Indian Institute of Science Bangalore
Department of Computational and Data Sciences (CDS)
DS284: Numerical Linear Algebra
Assignment 5 [Posted Oct 29, 2024]
Faculty Instructor: Dr. Phani Motamarri
TAs: Gourab Panigrahi, Srinibas Nandi, Nihar Shah, Rushikesh Pawar, Surya Neelakandan

Solution to first 5 problems need to be submitted by Nov 17, 2024

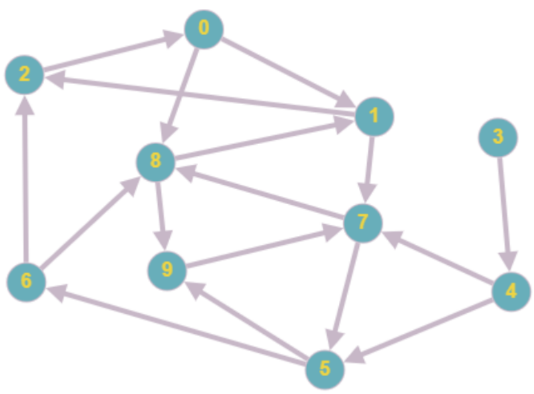
Problem 1

This question will help you to appreciate the usefulness of power iteration in a real-world application. In this question, you need to calculate the Page rank of each of the nodes of the graph given below by first constructing its Markov transition matrix and subsequently, calculating the most dominant eigenvector of this transition matrix i.e the eigenvector corresponding to the largest eigenvalue. Use the power iteration algorithm discussed in the class to compute the dominant eigenvector by implementing it on a computer.

Justify your result showing the following data/figures:

- Plot of the 2-norm of the residual of eigenvalue problem involving the dominant eigenvector with iteration number. (Use the unit vector generated at the end of each iteration in your power iteration algorithm to compute the eigenvalue problem residual at each iteration)
- Plot of the 2-norm of the difference between the vectors corresponding to the successive iterates of your power iteration with number of iterations.
- Compute the Rayleigh quotient at each iteration of your algorithm and plot the convergence with respect to iteration number.
- State the node numbers with least and highest page ranks.

Refer to the write-up below for description about building the Markov Transition matrix of a given graph.



Problem 2

Assert if the following statements are True or False. Give a detailed reasoning for your assertion.

(a) An eigenvalue solver can be designed to compute eigenvalues and eigenvectors of a given matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$ in a finite number of steps, using exact arithmetic.

(b) An eigenvalue solver designed to compute all eigenvalues and eigenvectors of a symmetric dense matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$ requires at most $O(m^3)$ work, if it is not initially reduced to tri-diagonal form in Phase 1.

(c) Power iteration produces a sequence of vectors $\mathbf{v}^{(i)}$ that converges to the eigenvector corresponding to the largest eigenvalue of $\mathbf{A} \in \mathbb{R}^{m \times m}$ starting with any initial guess vector $\mathbf{v}^{(0)} \neq \mathbf{0}$.

(d) Let $\mathbf{F} \in \mathbb{R}^{m \times m}$ denote the Householder reflector that introduces zeros below the diagonal entry of symmetric matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$ in the 1st column when pre-multiplied with $\mathbf{A}$. Then eigenvalues of $\mathbf{F}\mathbf{A}\mathbf{F}^T$ and $\mathbf{A}$ are the same.

Problem 3

We are usually confronted with large sparse matrix eigenvalue problems arising from the discretization of a partial differential equation (eigenproblem), where the unknown eigenfunctions are approximated in a finite-dimensional subspace as a linear combination of localized basis functions spanning the subspace. Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times m}$ be two such large symmetric positive definite sparse matrices and the problem of finding eigenvector, eigenvalue pairs $(\lambda_i, \mathbf{u}_i)$ satisfying the equation

$$\mathbf{A}\mathbf{u}_i = \lambda_i \mathbf{B}\mathbf{u}_i \tag{1}$$

is called a *generalized* eigenvalue problem (Note: The standard eigenvalue problem you are familiar with is a special case having $\mathbf{B} = \mathbf{I}$).

(a) Rewrite the generalized eigenvalue problem in equation (1) as a standard eigenvalue problem $\mathbf{H}\mathbf{v}_i = \lambda_i \mathbf{v}_i$ and in doing so, ensure that $\mathbf{H}$ is a symmetric matrix. Express $\mathbf{u}_i$ in terms of $\mathbf{v}_i$. (Hint: Use the property of $\mathbf{B}$)

(b) Devise an iterative algorithm to find the eigenvalue λ_i closest to 2.0 (assume 2.0 is not an eigenvalue of equation (1)) and a corresponding eigenvector. Will your algorithm always converge, or is there any condition that needs to be satisfied for convergence?

(c) Identify the computationally dominant step in your algorithm and describe your method of choice for performing this step.

(d) How will you modify the algorithms proposed in parts (b) and (c) if you are told that the eigenvalue closest to 2.0 is, in fact, the lowest eigenvalue? What do you gain in doing so?

