ENGINEERING TRIPOS PART IIA

Thursday 2 May 2013 9.30 to 11

Module 3M1

MATHEMATICAL METHODS

Answer not more than ***three*** *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.*

Attachment:

3M1 data sheet (4 pages).

STATIONERY REQUIREMENTS
Single-sided script paper

SPECIAL REQUIREMENTS
Engineering Data Book
CUED approved calculator allowed

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

Final version

1 (a) Explain what is meant by "a symmetric matrix **A** is positive definite". [10%]

(b) For what range of s is the matrix **A** positive definite, where

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 1 & -s \\ -1 & -s & 1 \end{bmatrix} \quad ?$$ [15%]

(c) For the case $s = 0.5$, express **A** in the form

$$\mathbf{A} = \mathbf{L}\mathbf{L}^t$$

where **L** is a lower triangular matrix with positive diagonal elements. [25%]

(d) Show that a symmetric matrix **A** is positive definite if, and only if, all of its eigenvalues are positive. [25%]

(e) For any positive definite symmetric matrix **A**, show that **A** can be written in the form

$$\mathbf{A} = \mathbf{B}^2$$

where the symmetric matrix **B** is *negative definite*. [25%]

2 An engineer is designing a solid cylindrical heat sink of height H and diameter D to be mounted on a thermally insulated surface, as shown schematically in Fig. 1.

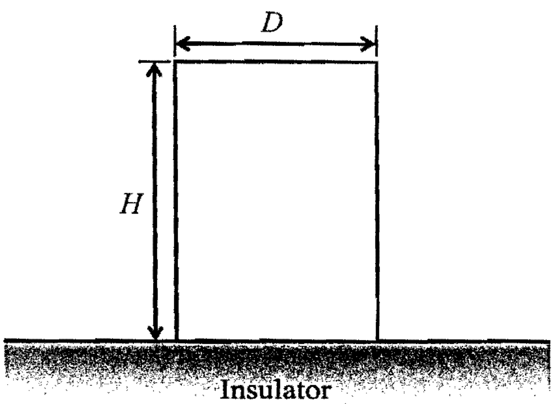


Fig. 1

The objective is to maximise the rate of heat transfer to the surrounding air. The rate of heat transfer is proportional to the exposed surface area of the cylinder A. The mass M of the cylinder must not exceed a specified maximum value M_{max}. For acceptable structural stability of the cylinder the height-to-diameter ratio $R = H/D$ must not exceed a specified maximum value R_{max}. Space constraints limit permissible values of D to the range $0 < D \le D_{max}$.

(a) Show that a suitable objective function to be minimised is

$$f(D,R) = -CD^2\left(R + \tfrac{1}{4}\right)$$

where C is a constant. [10%]

(b) Express the constraints that apply in terms of D and R. [10%]

(c) By considering the gradient of the objective function, show that there can be no unconstrained minimum for this problem. [15%]

(d) Assuming that the constraint on M is active at the minimum, reformulate the task as a univariate constrained minimisation problem in terms of D. What condition must be met by the specified values of R_{max} and D_{max} so as not to preclude the possibility that the constraint on M is active? [20%]

(e) By considering appropriate optimality criteria, determine whether an unconstrained minimum exists for this univariate problem, and, if so, under what conditions. [20%]

(f) Using your findings from the earlier parts of the question, set out the possible solutions to this minimisation problem and explain how they depend on the specified values of M_{max}, R_{max} and D_{max}. [25%]

3 The enrichment of uranium is defined as the ratio (kg per kg) of the mass of the isotope U-235 to the total mass of uranium.

A uranium enrichment plant is fed with F kg of natural uranium, which has an enrichment x_f of 0.00715. The plant is to produce 1 kg of product with an enrichment x_p of 0.035, and in doing so will produce W kg of waste with enrichment x_w, where $x_w < x_f < x_p$.

Simple conservation of U-235 and of uranium requires that

$$F = 1 + W \quad \text{and} \quad x_f F = x_p + x_w W$$

where x_f and x_p are fixed at the values given above.

The cost per unit mass of the product C_p is given by

$$C_p = C_f F + C_s\left[F \ln(x_f) - \ln(x_p) - W \ln(x_w)\right]$$

where C_f is the cost per unit mass of the feed, C_s is cost of running the plant per unit of 'separative work' and the separative work is given by the expression in square brackets.

