ENGINEERING TRIPOS PART IIA

Monday 30 April 2012 9 to 10.30

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 the *pseudo inverse*, $\mathbf{A}^{+}$, of a matrix **A**. [15%]

(b) By performing singular value decomposition on the matrix **A**, where

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

find a pseudo inverse of **A**. [50%]

(c) Explain how an $n \times n$ symmetric matrix **S** can be written in the form

$$\mathbf{S} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{t}$$

where **Q** is an orthogonal matrix and $\boldsymbol{\Lambda}$ is a diagonal one. [10%]

(d) Using a singular value decomposition, or otherwise, explain how a non-symmetric $n \times n$ matrix **B** can be written in the form

$$\mathbf{B} = \mathbf{Q}\mathbf{T}$$

where **Q** is an orthogonal matrix and **T** is a symmetric matrix whose eigenvalues cannot be negative. [25%]

2 The principle of stationary potential energy states that the equilibrium states of structural and mechanical systems are characterised by minima of the potential energy of the system. This principle can be used to find the equilibrium displacements of a structural system under load.

Consider the symmetric two-bar truss shown in Fig. 1. The structure is subject to a load W at node C. Under the action of this load node C moves to a point C′.

Assuming small displacements, the total potential energy of the system is given by:

$$P(x,y) = \frac{EA}{2L}\left[(x\cos\beta + y\sin\beta)^2 + (-x\cos\beta + y\sin\beta)^2\right] - W\left[x\cos\theta + y\sin\theta\right]$$

where E is Young's modulus of the material from which the truss is made, A is the cross-sectional area of the bars, L is the original length of the bars, x and y are the horizontal and vertical displacements of node C, and the angles β and θ are as defined in Fig. 1.

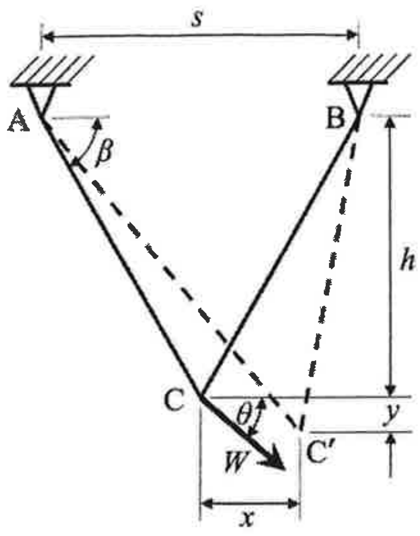


Fig. 1

(a) Show that for the case where $A = 10^{-5}$ m^2, $h = 1.0$ m, $s = 1.5$ m, $\theta = 30^\circ$, $W = 10$ kN and $E = 200$ GPa, the potential energy of the system is given by:

$$P(x, y) = 5.76\times10^{5}x^{2} + 1.024\times10^{6}y^{2} - 8.66\times10^{3}x - 5\times10^{3}y$$

h and s are defined in Fig. 1. [15%]

(b) Find the displacements x and y that minimize P using standard optimality criteria. [20%]

(c) Starting from the Taylor series expansion given on the 3M1 data sheet for the value of a multivariate function $f(\mathbf{x})$ near a point $\mathbf{x}_k$, show that the step size α_k that minimizes the function in a search direction $\mathbf{d}_k$ from $\mathbf{x}_k$ is given by:

$$\alpha_k = -\frac{\nabla f(\mathbf{x}_k)^T \mathbf{d}_k}{\mathbf{d}_k^{\ T} H(\mathbf{x}_k)\mathbf{d}_k}$$ [15%]

(d) Starting from a point $(x_1, y_1) = (0,0)$ and using the result in (c) execute two iterations of the Steepest Descent Method on $P(x, y)$. [35%]

(e) Comment on the performance of the Steepest Descent Method observed in (d) in light of the result in (b). How would you expect Newton's Method and the Conjugate Gradient Method to perform on this problem? [15%]

3 An engineer is designing a shell-and-tube heat exchanger. The design variables are the diameter D and length L of the heat exchanger. To provide sufficient heat transfer the total length of tubing T within the heat exchanger must be at least 100 m. For the design under consideration, T is related to D and L by the equation:

$$T = 15D^2L$$

The cost of the installation (in arbitrary units) has three components:

1. The cost of the tubes: $C_T = 150D^2L$
2. The cost of the shell: $C_S = 25D^{2.5}L$
3. The cost of the floor space occupied by the heat exchanger: $C_F = 20DL$

In all the equations above the values of D and L should be given in metres.