Problem 4

If $\mathbf{A} \in \mathbb{R}^{m \times m}$ is a symmetric invertible matrix with an exceedingly large condition number. To this end, let one eigenvalue of $\mathbf{A}$ be much smaller than others in absolute value, i.e. $|\lambda_1| <<< |\lambda_2| \leqslant |\lambda_3| ... \leqslant |\lambda_m|$ (i.e. assume $\mathbf{A}$ is an ill-conditioned matrix with $\kappa(\mathbf{A})$ on the order of ϵ_M^{-1}). Let $\mathbf{q}_1, \mathbf{q}_2, ... \mathbf{q}_m$ be the corresponding eigenvectors. Now answer the following questions:

(a) If the system of equations $\mathbf{A}\mathbf{w} = \mathbf{v}$ is solved using some backward stable algorithm for a given $\mathbf{v} \in \mathbb{R}^m$ yielding a computed vector $\widetilde{\mathbf{w}} = \mathbf{w} + \delta\mathbf{w}$, show that $\delta\mathbf{w} = -(\mathbf{A} + \delta\mathbf{A})^{-1}(\delta\mathbf{A})\mathbf{w}$ for some $\delta\mathbf{A}$ such that $\|\delta\mathbf{A}\| = O(\epsilon_M)\|\mathbf{A}\|$.

(b) Assuming that $\mathbf{v}$ is a vector with components in the directions of all eigenvectors $\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_m$ of $\mathbf{A}$, show that $\frac{\mathbf{w}}{\|\mathbf{w}\|_2} \approx \mathbf{q}_1$, where $\mathbf{w} = \mathbf{A}^{-1}\mathbf{v}$ is the exact solution of the system of equations given in (a). [**Hint**: First show that $\mathbf{w}$ is approximately in the direction of $\mathbf{q}_1$]

(c) Using the Taylor series expansion $(\mathbf{A} + \delta\mathbf{A})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}(\delta\mathbf{A})\mathbf{A}^{-1} + O(\epsilon_M^2)$, show that $\frac{\widetilde{\mathbf{w}}}{\|\widetilde{\mathbf{w}}\|_2} \approx \mathbf{q}_1$. [**Hint**: Using the expression $\delta\mathbf{w}$ derived in (a) and the fact $\delta\mathbf{A}$ is random roundoff perturbation, show that $\delta\mathbf{w}$ is in the direction of $\mathbf{q}_1$]

[This sequence of steps in (a), (b) and (c) show that, though the computed solution $\widetilde{\mathbf{w}}$ is far away from $\mathbf{w}$ for the ill-conditioned system $\mathbf{A}\mathbf{w} = \mathbf{v}$, but $\frac{\widetilde{\mathbf{w}}}{\|\widetilde{\mathbf{w}}\|_2}$ need not be far from $\frac{\mathbf{w}}{\|\mathbf{w}\|_2}$.]

Problem 5

Consider the eigenvalue problem corresponding to a large sparse and diagonally dominant symmetric matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$. Let us say, we are interested in solving $\mathbf{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i$ for $i = 1 \cdots n$ smallest eigenvalue and eigenvector pairs of $\mathbf{A}$ ($n << m$). Answering the following questions will make you deduce an iterative algorithm for solving this eigenvalue problem different from the methods discussed in the class.

(a) Let us begin the iteration $k = 0$ with a trial guess of orthogonal vectors spanning the n-dimensional subspace $\mathbb{V}_{(0)}^n = \{\widetilde{\mathbf{x}}_1^{(0)}, \widetilde{\mathbf{x}}_2^{(0)}, \ldots, \widetilde{\mathbf{x}}_n^{(0)}\}$. Assume each of these vectors act as an approximation to the corresponding exact eigenvectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ of the matrix $\mathbf{A}$. Let $\mathbf{t}_i = \mathbf{x}_i - \widetilde{\mathbf{x}}_i^{(0)}$ for $i = 1, 2, \ldots, n$ denote the correction vectors to the approximate eigenvectors. Show that $\mathbf{t}_i$ can be computed by solving the corrector equation $(\mathbf{A} - \epsilon_i \mathbf{I})\,\mathbf{t}_i = (\epsilon_i \mathbf{I} - \mathbf{A})\,\widetilde{\mathbf{x}}_i^{(0)}$, where ϵ_i is the exact eigenvalue of $\mathbf{A}$. Assuming that you know the exact eigenvalue ϵ_i, can this corrector equation be solved for $\mathbf{t}_i$?

(b) *A priori* one does not know the value of ϵ_i. In this regard what is the best approximation that can be used for ϵ_i and why? (Hint:- Think how the space $\mathbb{V}_{(0)}^n$ is constructed). If ϵ_i is approximated to be $\widetilde{\epsilon}_i^{(0)}$, then the corrector equation becomes $(\mathbf{A} - \widetilde{\epsilon}_i^{(0)}\mathbf{I})\mathbf{t}_i = (\widetilde{\epsilon}_i^{(0)}\mathbf{I} - \mathbf{A})\widetilde{\mathbf{x}}_i^{(0)}$. To solve for $\mathbf{t}_i$ efficiently, one approximates $\mathbf{A}$ in L.H.S of the above corrector equation to the matrix $\mathbf{D} = diag(\mathbf{A})$. To this end, write an expression for solution to the corrector equation (let $\widetilde{\mathbf{t}}_i$ denote this solution vector, an approximation to the correction vector $\mathbf{t}_i$).

