JMAT 7304

Degree Master of Science in Mathematical Modelling and Scientific Computing

Numerical Linear Algebra & Continuous Optimization

TRINITY TERM 2015
Friday 24th April 2015, 9.30 a.m. – 11:30 a.m.

Candidates should submit answers to a maximum of four questions that include an answer to at least one question in each section.

Please start the answer to each question on a new page.

All questions will carry equal marks.

Do not turn over until told that you may do so.

Part A — Numerical Linear Algebra

Question 1

(a) Let A be a matrix with eigenvalue λ_1 of algebraic and multiplicity one and with magnitude strictly greater than the magnitude of any other eigenvalue of A. State the power method for estimating λ_1. Derive a rate of convergence of the approximation to λ_1 supplied by the power method.

[7 marks]

(b) Define the rotation matrix $J(i, j, \theta)$ with k^{th} columns

$$J(i,j,\theta)_k := \begin{cases} \cos(\theta)e_i - \sin(\theta)e_j & \text{for } k = i \\ \sin(\theta)e_i + \cos(\theta)e_j & \text{for } k = j \\ e_k & \text{otherwise} \end{cases}$$

and let A be a real valued symmetric matrix. What value of θ should be selected so that the (i, j) entry of $J(i, j, \theta)^*AJ(i, j, \theta)$ is zero? State an algorithm for approximating the eigenvalues of A using matrices of the form $J(i, j, \theta)$. What are the number of floating point operations to compute $J(i, j, \theta)^*AJ(i, j, \theta)$? Explain, but do not prove, why this algorithm will converge to the eigenvalues of A.

[10 marks]

(c) State the singular value decomposition of a matrix $A \in \mathbb{R}^{m\times n}$. Let $A_0 = A$ and $A_i = U_iA_{i-1}V_i$ for $i = 1, 2, \ldots, r = \min(m, n)$. State formulae for unitary matrices $U_i \in \mathbb{R}^{m\times m}$ and $V_i \in \mathbb{R}^{n\times n}$ such that A_r is nonzero only on its entries with index (k, k) for $k = 1, 2, \ldots, r$ and $(k, k + 1)$ for $k = 1, 2, \ldots, r - 1$.

[8 marks]

Question 2

(a) Determine a unitary matrix Q such that for any $A \in \mathbb{R}^{m\times m}$ the product QAQ^* is upper Hessenberg, that is the (i, j) entry of QAQ^* is equal to zero for $i > j + 1$. Explain how this transformation is an important aspect of making QR-iteration computationally efficient.

[7 marks]

(b) Let T be a symmetric tridiagonal matrix. State a decomposition of T into a matrix with two tridiagonal blocks, $\left(\begin{array}{c|c} T_1 & 0 \\ \hline 0 & T_2 \end{array}\right)$, plus a rank one matrix. State a relationship between the eigenvalues of T and those of T_1 and T_2.

[10 marks]

(c) Let $A \in \mathbb{R}^{m\times n}$ have a QR factorisation, namely $A = QR$, where Q is unitary and R is upper triangular. Construct a sequence of unitary matrices V_i such that $RV_1V_2\cdots V_{n-1} = B$ is nonzero only on its entries with index (k, k) for $k = 1, 2, \ldots, n$ and $(k, k + 1)$ for $k = 1, 2, \ldots, n - 1$. Explain how the above factorisation can be further processed to approximate the singular values and what are the computational advantages of this method for computing B.

[8 marks]

Section B — Continuous Optimization

Question 3

Consider the following unconstrained optimization problem

$$\min_{x \in \mathbb{R}^n} \ f(x), \tag{1}$$

where

$$f(x) = g^T x + \frac{1}{2} x^T H x, \tag{2}$$

for some given vector $g \in \mathbb{R}^n$ and $n \times n$ matrix H that is positive definite. Apply the steepest descent method with linesearch to (1), where the linesearch is performed on each iteration $k \geq 0$ as follows: choose the stepsize α^k as the largest stepsize value allowed by the Armijo condition.

(a) Find the minimizer(s) of f in (2) and show that f is bounded below. State the steepest descent iteration and the Armijo condition. Calculate the stepsize α^k prescribed by the given linesearch and find a uniform lower bound on α^k independent of k.

[10 marks]

(b) Show global convergence of this steepest descent variant when applied to (1), namely, show that $\nabla f(x^k) \to 0$ as $k \to \infty$, where $\nabla f(\cdot)$ is the gradient of (2), x^k are the iterates of this method and x^0 is an arbitrary starting point.

[7 marks]

(c) Let $\epsilon \in (0, 1]$. Possibly using relations derived in (b), provide an upper bound as a function of the accuracy ϵ on the number of iterations until the first iteration i such that $\|\nabla f(x^i)\| \leq \epsilon$.

