MTH 373/573: Scientific Computing	Solutions to Midsem Exam	IIIT-Delhi, Monsoon 2022 Kaushik Kalyanaraman

1 Multiple Choice Questions with Justification

Problem 1: Well Posedness — 5 points

What property of a problem is not needed for it to be *well posed*?

(A) It has to have a solution.
(B) The solution should be unique.
(C) The unique solution should continuously depend on the input data.
(D) The problem should be well conditioned.

(D) A well posed problem is a one for which a solution exists, is unique and continuously depends on its inputs. Such a problem may, however, be poorly conditioned and so, conditioning is not part of the definition for well posedness.

Problem 2: Stable Algorithm — 5 points

Which of the following always produces an *accurate* result?

(A) Applying a stable algorithm to a well conditioned problem
(B) Applying a stable algorithm to an ill conditioned problem
(C) Applying an unstable algorithm to a well conditioned problem
(D) None of the above

(A) A stable algorithm when applied to an ill conditioned problem or an unstable algorithm applied to a well conditioned problem are both not guaranteed to provide an accurate solution. However, a stable algorithm applied to a well conditioned problem does generate an accurate solution.

Problem 3: Cancellation — 5 points

Floating point *cancellation* occurs when ...

(A) you add two numbers of the same sign and nearly the same magnitude.
(B) a result is so small that it can only be represented using a denormal floating point number.
(C) you subtract two numbers of the same sign and nearly the same magnitude.
(D) you subtract a small number from a big number, affecting only a few trailing digits of the big number.

(C) Cancellation in a floating point number system occurs when two similar magnitude numbers but with opposite signs are added together. In such an instance, the leading significant digits are *cancelled* in the result leading to an overall loss in accuracy of the represented solution. For example, $1.124496 - 1.121337$ in a system with 4 bits of precision will be computed as $1.124 \times 10^0 - 1.121 \times 10^0 = 3.xyz \times 10^{-3}$ where the x, y and z are either unknown or just wrong. for example, each $x = y = z = 0$.

Problem 4: Matrix Condition Number 1 — 5 points

Let M be a square, real-valued matrix. Consider the following two statements:

(1) $\|M\|_p$ is small.
(2) The condition number $\kappa_p(M)$ with respect to the matrix norm $\|\cdot\|_p$ is small.

How are (1) and (2) related? Choose the best answer.

(A) (1) $\implies$ (2) — If (1), then (2)
(B) (1) $\impliedby$ (2) — If (2), then (1)
(C) (1) $\iff$ (2) — (1) if and only if (2)
(D) (1) and (2) are not related to each other

(D) Consider the diagonal matrix $\begin{bmatrix} 1 & 0 \\ 0 & 10^{-20} \end{bmatrix}$ whose norm is 1 but condition number is 10^{20} since the inverse matrix has a norm of 10^{20}. Thus, (1) $\not\Longrightarrow$ (2).
Next, consider the diagonal matrix $\begin{bmatrix} 10^{20} & 0 \\ 0 & 10^{20} \end{bmatrix}$. It's norm is large at 10^{20} but the norm of its inverse is 10^{-20} ensuring that the matrix is perfectly conditioned with a condition number of 1. Thus, (2) $\not\Longrightarrow$ (1).s

Problem 5: Matrix Condition Number 2 — 5 points

Let $A \in \mathbb{R}^{n \times n}$ be a perfectly conditioned matrix, that is, it has $\kappa_\infty(A) = 1$. Which other matrix among the following matrices will necessarily share this property of ∞-norm condition number being 1.

(A) AB where $B \in \mathbb{R}^{n \times n}$ is any nonsingular matrix
(B) A^{-1}, the inverse of A
(C) DA where D is a nonsingular diagonal matrix
(D) A^T, the tranpose of A

(B). For an arbitrary nonsingular (including diagonal) matrix B, $\kappa_\infty(AB) = \kappa_\infty(A)\kappa_\infty(B)$ or $\kappa_\infty(BA) = \kappa_\infty(B)\kappa_\infty(A)$ where $\kappa_\infty(B)$ need not equal 1. For the transpose of a matrix, $\|A^T\|_\infty = \|A\|_1$ and the 1-norm and ∞-norm of a matrix need not be equal. Thus, only for A^{-1} and readily by definition $\kappa_\infty(A^{-1}) = \kappa_\infty(A) = 1$.