(c) Now we construct a $2n$ dimensional space $\mathbb{V}^{2n}_{(0)} = \{\widetilde{\mathbf{x}}_1^{(0)}, \widetilde{\mathbf{t}}_1^{(0)}, \widetilde{\mathbf{x}}_2^{(0)}, \widetilde{\mathbf{t}}_2^{(0)}, \ldots, \widetilde{\mathbf{x}}_n^{(0)}, \widetilde{\mathbf{t}}_n^{(0)}\}$. Argue why is $\mathbb{V}^{2n}_{(k)}$ a better subspace than $\mathbb{V}^{n}_{(k)}$ to look for eigenvectors of $\mathbf{A}$ at any given iteration k. $(\widetilde{\mathbf{x}}_i^{(1)}, \epsilon_i^{(1)})$

(d) We now look for approximate eigenvector, eigenvalue pair of $\mathbf{A}$ in the space $\mathbb{V}^{2n}_{(0)}$. Let us denote $\widetilde{\mathbf{x}}_i^{(1)} \in \mathbb{V}^{2n}_{(0)}$ for $i = 1, 2, \ldots, n$ to be the eigenvector approximations we seek to find in $\mathbb{V}^{2n}_{(0)}$. Define the residual vector $\mathbf{r}_i = \mathbf{A}\widetilde{\mathbf{x}}_i^{(1)} - \widetilde{\epsilon}_i^{(1)}\widetilde{\mathbf{x}}_i^{(1)}$ where $\widetilde{\epsilon}_i^{(1)}$ is the best approximation to the eigenvalue corresponding to the eigenvector approximation $\widetilde{\mathbf{x}}_i^{(1)}$. This approximate eigenvector, eigenvalue pair $(\widetilde{\mathbf{x}}_i^{(1)}, \widetilde{\epsilon}_i^{(1)})$ is obtained by imposing the Galerkin condition that states that $\mathbf{r}_i$ is orthogonal to the space $\mathbb{V}^{2n}_{(0)}$. Mathematically deduce the consequences of the imposition of this Galerkin condition and subsequently elaborate how should one go about finding $(\widetilde{\mathbf{x}}_i^{(1)}, \widetilde{\epsilon}_i^{(1)})$ after imposing the Galerkin condition. Finally, $(\widetilde{\mathbf{x}}_i^{(1)}, \widetilde{\epsilon}_i^{(1)})$ forms the eigenvector, eigenvalue approximations for $k = 1$ iteration.

Note: Once $(\widetilde{\mathbf{x}}_i^{(1)}, \widetilde{\epsilon}_i^{(1)})$ is obtained from $\mathbb{V}^{2n}$, the trial subspace gets updated to $\{\widetilde{\mathbf{x}}_i^{(1)}, \widetilde{\mathbf{t}}_i^{(1)}\}$ to seek $\{\widetilde{\mathbf{x}}_i^{(2)}, \widetilde{\mathbf{t}}_i^{(2)}\}$ and iterations are continued till convergence is reached.

Problem 6

If $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 3 & 3 \end{bmatrix}$, answer the following questions:

(a) Write the characteristic equation associated with the above matrix $\mathbf{A}$ and subsequently compute its eigenvalues

(b) Set up the Arnoldi iteration with the starting vector $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ to find the orthonormal basis vectors $\{\mathbf{q}_1, \mathbf{q}_2\}$ spanning the two dimensional Krylov subspace $\mathcal{K}_2 = \langle \mathbf{b}, \mathbf{Ab} \rangle$

(c) Find the orthogonal projection $\mathbf{H}$ of $\mathbf{A}$ onto $\mathcal{K}_2$ represented in the basis $\{\mathbf{q}_1, \mathbf{q}_2\}$ obtained in (b) above, and then compute the eigenvalues of this $\mathbf{H}$ (also called Ritz values). Find the absolute error between the smallest Ritz value and the smallest eigenvalue of $\mathbf{A}$ and similarly compute the absolute error between the largest Ritz value and largest eigenvalue of $\mathbf{A}$.

(d) Consider the system of equations $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 3 & 3 \end{bmatrix}$ as given above and $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ as given in (b). Find the exact solution $\mathbf{x}^*$ which solves $\mathbf{Ax} = \mathbf{b}$ using forward substitution. Subsequently find the vector $\hat{\mathbf{x}} \in \mathcal{K}_2$ that minimizes the norm $\|\mathbf{Ac} - \mathbf{b}\|_2$ over all possible vectors $\mathbf{c} \in \mathcal{K}_2$, where $\mathcal{K}_2$ is the Krylov subspace constructed in (b). Finally, find the norm of error between exact solution $\mathbf{x}^*$ and $\hat{\mathbf{x}}$.