MATH3322 Matrix Computation
Homework 6

Due date: 17 May, Monday

1. Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ be nonsingular and symmetric. We can generalize the inverse power iteration for finding r unit eigenvectors of $\boldsymbol{A}$ corresponding to the r smallest eigenvalues in magnitude. Write out the algorithm, and explain why it works.

2. To accelerate the convergence of QR algorithm, we can also consider a *shifted* version of QR algorithm

   Choose $\boldsymbol{A}^{(0)} = (\boldsymbol{Q}^{(0)})^T \boldsymbol{A}\boldsymbol{Q}^{(0)}$ with orthogonal $\boldsymbol{Q}^{(0)} \in \mathbb{R}^{n\times n}$
   for $k = 1, 2, \dots,$
       Choose a shift μ_k
       Compute QR decompostion $\boldsymbol{A}^{(k-1)} - \mu_k \boldsymbol{I} = \boldsymbol{Q}^{(k)}\boldsymbol{R}^{(k)}$
       $\boldsymbol{A}^{(k)} = \boldsymbol{R}^{(k)}\boldsymbol{Q}^{(k)} + \mu_k \boldsymbol{I}$
   end

   With a proper choice of μ_k, the above algorithm will converge super fast. Prove that $\boldsymbol{A}^{(k)}$ is similar to $\boldsymbol{A}$.

3. Find the best rank-1 approximation in $\|\cdot\|_F$ of the matrix

$$\boldsymbol{A} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \\ 2 & 0 \\ 0 & 1 \end{bmatrix}$$

   by the following steps.

   (a) Compute eigenvalues λ_1, λ_2 with $\lambda_1 \geq \lambda_2$ and the corresponding eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2$ of the matrix $\boldsymbol{A}^T\boldsymbol{A}$.

   (b) Define $\sigma_i = \sqrt{\lambda_i}$ for $i = 1, 2$. Compute $\boldsymbol{u}_i$ by $\boldsymbol{A}\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$ for $i = 1, 2$.

   (c) Compute $\boldsymbol{A}_1 = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^T$. This is the best rank-1 approximation to $\boldsymbol{A}$ in $\|\cdot\|_F$.

4. Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ $(m \geq n)$ be a full rank matrix with SVD $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$. Compute SVDs of the following matrices in terms of $\boldsymbol{U}$, $\boldsymbol{V}$, and $\boldsymbol{\Sigma}$:

   (a) $(\boldsymbol{A}^T\boldsymbol{A})^{-1}$

   (b) $(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T$

   (c) $\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-1}$

   (d) $\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T$

5. Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ with $m \geq n$ be full rank and $\boldsymbol{A} = \boldsymbol{U\Sigma V}^T$ be its SVD. We use it to solve the least squares problem (called linear regression in statistics)

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{Ax} - \boldsymbol{b}\|_2^2. \tag{1}$$

   (a) Show that the solution of the least squares problem (1) is given by $\boldsymbol{x} = \boldsymbol{V\Sigma}^{-1}\boldsymbol{U}^T b$. (*Hint: You may use the normal equation* $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$.)

   (b) When $\boldsymbol{A}$ is very close to rank-deficient, the solution of (1) is not stable. Instead, we solve the following *Ridge Regression*

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{Ax} - \boldsymbol{b}\|_2^2 + \lambda\|\boldsymbol{x}\|_2^2, \tag{2}$$

   where $\lambda > 0$ is a parameter. Use SVD of $\boldsymbol{A}$ to give the solution to the ridge regression (2). (*Hint: Convert* (2) *it to a least squares problem.*)