Problem 6: Normal Equations Matrix — 5 points

Let $A \in \mathbb{R}^{m \times n}$ with $m \geq n$. Then, the matrix $A^T A$ is always:

(A) nonsingular (B) symmetric (C) positive definite (D) nonzero

(B). If A is not full rank, then $A^T A$ is singular and only positive semidefinite. If A is the zero matrix, $A^T A$ will also be the zero matrix. Clearly though $\left(A^T A\right)^T = A^T A$.

Problem 7: Orthonormal Matrix — 5 points

Let $Q \in \mathbb{R}^{m \times n}$, $m \geq n$ be a matrix with orthonormal columns. Consider the following two statements:

(1) $Q^T Q = I$, (2) $QQ^T = I$,

where I is an appropriately sized identity matrix. Which of the following options correctly characterizes these two statements?

(A) Only (1) is True (B) Only (2) is True
(C) Both (1) and (2) are True (D) Both (1) and (2) are False

(A). Consider $Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$. Then, $Q^T Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $QQ^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Problem 8: Gaussian Elimination with Partial Pivoting — 5 points

Using Gaussian Elimination with partial pivoting to solve the following matrix problem, what is the first pivot element?

$$\begin{array}{r} \\ \text{Row 1} \\ \text{Row 2} \\ \text{Row 3} \\ \text{Row 4} \end{array} \begin{array}{c} \begin{array}{cccc} \text{Column 1} & \text{Column 2} & \text{Column 3} & \text{Column 4} \end{array} \\ \begin{bmatrix} 100 & 10 & 20 & -1000 \\ 1 & 2 & -2 & 1 \\ 10 & 20 & -400 & 100 \\ -200 & 10 & 50 & 1000 \end{bmatrix} \end{array} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

(A) Row 1, Column 1 (B) Row 1, Column 4
(C) Row 4, Column 1 (D) Row 4, Column 4

(C). By definition, in Gaussian elimination with partial pivoting, we pick the pivot to be the largest magnitude entry in the sub column which is being processed. Since this is the first column, we pick the largest magnitude entry in this column.

Problem 9: Householder QR Factorization 5 points

Suppose that you are computing the QR factorization of the matrix:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix},$$

by Householder transformations. How many Householder transformations are required?

(A) 2 (B) 1 (C) 3 (D) Cannot be determined

(C). To convert this matrix into an upper triangular matrix of size 4×3 using Householder transformations, we need one such transformation for each column of the matrix.

Problem 10: Backward Error 5 points

Consider the following approximation of the square root $\sqrt{1-x} \approx 1 - \frac{1}{2}x$. What is the magnitude of the *absolute backward error* from using this approximation when $x = 1/2$?

(A) $3/4 - 1/\sqrt{2}$ (B) $7/16$ (C) $1/16$ (D) $3/4$

(C). Using the approximation, for $x = 1/2$, $1 - x/2 = 1 - 1/4 = 3/4$. Now, we seek an $\hat{x}$ such that $\sqrt{1-\hat{x}} = 3/4 \implies \hat{x} = 7/16$. Thus, the absolute backward error is $|x - \hat{x}| = |1/2 - 7/16| = 1/16$.

Problem 11: Gaussian Elimination Backward Stability 5 points

Suppose you are solving an $n \times n$ nonsingular linear system $Ax = b$ by Gaussian elimination with partial pivoting. If the matrix A is highly ill conditioned, what are the likely consequences?

(A) The error in the computed solution will be relatively large.
(B) The solution will be relatively insensitive to changes in the data.
(C) The residual will be relatively large.
(D) All of the above.

