



Exam in : **MAT-2201**

Date of exam : **June 8, 2007**

Time : **0900 - 1300**

Place : **Åsgårdveien 9.**

Aids: : **Rottmann's tables; 2 A4-sheets with
handwritten notes on both sides;
approved calculator.**

The set of exam problems is 6 pages including the front page

Contact person during exam: Einar Mjølhus, tlf. 776 44013.

Problem 1.

In a computer which we shall call Marc-32, floating point numbers in single precision are represented by 32 bits, where 1 bit contains the sign, 8 bits contain the exponent and 23 bits contain the mantissa. The mantissa has the form $(1.b_1\dots b_{23})_2$. For any real number r , $fl(r)$ denotes the floating point number which is nearest to r .

a) What is the smallest possible *positive* value of

$$g(x) = fl(x) - 2, \quad x \in \mathbf{R}?$$

b) An industrial mathematician, Marit, writes a program for Marc-32, where she has to compute

$$(1.1) \quad y = \sqrt{4 + x^2} - 2.$$

Unfortunately, Marit is not well educated, and she discovers to her surprise that the computed values of y become 0 for small values of x . She is, however, not helpless, and she finds, after some thought, what to do. Determine a value $r_c > 0$ such that (1.1), programmed as it stands, will give $y = 0$ for $|x| < r_c$. Propose also an alternative way of computation which will give Marit good approximations to y also for $|x| < r_c$.

c) Marit has to compute the derivative of a function f in her program. She then intends to use the formula

$$Df(x, h) = \frac{f(x+h) - f(x-h)}{2h}$$

in order to approximate $f'(x)$. She has reasons to assume that f has derivatives to all orders. Derive an expression for the truncation error

$$E(f, x, h) = f'(x) - Df(x, h).$$

d) According to the result from c), one would think that the error can be made as small as one wants, by choosing h sufficiently small. But Marit has made her experiences, and she knows that such is not the case. Explain why. Derive also for what value of h (order of magnitude) one obtains the most accurate result, and estimate the error in that case (order of magnitude) .

Problem 2.

a) Determine a LU factorization of the matrix

$$A = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 6 \end{bmatrix}$$

b) Solve the system of equations

$$Ax = b$$

where

$$b = \begin{bmatrix} 12 \\ 22 \\ 28 \end{bmatrix}$$

by means of the obtained LU factorization of A .

c) Determine a QR factorization of A , i. e.,

$$A = QR = [q_1 \mid q_2 \mid q_3] \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix}$$

where the columns of Q are orthonormal vectors.

Problem 3.

Let $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, $x \in \mathbf{R}^n$ where m may be larger than n .

a) Explain what is meant by a least squares solution to

$$Ax = b.$$

b) Let a least squares solution be $\tilde{x}$. Show that it must satisfy

$$A^T A \tilde{x} = A^T b,$$

where A^T denotes the transpose of A .

c) Show that the least squares solution can be computed by from

$$R\tilde{x} = Q^T b$$

where $A = QR$ is a QR factorization of A . Comment on the advantages of this.

Problem 4.

a) Find a second-degree polynomial which interpolates $\sin x$ in the points $x = 0, \pi/2$, og π . Explain why there can be only one such polynomial.

b) We wish to derive a formula for numerical approximation to the integral

$$(4.1) \quad \int_{-h}^h f(x) dx$$

by replacing f by a second-degree polynomial which interpolates f in the points $-h$, 0 and h . Write up this polynomial $P_2(x)$, and derive an approximate formula for the integration (4.1) by computing

$$(4.2) \quad \int_{-h}^h P_2(x) dx$$

What is this formula called?

c) Simpson's formula for approximating (4.1) is

$$\int_{-h}^h f(x) dx = \frac{1}{3}h[f(-h) + 4f(0) + f(h)] + E$$

where E is the local error. What is meant by saying that a formula of integration has degree of precision p ? What is the degree of precision of Simpson's formula. Give argument for your answer.

d) Statement: If a formula of integration has degree of precision p , then the local error satisfies

$$|E| \leq Ch^{p+2}$$

for a constant C which is independent of h when $h \rightarrow 0$. Prove that this is correct for Simpson's formula, under the assumption that the function f is differentiable a sufficient number of times. The proof can be based upon Taylor's formula.