MSBD5007 Optimization and Matrix Computation
Homework 2

Due date: 16 March, Sunday

1. Let $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 1 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} 6 \\ 5 \\ 7 \\ 4 \end{bmatrix}$. Use the QR factorization to solve the least square problem:

$$\min_{\boldsymbol{x} \in \mathbb{R}^4} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2.$$

2. Find the best rank-1 approximation with respect to the Frobenius norm to the matrix

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

3. Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two matrices of dimensions $m \times n$. Consider the optimization problem

$$\min_{\boldsymbol{Q} \in \mathbb{R}^{m \times m}} \|\boldsymbol{X} - \boldsymbol{Q}\boldsymbol{Y}\|_F \quad \text{subject to} \quad \boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}_m,$$

where $\|\cdot\|_F$ denotes the Frobenius norm and I_m is the $m \times m$ identity matrix. This problem arises when we want to align two datasets $\boldsymbol{X}$ and $\boldsymbol{Y}$ up to an orthogonal linear transformation. Prove that the solution is given by $\boldsymbol{Q} = \boldsymbol{U}\boldsymbol{V}^T$, where $\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$ is the singular value decomposition of $\boldsymbol{X}\boldsymbol{Y}^T$, *i.e.*, $\boldsymbol{X}\boldsymbol{Y}^T = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$.

4. Let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ be a matrix with $m \geq n$ and full column rank, i.e., $\text{rank}(\boldsymbol{A}) = n$. A *Givens rotation* matrix $\boldsymbol{G}_{i,j,\theta} \in \mathbb{R}^{n \times n}$ (for indices $1 \leq i < j \leq n$) is defined by

$$\boldsymbol{G}_{i,j,\theta} = \begin{bmatrix} \boldsymbol{I}_{i-1} & & & & \\ & \cos\theta & & -\sin\theta & \\ & & \boldsymbol{I}_{j-i-1} & & \\ & \sin\theta & & \cos\theta & \\ & & & & \boldsymbol{I}_{n-j} \end{bmatrix},$$

where $\boldsymbol{I}_k$ denotes the $k \times k$ identity matrix, and all unspecified entries are zero. Equivalently, $\boldsymbol{G}_{i,j,\theta}$ is identical to the identity matrix except for the four entries in its ith and jth rows and columns:

$$g_{ii} = \cos\theta, \quad g_{ij} = -\sin\theta, \quad g_{ji} = \sin\theta, \quad g_{jj} = \cos\theta.$$

Answer the following questions:

(a) Prove that $\boldsymbol{G}_{i,j,\theta}$ is an orthogonal matrix, i.e., show that

$$\boldsymbol{G}_{i,j,\theta}^T \boldsymbol{G}_{i,j,\theta} = \boldsymbol{I}_n.$$

(b) Let $\boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$. Prove that for any indices i, j with $1 \le i < j \le n$, there exists an angle θ such that

$$\boldsymbol{G}_{i,j,\theta}\, \boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_{i-1} \\ \sqrt{v_i^2 + v_j^2} \\ v_{i+1} \\ \vdots \\ v_{j-1} \\ 0 \\ v_{j+1} \\ \vdots \\ v_n \end{bmatrix}.$$

In other words, show that the Givens rotation can eliminate the jth entry of $\boldsymbol{v}$ by affecting only the ith and jth components.

(c) An $m \times n$ matrix $\boldsymbol{A}$ is called an *upper Hessenberg matrix* if

$$a_{ij} = 0 \quad \text{whenever } i - j \ge 2.$$

For $m \ge n$, the pattern of nonzero entries in $\boldsymbol{A}$ is depicted as

$$\begin{bmatrix} \times & \times & \cdots & \times \\ \times & \times & \cdots & \times \\ 0 & \times & \ddots & \vdots \\ \vdots & \ddots & \ddots & \times \\ \vdots & \ddots & \ddots & \times \\ \vdots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}.$$

Since $\boldsymbol{A}$ already contains many zeros in its lower triangular portion, the standard Householder QR algorithm can be optimized. Develop an algorithm using Givens rotations to compute the QR decomposition of an upper Hessenberg matrix $\boldsymbol{A}$. Your algorithm should achieve a computational cost of $O(n^2)$ for the reduced (economy-size) QR factorization and $O(mn)$ for the full QR factorization.