the implication on the nature of the point $\mathbf{x}^*$. [30%]

 (TURN OVER

3 For an experiment it is necessary to compute the expected value of a function. The expected value, μ, can be expressed as

$$\mu = \int h(\boldsymbol{x})p(\boldsymbol{x})d\boldsymbol{x}$$

where $h(\boldsymbol{x})$ is the value of the function for point $\boldsymbol{x}$, $\boldsymbol{x} \in \mathbb{R}^D$, and $p(\boldsymbol{x})$ is the probability of selecting the point $\boldsymbol{x}$. It is possible to obtain the value for the function $h(\boldsymbol{x})$ for any value of $\boldsymbol{x}$.

(a) Initially it is assumed that it is possible to draw samples from $p(\boldsymbol{x})$. N samples are drawn from $p(\boldsymbol{x})$, $\{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N)}\}$.

(i) Derive the Monte-Carlo approximation for the value of μ, $\hat{\mu}_N$, in terms of the N samples, $\{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N)}\}$. [10%]

(ii) Show that the variance of the estimate $\hat{\mu}_N$ using N samples, $\text{var}_p(\hat{\mu}_N)$, can be expressed as

$$\text{var}_p(\hat{\mu}_N) = \frac{\sigma^2}{N}; \quad \text{where } \sigma^2 = \int (h(\boldsymbol{x}) - \mu)^2 \, p(\boldsymbol{x})d\boldsymbol{x}$$

Comment on this result as the dimensionality of $\boldsymbol{x}$ increases. [30%]

(b) It is not possible to sample from $p(\boldsymbol{x})$. Instead N samples from a distribution $q(\boldsymbol{x})$, $\{\tilde{\boldsymbol{x}}^{(1)}, \ldots, \tilde{\boldsymbol{x}}^{(N)}\}$ can be drawn. It is possible to compute $p(\boldsymbol{x})$ and $q(\boldsymbol{x})$ for any value of $\boldsymbol{x}$.

(i) Show that a new approximation for the integral, $\tilde{\mu}_N$, can be expressed as

$$\tilde{\mu}_N = \frac{1}{N}\sum_{i=1}^{N} h(\tilde{\boldsymbol{x}}^{(i)})w^{(i)}$$

and give the expression for $w^{(i)}$. [20%]

(ii) What are the requirements for this approximation to converge to μ as $N \rightarrow \infty$? [10%]

(iii) When the requirements in b(ii) are satisfied, show that the difference between the variance of the estimate in Part (a)(i), $\text{var}_p(\hat{\mu}_N)$, and the variance of the estimate in Part (b)(i), $\text{var}_q(\tilde{\mu}_N)$, can be expressed as

$$\text{var}_p(\hat{\mu}_N) - \text{var}_q(\tilde{\mu}_N) = \frac{1}{N}\int h(\boldsymbol{x})^2 \left(1 - \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})}\right) p(\boldsymbol{x})d\boldsymbol{x}$$

Comment on the implications of this expression for the selection of $q(\boldsymbol{x})$. [30%]

END OF PAPER