Information and Instructions

- Open-source and citation. This is an open-source exam, on the condition that the sources be properly cited. Proper citation reflects critical selection and use of existing knowledge and eliminates intentional or unintentional plagiarism.

  You may bring in books and notes. The citation includes the page numbers as well as the book titles and authors. You may also search on the Internet, cite the exact web addresses.

- No-live communication is allowed with anyone else, in person or via any electronic media.

- Problem set and grading. Each individual selects **4** out of the 5 problems as the primary ones for the exam grading. Please mark with X the numbers for the 4 problems you choose as the primary.

  (1) ______ (2) ______ (3) ______ (4) ______ (5) ______

  The additional problem, if you wish to submit as well, will not get a score higher than the average of the primary ones.

- Concise answers and arguments are preferred; irrelevant and incorrect comments are subject to point deduction.

- The grading process is blind to individual names. Please mark every page of your answers with the ID provided to you.

- Upon completion, please submit your answers and return the problem description to the DGS office.

- If you decide not to continue and finish the exam, please inform the DGS office so, on the returned problem set.

Confirmation:
I have read the above instructions on Jan. ______, 2021.

PLEASE READ EACH PROBLEM CAREFULLY: Unless a proof or disproof is clearly requested, some of the problems ask for factual statements or valid conditions as answers, which are supported by credible references, or direct proofs, or both. Statements made without any support are considered as incomplete answers. If any description is confusing, do not hesitate to ask the proctor to help clarify.

1. `Elementary operations with large sparse or structured matrices.`

   By arithmetic complexity for a computational procedure in numerical computation, we mean the total number of arithmetic operations used in the procedure.

   The complexity for a matrix computation procedure is often described in terms of matrix sizes and sparsity. Specifically, for matrix B, the sizes of B are the number of rows and the number of columns in B, and $\mathrm{nnz}(B)$ is the number of nonzero elements.

   In the following subproblems, A is an $m \times n$ matrix, $m, n > 0$ are large. We are concerned with computational efficiency in computing the matrix-vector product

   $$y := A\,x$$

   with $x \neq 0$ as a single column vector, or the matrix-matrix product

   $$Y := A\,X$$

   with X consists of q nonzero column vectors. (One may imagine that A represents a bipartite network or query table, x represents a single query vector, and X as multiple query vectors, if this helps.)

   (a) Assume that A is sparse, $\mathrm{nnz}(A) < m\,n$. Estimate the worst-case arithmetic complexity of computing matrix-vector product y.

   (b) Assume that A has a readily available rank-k factorization, $A = BC$, k is much smaller than $\min\{m, n\}$. Describe an economic way of computing y via utilizing the low-rank factorization and estimate the worst-case arithmetic complexity.

   (c) Assume that A can be split into two terms, $A = A_1 + A_2$, A_1 is sparse and A_2 is of low rank. Describe the conditions under which computing y via the split is more economic.

   (d) Assume that both A and X are sparse, $\mathrm{nnz}(A) < m\,n$ and $\mathrm{nnz}(A) < n\,q$. Estimate the worst-case complexity of arithmetic complexity of computing the product Y.

   (e) ***Optional.*** Give an analysis on the average-case complexity of computing Y, with a clear description of the probabilistic model assumed for the analysis.

2. `Numerical solution to a system of linear equations.`

   Let matrix P be column-wise stochastic. Let α be a positive scalar $0 < \alpha < 1$. Consider the solution to the personalized or content-specific page-ranking equations

   $$(I - \alpha P)x = (1 - \alpha)v, \tag{1}$$

   where I is the identity matrix, v is a content-specific distribution, i.e., $v \neq 0$ and $\mathbf{1}^{\intercal} v = 1$.

   (a) Verify that under the given conditions, the solution to the ranking equations exists uniquely, and it is a distribution. (The solution is referred to as the page-rank distribution).

   (b) Prove or disapprove that the problem (1) is equivalent to the following minimal-residual problem,

   $$x = \arg\min_{x'} \|(1 - \alpha P)x' - (1 - \alpha)v\|_p^2, \tag{2}$$

   where $\|b\|_p$ is the p-norm of vector b , $p \geq 1$.

3. `Numerical solution to a system of non-linear equations.`

Consider numerical solution to a system of n non-linear equations for n unknown variables, $n > 5$.

$$F(x) = 0, \quad F : \mathbb{R}^n \to \mathbb{R}^n \tag{3}$$

(a) Give one or more than one reasons why iterative methods are used for numerical solution to the problem (3) in general.

(b) Give a condition under which problem (3) is equivalent to the following problem

$$x = x + C(x)F(x) \tag{4}$$

(c) Assume that the system (4) is equivalent to the system (3). Consider the following iterative process,

$$x_{k+1} = x_k + C(x_k)F(x_k), \quad k = 0, 1, 2, \cdots \tag{5}$$

starting with an initial vector x_0.

Give a sufficient condition under which the iteration converges theoretically,

Practically, an iteration method must be equipped with termination criteria. Give a brief explanation.

(d) ***Optional.*** Make a specific connection of the generic iteration scheme (5) to the Newton's iteration method, describe the condition(s) for Newton's method to converge.

4. `Iterative solutions to optimization problems.`

The gradient descent (GD) method is used frequently for solving various minimization problems in Machine Learning (ML) with a large vector of variables to be determined, learned or predicted.

(a) Describe briefly the GD procedure. The description must include termination criteria for practical use.

(b) Give some conditions under which the GD procedure converges and converges to the global minimum.

(c) Give an example that the GD procedure converges to a local minimum, far greater than the global one.

(d) ***Optional***. Give an example that the GD procedure fails to converge.

5. `Numerical solution to a dynamic system.`

A simple epidemic model is in the form of an ordinary differential equation (ODE)

$$\frac{dy(t)}{dt} = \beta y(t)(N - y(t)), \tag{6}$$

provided with the initial condition (I.C.) at time 0, at which data $y(0)$ is available, where N is the number of hosts in the population under consideration, each host is either in the infectious state or in the susceptible state, $y(t)$ is the number of infectious hosts at time t, and β is the pairwise infection rate. Thus, $N - y(t)$ is the number of susceptible hosts at time t.

(a) Describe the forward Euler (FE) method for prediction of $y(t)$ at time $t > 0$.
Find a condition under which the FE procedure under-estimates or over-estimate the model solution.

(b) Describe the backward Euler (BE) method for prediction of $y(t)$ at time $t > 0$.
Find a condition under which the FE procedure under-estimates or over-estimate the model solution.

(c) ***Optional.*** Find a condition under which $y(t)$ increases over a period of time.

Reference. *Epidemic Modeling: An Introduction*, by D. J. Daley and J. Gani, Cambridge University Press, 1999