(a) The plant engineer wants to minimise C_p by choosing appropriate values for x_w, F and W. Formulate this task as an equality constrained optimisation problem in standard form. [10%]

(b) Although this problem can be solved by eliminating control variables using the constraint equations, it is much easier to solve it using the method of Lagrange multipliers. Find appropriate values of the Lagrange multipliers, and, hence, show that at the optimal solution

$$\frac{x_f}{x_w} = 1 + \frac{C_f}{C_s} + \ln\left(\frac{x_f}{x_w}\right)$$ [60%]

(c) Find the optimal values of x_w, F and W if $\frac{C_f}{C_s} = 1$. [20%]

(d) Discuss the significance of the values of the Lagrange multipliers at the optimum. [10%]

4 (a) Explain what is meant by the term *a limiting distribution* for a finite state space Markov chain. [10%]

(b) Describe the properties of the Markov chain under which you would expect a limiting distribution *not* to be achievable in practice. [20%]

(c) A system X can exhibit three states (a, b, c). The system undergoes a series of transitions through states $X_0, X_1, X_2, \ldots$ whereby at each stage of the process:

(i) the probability that state a changes to state b is 1/4 and the probability that it changes to state c is 1/4;

(ii) the probability that state b changes to state a is 1/8 and the probability that it changes to state c is 3/8;

(iii) the probability that state c changes to state a is 1/4 and the probability that it changes to state b is 3/4.

Find the transition matrix **P** that governs the process. [10%]

(d) Find the limiting distribution and show, carefully, that it will be attained. [40%]

(e) Explain what is meant by a limiting distribution being in *dynamic balance* and show that the limiting distribution found in (d) is in dynamic balance. [20%]

[You may quote, without proof, the results that the sum of the eigenvalues of a matrix is equal to the sum of the diagonal terms and that the product of the eigenvalues is equal to the determinant.]

END OF PAPER

3M1

OPTIMIZATION

DATA SHEET

1. **Taylor Series Expansion**

For one variable:

$$f(x) = f(x^*) + (x-x^*)f'(x^*) + \frac{1}{2}(x-x^*)^2 f''(x^*) + R$$

For several variables:

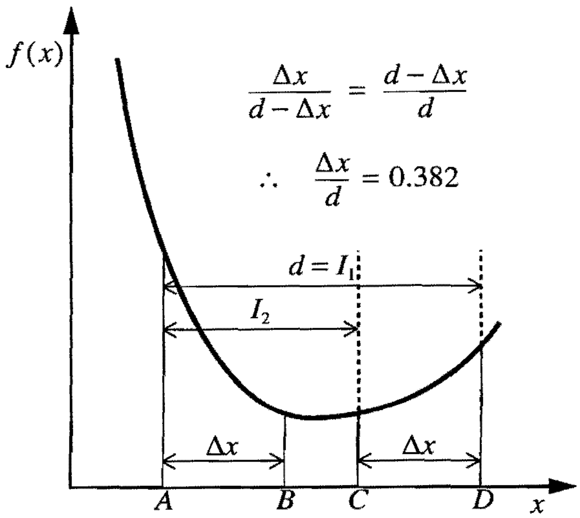
$$f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T(\mathbf{x}-\mathbf{x}^*) + \frac{1}{2}(\mathbf{x}-\mathbf{x}^*)^T \boldsymbol{H}(\mathbf{x}^*)(\mathbf{x}-\mathbf{x}^*) + R$$

where

$$\text{gradient } \nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \quad \text{and hessian } \boldsymbol{H}(\mathbf{x}) = \nabla(\nabla f(\mathbf{x})) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

$\boldsymbol{H}(\mathbf{x}^*)$ is a symmetric $n \times n$ matrix and R includes all higher order terms.

2. **Golden Section Method**



(a) Evaluate $f(x)$ at points A, B, C and D.

(b) If $f(B) < f(C)$, new interval is $A-C$.

If $f(B) > f(C)$, new interval is $B-D$.
If $f(B) = f(C)$, new interval is either $A-C$ or $B-D$.

(c) Evaluate $f(x)$ at new interior point. If not converged, go to (b).

3. **Newton's Method**

(a) Select starting point $\mathbf{x}_0$

(b) Determine search direction $\mathbf{d}_k = -H(\mathbf{x}_k)^{-1}\nabla f(\mathbf{x}_k)$

(c) Determine new estimate $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k$

(d) Test for convergence. If not converged, go to step (b)

4. **Steepest Descent Method**

(a) Select starting point $\mathbf{x}_0$

(b) Determine search direction $\mathbf{d}_k = -\nabla f(\mathbf{x}_k)$

(c) Perform line search to determine step size α_k or evaluate $\alpha_k = \dfrac{\mathbf{d}_k^T \mathbf{d}_k}{\mathbf{d}_k^T H(\mathbf{x}_k)\mathbf{d}_k}$

(d) Determine new estimate $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$

(e) Test for convergence. If not converged, go to step (b)

5. **Conjugate Gradient Method**

(a) Select starting point $\mathbf{x}_0$ and compute $\mathbf{d}_0 = -\nabla f(\mathbf{x}_0)$ and $\alpha_0 = \dfrac{\mathbf{d}_0^T \mathbf{d}_0}{\mathbf{d}_0^T H(\mathbf{x}_0)\,\mathbf{d}_0}$

(b) Determine new estimate $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$

(c) Evaluate $\nabla f(\mathbf{x}_{k+1})$ and $\beta_k = \left[\dfrac{|\nabla f(\mathbf{x}_{k+1})|}{|\nabla f(\mathbf{x}_k)|}\right]^2$

