Practice and Instructions for the Final Exam

Information and rules for the final

- The final exam is 6:00 to 7:50 on Thursday, December 18.
- No electronic devices may be visible during the exam.
- The exam is "closed book". No materials are allowed except for one "cheat sheet", which is one standard piece of paper that you may prepare in advance with any information you want.
- Please explain all answers with a few words or sentences. You may lose points for a formula or answer with no explanation.
- You will get credit for questions left blank according to the schedule
  - 75% for a True/False question
  - 50% for a multiplie choice question
  - 25% a full answer question.

  Points may be deducted for wrong answers, which can bring the credit to zero.
- Please cross out any answers you think are wrong. You may lose points for wrong answers even if you also give the right answer.
- Write answers in an exam book ("blue book").

See also practice final exam questions from recent years.

True/False

In each case, say whether the statement is true or false and give few words or sentences to justify your answer. Use the mathematician's understanding that if a statement has a counter-example then it is false, even if it's usually true. For questions that involve judgement, give the answer you would give to a client if you were a Scientific Computing Consultant.

1. Suppose we generate random numbers U using $U = S/n$, where S is an integer in the range $0 \leq S \leq n$ and n is a 1024 bit integer. Then U could be any double precision floating point number in the range $0 \leq U \leq 1$. Assume the operation S/n is done according to the rule "exact answer correctly rounded".
2. Suppose $f(x)$ has $f(0) = 0$ and $f'(x) > 5$ for all $x \geq 0$. Suppose $0 \leq x \leq 1$ and $0 \leq y \leq 1$. Then $f(y) = f(x)$ only if $y = x$ in double precision floating point.
3. If we start a Monte Carlo calculation by initializing the random number generator with one of the two (identical except that the seeds are slightly different):

   ```
   rng_1 = np.random.default_rng(seed = 100000)
   rng_2 = np.random.default_rng(seed = 100001)
   ```

   The codes produce answers $\widehat{A}_1$ and $\widehat{A}_2$, which are unbiased Monte Carlo approximations to the true answer A (same algorithm, different seed). Then the statistical errors $\widehat{A}_1 - A$ and $\widehat{A}_2 - A$ are close to each other.

4. The routine trap(f,n) applies the trapezoid rule with n uniformly spaced points and returns

$$\widehat{A}_n \approx A = \int_0^1 f(x)\,dx \;.$$

Suppose $f(x)$ is a smooth function of $x \in [0,1]$ (including endpoints) The following code will eventually find an n with $\left|\widehat{A}_n - \widehat{A}_{2n}\right| < \epsilon$:

```
n   = 10      # starting guess
eps = .1      # desired accuracy
while(1):
   A_1 = trap(f,n)
   A_2 = trap(f,2*n)
   if ( np.abs( A_2-A_1 ) < eps ):
      break
   n = 2*n
```

The integrand is a smooth function of x. Monte Carlo is likely to be the cheapest (least computer time, fewest evaluations of f) way to estimate A accurately.

5. Let $V(x)$ be a smooth function of d variables. Let $H(x)$ be the $d\times d$ hessian matrix of V at x Then $p = -H(x)\nabla V(x)$ is a descent direction.

6. If a mathematical problem to compute some number A that depends on a parameter x, and the condition number is $\kappa = 10^7$, then there may still be a very stable algorithm that computes the answer to 1% relative accuracy in single precision arithmetic..

7. If $A(h)$ is a second order accurate approximation to A, then $A(h)$ has an asymptotic error expansion of the form $A(h) \sim A + A_2 h^2 + \cdots$.

8. If $|x-y| < 10^{-6}$ then x is close to y.

Full answer

Warning. These questions may involve more computation or longer answers than would be practical on the actual exam. Please do not panic at the difficulty of these questions. These questions do not cover every topic in the class. Please look at sample exams from other years also.

1. Let $x(t)$ be a trajectory of an ODE system $\dot{x} = f(x)$ with $x(0) = x_0$ and T given. Give an algorithm to find

$$A(T) = \int_0^T x_1(t)\,dt.$$

The error should be fourth order in a time step parameter Δt. Here, $x(t) = (x_1(t), \cdots, x_d(t))$.

2. Suppose $A(n)$ is the midpoint rule approximation

$$I = \int_0^1 f(x)\,dx$$

$$\widehat{I}_n = \Delta x \sum_{k=0}^{n} f(x_{k+\frac{1}{2}})$$

Use the usual definitions $x_k k = k\Delta x$ and $\Delta x = 1/n$. Find a way to combine $\widehat{I}_n$ and $\widehat{I}_{n+1}$ to make an approximation to I that is more than second order accurate. *Hint.* Use the asymptotic error expansion for the midpoint rule and the approximation

$$\begin{aligned}
\frac{1}{(n+1)^2} &= \frac{1}{n^2+2n+1} \\
&= \frac{1}{n^2}\,\frac{1}{1+\frac{2}{n}+\frac{1}{n^2}} \\
&\sim \frac{1}{n^2}\left[1-\left(\frac{2}{n}+\frac{1}{n^2}\right)+\left(\frac{2}{n}+\frac{1}{n^2}\right)^2-\cdots\right] .
\end{aligned}$$

You need only a small part of this asymptotic expansion of $(n+1)^{-2}$ to answer the question.

3. Suppose $A(x)$ is an $n \times n$ matrix with entries that depend on x in a smooth way. Suppose that the eigenvalues of $A(x)$ are real and distinct, and that $\lambda_1(x_0)$ is close to zero. Describe an algorithm that uses eigenvalue/eigenvector perturbation theory together with Newton's method to find x_* with $\lambda_1(x_*) = 0$.

4. Suppose $X + \Delta X$ is an $n \times d$ data matrix and that $\sigma_1(X + \Delta X)$ is the largest singular value of X. Suppose ΔX is very small relative to X and that the entries of ΔX are independent mean zero Gaussian random variables with variance $s^2 \ll 1$. Use SVD perturbation theory to find the approximate distribution of σ_1.

5. Describe an algorithm that uses one or more matrix factorizations to solve the following regularized linear least squares problem. Assume W is symmetric and positive definite (a "weight" matrix)
$$\min \|Ax - b\|_2^2 + \lambda x^T W x \ .$$

6. Consider the iterative minimization algorithm with "momentum" variable p and iterates x_n and p_n given by
$$\begin{aligned}
p_{n+1} &= (1-t)p_n + t\nabla V(x_n) \\
x_{n+1} &= x_n - s p_n
\end{aligned}$$
Suppose t and s are positive. Give the convergence rate of this iteration if $V(x) = \frac{1}{2}x^T A x$.

7. Suppose $f(x)$ is a probability density that is expensive to evaluate and that $g(x)$ is another PDF that is easy to evaluate and sample. Suppose the acceptance probability for sampling f using proposals from g is $p \approx 10\%$. Suppose we want to know
$$A = \mathrm{E}_f[\, u(X)\,] \ .$$
Is it better (more accurate estimates of A for the same number of f evaluations) to use samples $X_k \sim f$ or samples $X_k \sim g$ and importance sampling? You can't know in advance with the information given, but you can tell after you do the importance sampling strategy. How?