(a) Formulate the task of optimizing the design of the heat exchanger to minimize the cost of installation as a constrained minimization problem in standard form. [10%]

(b) Assuming that the constraint on the length of tubing is active at the optimum, use the Lagrange multiplier method to show that the optimal heat exchanger design has $D = 1.368$ m and $L = 3.562$ m. There is no need to check the second-order optimality conditions. [40%]

(c) If an additional constraint is introduced requiring that the floor space $A = DL$ occupied by the heat exchanger must not exceed 4 m^2, use the Kuhn-Tucker multiplier method to find the new optimal design. [50%]

4 (a) Explain what is meant by "A stochastic process $X_0, X_1, X_2, \ldots$ is a Markov chain" and how the properties of a finite state space Markov chain are determined by a transition matrix **P**. [15%]

(b) Prove that, for a homogeneous Markov chain, for any $s < n$

$$P(X_{i+n} = k \mid X_i = j) = \sum_l P(X_{i+s} = l \mid X_i = j) P(X_{i+n} = k \mid X_{i+s} = l)$$ [15%]

(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 the state does not change between X_n and X_{n+1} is 1/2;

(ii) the probability that state a changes to state b is 1/3 and the probability that it changes to state c is 1/6;

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

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

If the system is initially in state a, find the expected number of transitions before it is again in state a. [35%]

(d) A large number of processes of the type described in (c) are performed independently in parallel. Show that, after many transitions, the proportion of processes which are in any given state is independent of the starting states and find the proportion in each of the states a, b and c. [35%]

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:

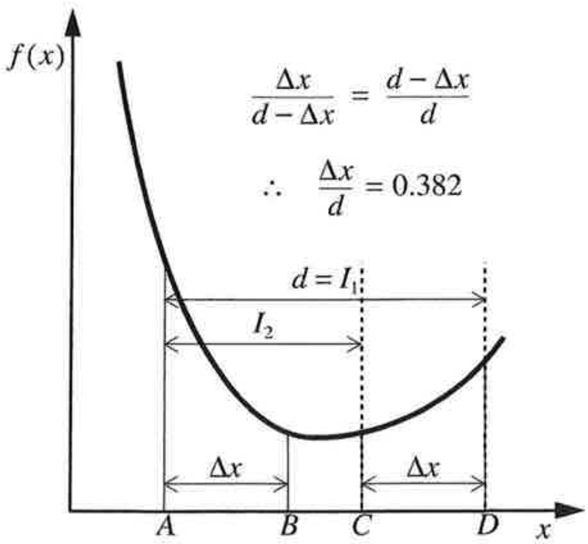
$$f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T (\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^T \boldsymbol{H}(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + R$$

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 $\mathbf{d}_k = -\boldsymbol{H}(\mathbf{x}_k)^{-1}\nabla f(\mathbf{x}_k)$

(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 $\alpha_k = \dfrac{\mathbf{d}_k^T\mathbf{d}_k}{\mathbf{d}_k^T\boldsymbol{H}(\mathbf{x}_k)\mathbf{d}_k}$

(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 $\alpha_0 = \dfrac{\mathbf{d}_0^T\mathbf{d}_0}{\mathbf{d}_0^T\boldsymbol{H}(\mathbf{x}_0)\,\mathbf{d}_0}$

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

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

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

(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 $g_i(\mathbf{x}) \le 0$, $i = 1, \ldots, p$, 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

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

3M1 Mathematical Methods 2012

Answers

Q1 (b) $\begin{bmatrix} \frac{1}{24} & \frac{1}{6} & -\frac{5}{24} \\ -\frac{5}{24} & \frac{1}{6} & \frac{1}{24} \end{bmatrix}$ (other equivalent results are possible)

Q2 (b) $x = 7.52\times10^{-3}$ m; $y = 2.44\times10^{-3}$ m

(d) $\mathbf{x}_2 = \begin{bmatrix} 6.294\times10^{-3} \\ 3.634\times10^{-3} \end{bmatrix}$; $\mathbf{x}_3 = \begin{bmatrix} 7.066\times10^{-3} \\ 2.295\times10^{-3} \end{bmatrix}$

Q3 (c) $D = \frac{5}{3}$ m; $L = \frac{12}{5}$ m

Q4 (c) 3

(d) $\left(\frac{9}{27} \quad \frac{10}{27} \quad \frac{8}{27}\right)$