(d) Determine search direction $\mathbf{d}_{k+1} = -\nabla f(\mathbf{x}_{k+1}) + \beta_k \mathbf{d}_k$

(e) Determine step size $\alpha_{k+1} = -\dfrac{\mathbf{d}_{k+1}^T \nabla f(\mathbf{x}_{k+1})}{\mathbf{d}_{k+1}^T H(\mathbf{x}_{k+1})\,\mathbf{d}_{k+1}}$

(f) Test for convergence. If not converged, go to step (b)

6. **Gauss-Newton Method (for Nonlinear Least Squares)**

If the minimum squared error of residuals $\mathbf{r}(\mathbf{x})$ is sought:

$$\text{Minimise } f(\mathbf{x}) = \sum_{i=1}^{m} r_i^2(\mathbf{x}) = \mathbf{r}(\mathbf{x})^T \mathbf{r}(\mathbf{x})$$

(a) Select starting point $\mathbf{x}_0$

(b) Determine search direction $\mathbf{d}_k = -[\, J(\mathbf{x}_k)^T J(\mathbf{x}_k)\,]^{-1} J(\mathbf{x}_k)^T \mathbf{r}(\mathbf{x}_k)$

$$\text{where } \boldsymbol{J}(\mathbf{x}) = \begin{bmatrix} \nabla r_1(\mathbf{x})^T \\ \vdots \\ \nabla r_m(\mathbf{x})^T \end{bmatrix} = \begin{bmatrix} \frac{\partial r_1}{\partial x_1} & \cdots & \frac{\partial r_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial r_m}{\partial x_1} & \cdots & \frac{\partial r_m}{\partial x_n} \end{bmatrix}$$

(c) Determine new estimate $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k$

(d) Test for convergence. If not converged, go to step (b)

7. **Lagrange Multipliers**

To minimise $f(\mathbf{x})$ subject to m equality constraints $h_i(\mathbf{x}) = 0$, $i = 1, \ldots, m$, solve the system of simultaneous equations

$$\begin{aligned} \nabla f(\mathbf{x}^*) + [\nabla \mathbf{h}(\mathbf{x}^*)]^T \boldsymbol{\lambda} &= 0 \quad (n \text{ equations}) \\ \mathbf{h}(\mathbf{x}^*) &= 0 \quad (m \text{ equations}) \end{aligned}$$

where $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_m]^T$ is the vector of Lagrange multipliers and

$$[\nabla \mathbf{h}(\mathbf{x}^*)]^T = \left[\nabla h_1(\mathbf{x}^*) \ \ldots \ \nabla h_m(\mathbf{x}^*) \right] = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_m}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_1}{\partial x_n} & \cdots & \frac{\partial h_m}{\partial x_n} \end{bmatrix}$$

8. **Kuhn-Tucker Multipliers**

To minimise $f(\mathbf{x})$ subject to m equality constraints $h_i(\mathbf{x}) = 0$, $i = 1, \ldots, m$ and p inequality constraints $g_i(\mathbf{x}) \le 0$, $i = 1, \ldots, p$, solve the system of simultaneous equations

$$\begin{aligned} \nabla f(\mathbf{x}^*) + [\nabla \mathbf{h}(\mathbf{x}^*)]^T \boldsymbol{\lambda} + [\nabla \mathbf{g}(\mathbf{x}^*)]^T \boldsymbol{\mu} &= 0 \quad (n \text{ equations}) \\ \mathbf{h}(\mathbf{x}^*) &= 0 \quad (m \text{ equations}) \\ \forall\ i = 1, \ldots, p, \quad \mu_i g_i(\mathbf{x}) &= 0 \quad (p \text{ equations}) \end{aligned}$$

where $\boldsymbol{\lambda}$ are Lagrange multipliers and $\boldsymbol{\mu} \ge 0$ are the Kuhn-Tucker multipliers.

9. **Penalty & Barrier Functions**

To minimise $f(\mathbf{x})$ subject to p inequality constraints $g_i(\mathbf{x}) \leq 0$, $i = 1, \ldots, p$, define

$$q(\mathbf{x}, p_k) = f(\mathbf{x}) + p_k P(\mathbf{x})$$

where $P(\mathbf{x})$ is a penalty function, e.g.

$$P(\mathbf{x}) = \sum_{i=1}^{p} (\max[0, g_i(\mathbf{x})])^2$$

or alternatively

$$q(\mathbf{x}, p_k) = f(\mathbf{x}) - \frac{1}{p_k} B(\mathbf{x})$$

where $B(\mathbf{x})$ is a barrier function, e.g.

$$B(\mathbf{x}) = \sum_{i=1}^{p} \frac{1}{g_i(\mathbf{x})}$$

Then for successive $k = 1, 2, \ldots$ and p_k such that $p_k > 0$ and $p_{k+1} > p_k$, solve the problem

$$\text{minimise } q(\mathbf{x}, p_k)$$