(A). Gaussian elimination with partial pivoting is a (backward) stable algorithm. Since the given problem is stated to be ill conditioned, the application of a stable algorithm to an ill conditioned problem will lead to a large error. Note that it can also be argued that for a stable algorithm, the residual will by definition be small due to errors not propagating but this residual will get magnified by the condition number to provide a large error in the solution.

Problem 12: Linear Least Squares Solution 5 points

Which of the following statements about a linear least squares problem, $Ax \simeq b$, is *False*?

(A) If A is not full-rank, there is still a linear least-squares solution but it is not unique.
(B) The columns of A are orthogonal to the linear least-squares solution residual.
(C) If b is not in the space spanned by the columns of A, there is no solution.
(D) If b is exactly orthogonal to the span of columns of A, then the only solution will be the trivial solution.

(C). A linear least squares problem will always admit a solution, and the solution may be unique or there can be infinitely many solutions depending on the rank of the least squares system matrix. Even if the right hand side vector is exactly orthogonal to the subspace spanned by the columns of the least squares system matrix, the solution will be unique (and trivial) or infinitely many (and in the kernel of the system matrix). In now case will the problem have *no solution*.

Problem 13: Gram-Schmidt Orthogonalization — 5 points

Carry out one step of the classical Gram-Schmidt procedure on the following two vectors:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \qquad u_2 = \begin{bmatrix} 3 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

Let q_1 and q_2 be the resulting vectors. Which of the following choices correctly provides these two vectors?

(A) $$q_1 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad q_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 2 \\ -1 \\ 0 \\ -1 \end{bmatrix}.$$

(B) $$q_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad q_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 2 \\ -1 \\ 0 \\ 1 \end{bmatrix}.$$

(C) $$q_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad q_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 2 \\ -1 \\ 0 \\ -1 \end{bmatrix}.$$

(D) $$q_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad q_2 = \begin{bmatrix} 2 \\ -1 \\ 0 \\ -1 \end{bmatrix}.$$

(C). Since $\|u_1\|_2 = \sqrt{3}$, normalizing u_1 we get q_1 as in either options (B) or (C). So, we can compute $r_{21} := \langle u_2, q_1 \rangle$ to be $5/\sqrt{3}$, and computing $u_2 - r_{21} q_1$, we get the required solution as in option (C).

Problem 14: Floating Point Machine Epsilon — 5 points

If ε_{M} is the machine epsilon in the IEEE double precision floating point system, then $(2 + \varepsilon_{\mathrm{M}}) - 2$ is equal to:

(A) $(1/2)\varepsilon_{\mathrm{M}}$ (B) ε_{M} (C) 0 (D) $2\varepsilon_{\mathrm{M}}$

(C). By definition, $\varepsilon_{\mathrm{M}} := \inf_{\varepsilon} \mathrm{fl}(1+\varepsilon) > \mathrm{fl}(1)$. In the IEEE double precision floating point number system where rounding to nearest (with ties broken by rounding to even), we have for ε_{M} that $\mathrm{fl}(1+\varepsilon) = \mathrm{fl}(1)$. Hence $\mathrm{fl}(2+2\varepsilon) = \mathrm{fl}(2)$.

Problem 15: Accuracy of Gaussian Elimination 5 points

Suppose that the condition number of a nonsingular, real-valued square matrix A is 10^{12}, and A and a right hand side vector b are stored in IEEE double precision floating point system. How many correct digits can we expect for the solution to this linear system using Gaussian elimination with partial pivoting?

(A) 4 (B) 8 (C) 12 (D) 16

(A). For the IEEE double precision floating point number system, in the decimal equivalent, the precision is 17. That is, there are 16 digits after the decimal place. Equivalently, $\varepsilon_{\mathrm{M}} \approx 10^{-16}$. Thus, the solution of the linear system will have an accuracy that is given by the condition number times the machine epsilon or approximately $10^{-16+12} = 10^{-4}$. In other words, there will be about 4 correct digits in any computed solution.