[3 marks]

(d) Find a linear change of variables $y = Ax$ where $y \in \mathbb{R}^n$, such that steepest descent with exact linesearches applied to the scaled form of (1) will terminate at the minimizer in one iteration. How many iterations will the Armijo linesearch variant take on the scaled function to reach within $\epsilon \in (0, 1]$ of the scaled minimizer?

[5 marks]

Question 4

(a) Consider the unconstrained optimization problem

$$\min_{x \in \mathbb{R}^n} \; f(x), \tag{3}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable. Let $\nabla f(\cdot)$ and $\nabla^2 f(\cdot)$ denote the gradient and the matrix of second-derivatives (ie, the Hessian) of f, respectively. Let x^* be a stationary point of f such that $\nabla^2 f(x^*)$ is nonsingular.

Apply Newton's method for optimization to (3), without linesearch. State the conditions required on f and its derivatives and on some Newton iterate such that Newton's method converges, with quadratic rate, to x^*. Prove this result (of local convergence of Newton's method).

[11 marks]

Hint: you may use in your proof the following property: when $\nabla^2 f$ is Lipschitz continuous in a neighbourhood $\mathcal{N}$ of some z^*, with Lipschitz constant L_*, then

$$\nabla f(z^*) = \nabla f(x) + \nabla^2 f(x)(z^* - x) + E(z^* - x), \text{ with } \|E(z^* - x)\| \leq \frac{3}{2} L_* \|z^* - x\|^2,$$

for all $x \in \mathcal{N}$. (No need to prove this property.)

(b) Consider the following function of one variable x,

$$f(x) = -\frac{x^6}{6} + \frac{x^4}{4} + 2x^2. \tag{4}$$

Calculate the stationary points of this problem and establish whether they are (local) minimizers or maximizers. Estimate the local rate(s) of convergence of Newton's method for optimization (without linesearch) applied to f, when the starting point is close to each of the stationary points of f that you found.

[9 marks]

(c) Find a starting point x^0 for Newton's method for optimization (without linesearch) applied to f in (4) such that the iterates satisfy $x^{k+1} = -x^k$ for all k and establish whether the limits points of these iterates are stationary points of (4).

[3 marks]

(d) What conclusions can you draw, based on the above example or otherwise, about the global convergence, from an arbitrary starting point, of Newton's method? Briefly describe one modification of Newton's method that attempts to improve the global behaviour of the method.

[2 marks]

Question 5

(a) Consider the following constrained optimization problem

$$\min_{x\in\mathbb{R}^n} f(x) \quad \text{subject to } Ax \geq b, \tag{5}$$

where f is twice continuously differentiable, A is an $m \times n$ matrix and $b \in \mathbb{R}^m$. State the KKT conditions for this problem. Is any local minimizer of (5) a KKT point? State the second-order necessary optimality conditions satisfied at a local minimizer.

[8 marks]

(b) Consider the quadratic programming problem

$$\min_{x\in\mathbb{R}^n} q(x) = g^T x + \frac{1}{2}x^T Hx \quad \text{subject to} \quad x \geq 0, \tag{6}$$

where H is a negative definite $n \times n$ matrix and $g \in \mathbb{R}^n$.

Write down the KKT conditions at a local minimizer x^* of (6). Show that if any component of a feasible point x of problem (6) is positive, then this point x is not a local minimizer of (6).

[8 marks]

(c) Show that problem (6) has no global minimizer.

[3 marks]

(d) Assume now that the same objective $q(x)$ in (6) is minimized over the unit hypercube, namely the set of all x such that $0 \leq x_i \leq 1$, for all $i = 1, \ldots, n$. Deduce that the global minimizer of this problem occurs at one of its vertices.

[6 marks]

Question 6

Consider the equality-constrained optimization problem,

$$\min_{x\in\mathbb{R}^n} f(x) \quad \text{subject to} \quad c(x) = 0, \tag{7}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $c : \mathbb{R}^n \to \mathbb{R}^m$ with $c(x) = (c_1(x), \ldots, c_m(x))^T$ are twice continuously differentiable functions.

(a) Write down the augmented Lagrangian function associated with (7) and describe its connection to the solution of problem (7).

[4 marks]

(b) State and prove the theorem of global convergence for the augmented Lagrangian method.

[11 marks]

(c) Using information from the theorem in part (b), briefly outline the basic steps of the augmented Lagrangian method applied to (7). Comment on this method's advantages in connection to the disadvantages of the quadratic penalty method.

[5 marks]

(d) In the augmented Lagrangian function, instead of a quadratic penalty term, consider using a quartic penalty term of the form $\|c(x)\|^4/(4\sigma)$. Does the global convergence result hold? Explain your reasoning.

[5 marks]